SIAM J. Optim.
Vol. 10, No. 2, pp. 359–383

ON LOCAL SOLUTIONS OF THE CELIS–DENNIS–TAPIA SUBPROBLEM*

XIONGDA CHEN† AND YA-XIANG YUAN†

Abstract. We discuss the distribution of the local solutions of the Celis–Dennis–Tapia (CDT) subproblem, which appears in some trust region algorithms for nonlinear optimization. We also give some examples to show the differences between the CDT subproblem and the single-ball-constraint subproblem. These results show that the complexity of the CDT subproblem does not depend on the complexity of the structure of the dual plane. Thus they provide the possibility to search for the global minimizer in the dual plane.

Key words. trust region subproblem, local solutions, optimality conditions

AMS subject classifications. 65K10, 90C20

PII. S1052623498335018

1. Introduction.

In this paper, we study some theoretical properties of local solutions to the following minimization problem with a quadratic objective and two quadratic constraints:

$$\min_{d\in\mathcal{R}^n} \Phi(d) = \frac{1}{2}d^TBd + g^Td \tag{1.1}$$

subject to

$$\|d\| \le \Delta, \tag{1.2}$$

$$\|A^Td + c\| \le \xi, \tag{1.3}$$

where $g \in \mathcal{R}^n$, $B \in \mathcal{R}^{n\times n}$, $A \in \mathcal{R}^{n\times m}$, $c \in \mathcal{R}^m$, $\Delta > 0$, $\xi \ge 0$, and B is a symmetric matrix. Throughout this paper, the norm $\|\cdot\|$ is the 2-norm. Problem (1.1)–(1.3) is a subproblem of some trust region algorithms for nonlinear programming (see Celis, Dennis, and Tapia [2] and Powell and Yuan [15]), and it is often called the CDT subproblem.

As an important application, the CDT subproblem was used as an inner iteration in the algorithm given by Powell and Yuan [15], whose superlinear convergence property is obtained under certain conditions. However, for general B and A, there is still no satisfactory method with which to find the global solution of problem (1.1)–(1.3) which is required in some trust region algorithms.

The properties of the CDT subproblem have been studied; see Yuan [16] and Peng and Yuan [13] for its extension. Under some additional assumptions, some algorithms have been given to solve it. For example, under the assumption that B is positive definite, different kinds of algorithm are presented by Ecker and Niemi [6], Mehrotra and Sun [12], Phan-Huy-Hao [14], Yuan [17], and Zhang [18]. Instead of the above

*Received by the editors March 6, 1998; accepted for publication (in revised form) March 1, 1999; published electronically December 22, 1999. This work was partially supported by Chinese NSF grants 19525101 and 19731001 and by State Key project 96-221-04-02-02.
http://www.siam.org/journals/siopt/10-2/33501.html
†State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific and Engineering Computing, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, People's Republic of China (cxd@lsec.cc.ac.cn, yyx@lsec.cc.ac.cn).

assumption, under the assumption that $A = I$ and B is semidefinite an algorithm is given by Heinkenschloss [9], which is modified by Chen [3]. A global algorithm for the case $A = I$ and general symmetric B is given by Martínez and Santos [11].

Some approximate methods are given since the CDT problem, used as a subproblem of nonlinear programming algorithms, is needed to obtain a sufficient descent feasible point instead of the global minimizer. See El-Alem and Tapia [7] and Fu, Luo, and Ye [8] for algorithms based on approximations of the feasible region. Byrd and Schnabel [1] and Dennis and Williamson [4] solve the CDT subproblem in the two-dimensional subspace $\{g, (A^T)^+c\}$. These alternative CDT subproblems work in some nonlinear programming to some extent—for example, in the so-called PNCDT method of El-Alem and Tapia [7]. However, it is not clear how to compute the global solution of the CDT problem efficiently.

As a subproblem, ξ can be chosen such that problem (1.2)–(1.3) has feasible points; see Dennis, El-Alem, and Maciel [5] and references therein. If the CDT subproblem has no interior point, it is deduced to a simple case which is discussed in Yuan [16]. So we assume that problem (1.2)–(1.3) has interior points and we do not discuss the choice of ξ.

The rest of this paper is organized as follows. Some basic results are restated in section 2. The structure of the dual plane of the CDT problem is investigated in section 3. The dual function of the CDT problem is extended to a closed region, and some properties of the extend dual function are given, in section 4. Sections 5 and 6 are the main part of this paper. The global minimizer of the CDT problem is divided into three cases. By defining a "related region," we prove that the Lagrangian multipliers corresponding to the global minimizer locate in the related region if the maximizer of the dual function does not correspond to a global minimizer. Then we find the smallest related region. The differences between the local minimizers of the trust region subproblem and those of the CDT problem are presented in section 7. Also shown is that the Lagrangian multipliers of the CDT problem corresponding to local minimizers are permuted in the way of the connected branches of the region where the Hessian has exactly one negative eigenvalue. Conclusions and some possible ways to solve the CDT problem are presented in section 8.

2. Some basic results. In this section, we restate some fundamental results of the CDT problem. For their proofs, see Yuan [16].

THEOREM 2.1. *Let d^* be a global solution of the problem* (1.1)–(1.3). *Assume that $\xi > \min_{\|d\| \le \Delta} \|A^T d + c\|$. Then there exist nonnegative constants λ, μ such that*

$$(B + \lambda I + \mu AA^T)d^* = -(g + \mu Ac), \tag{2.1}$$

where λ and μ satisfy the complementarity conditions

$$\lambda(\Delta - \|d^*\|) = 0, \tag{2.2}$$

$$\mu(\xi - \|A^T d^* + c\|) = 0. \tag{2.3}$$

Furthermore, the matrix

$$H(\lambda, \mu) = B + \lambda I + \mu AA^T \tag{2.4}$$

has at most one negative eigenvalue if the multipliers λ and μ are unique.

To say that $H(\lambda,\mu)$ has one negative eigenvalue means that the negative eigenvalue of $H(\lambda,\mu)$ is a single eigenvalue. For the case that the multipliers λ and μ are not unique, we have the following result.

THEOREM 2.2. *Assume that the conditions of Theorem* 2.1 *hold. Then there exists* $(\lambda,\mu)\in\Omega$ *such that the matrix* (2.4) *has at most one negative eigenvalue, where* Ω *is the set of Lagrangian multipliers.*

We have the following sufficient optimality condition for a global minimizer of problem (1.1)–(1.3).

THEOREM 2.3. *If* d^* *is a feasible point of* (1.2)–(1.3)*, if there are two multipliers* λ *and* μ *such that* (2.1)–(2.3) *hold, and if the matrix* (2.4) *is positive semidefinite, then* d^* *is a global solution of the problem* (1.1)–(1.3).

3. Structure of dual plane. A dual algorithm is given in Yuan [17] for solving subproblem (1.1)–(1.3) with B positive definite, based on the equivalent problem:

$$\min_{d\in\mathcal{R}^n}\Phi(d)=\frac{1}{2}d^TBd+g^Td \tag{3.1}$$

subject to

$$\|d\|^2\le\Delta^2, \tag{3.2}$$

$$\|A^Td+c\|^2\le\xi^2. \tag{3.3}$$

Similar to the single-ball constrained trust region subproblem, the CDT subproblem may be hard when the Hessian of the Lagrangian is positive semidefinite but not positive definite. Furthermore, the Hessian at the global solution may have one negative eigenvalue (see Theorem 2.1). The dual problem for (3.1)–(3.3) can also be defined when the Hessian of the Lagrangian is singular. This will be discussed in the next section.

First, we consider the case in which the Hessian of the Lagrangian is nonsingular. The Hessian of the Lagrangian is (2.4), where $\lambda\ge0$, $\mu\ge0$ are the Lagrangian multipliers of problem (3.1)–(3.3), and they are also the dual variables. Using the notations of Yuan [17], we define the vector

$$d(\lambda,\mu)=-H(\lambda,\mu)^{-1}(g+\mu Ac), \tag{3.4}$$

which satisfies the first equation of the well-known KKT system (2.1)–(2.3) of problem (3.1)–(3.3). We also define the Lagrangian dual function of problem (3.1)–(3.3) as

$$\Psi(\lambda,\mu)=\Phi(d(\lambda,\mu))+\frac{\lambda}{2}(\|d(\lambda,\mu)\|^2-\Delta^2)+\frac{\mu}{2}(\|A^Td(\lambda,\mu)+c\|^2-\xi^2) \tag{3.5}$$

and the region

$$\Omega_0=\{(\lambda,\mu)\in\mathcal{R}^2_+\mid H(\lambda,\mu)\text{ is positive semidefinite}\}, \tag{3.6}$$

where $d(\lambda,\mu)$ is defined by (3.4) and $\mathcal{R}^2_+=\{\lambda\ge0,\mu\ge0\}$. Direct calculations show that

$$\nabla\Psi(\lambda,\mu)=\frac{1}{2}\begin{pmatrix}\|d(\lambda,\mu)\|^2-\Delta^2\\ \|A^Td(\lambda,\mu)+c\|^2-\xi^2\end{pmatrix} \tag{3.7}$$

and

$$\nabla^2\Psi(\lambda,\mu) = -\begin{pmatrix} d(\lambda,\mu)^T H(\lambda,\mu)^{-1} d(\lambda,\mu) & d(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) \\ d(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) & y(\lambda,\mu)^T H(\lambda,\mu)^{-1} y(\lambda,\mu) \end{pmatrix}, \quad (3.8)$$

where $y(\lambda,\mu)$ is the vector

$$y(\lambda,\mu) = A(A^T d(\lambda,\mu) + c). \quad (3.9)$$

In order to study the dual function, we define the region

$$\Omega(\varepsilon) = \{(\lambda,\mu) \in \Omega_0 \mid \; \mathrm{dist}((\lambda,\mu), \partial\Omega_0) \geq \varepsilon\}, \quad (3.10)$$

where $\varepsilon > 0$, $\mathrm{dist}(\cdot,\cdot)$ is the 2-norm distance function, and $\partial\Omega$ denotes the boundary of a region Ω. It is easy to see that Ω_0 and $\Omega(\varepsilon)$ are convex sets. First we show a property of the Hessian on $\Omega(\varepsilon)$.

LEMMA 3.1. *For any* $(\lambda,\mu) \in \Omega(\varepsilon)$, $H(\lambda,\mu)$ *is positive definite.*

Proof. Define $\mathcal{B}_\eta$ to be the Euclidean ball in R^2 with radius η and

$$X \oplus Y = \{x + y \mid x \in X, y \in Y\} \quad (3.11)$$

for two sets X and Y. We have, for any $(\lambda,\mu) \in \Omega(\varepsilon)$,

$$(\lambda,\mu) \oplus \mathcal{B}_{\frac{\varepsilon}{2}} \subset \Omega(\varepsilon) \oplus \mathcal{B}_{\frac{\varepsilon}{2}} \subset \Omega\left(\frac{\varepsilon}{2}\right), \quad (3.12)$$

which implies that $(\lambda - \frac{\varepsilon}{2}, \mu) \in \Omega(\frac{\varepsilon}{2})$ and $H(\lambda - \frac{\varepsilon}{2}, \mu)$ is positive semidefinite. Therefore,

$$H(\lambda,\mu) = H\left(\lambda - \frac{\varepsilon}{2}, \mu\right) + \frac{\varepsilon}{2} I \quad (3.13)$$

is positive definite. □

From the above lemma, the dual function $\Psi(\lambda,\mu)$, its gradient, and its Hessian are well defined in $\mathrm{int}\Omega_0$, the interior of Ω_0, and the region $\Omega(\varepsilon)$ for any $\varepsilon > 0$. Thus by the concavity of $\Psi(\lambda,\mu)$ (from (3.8)), we can obtain the maxima of $\Psi(\lambda,\mu)$ on $\Omega(\varepsilon)$. If there is an $\varepsilon > 0$ such that

$$(\lambda_+,\mu_+) = \arg \max_{(\lambda,\mu)\in\Omega(\varepsilon)} \Psi(\lambda,\mu) \in \; \mathrm{int}\Omega(\varepsilon), \quad (3.14)$$

then $H(\lambda_+,\mu_+)$ is positive definite and we can prove the following theorem. Actually, this theorem holds for any positive definite $H(\lambda_+,\mu_+)$ if $(\lambda_+,\mu_+) \in \Omega_0$.

THEOREM 3.2. *Suppose that* (λ_+,μ_+) *is any point defined by* (3.14). *Then the global solution of* (3.1)–(3.3) *is* $d(\lambda_+,\mu_+)$ *given by* (3.4).

Proof. Because $H(\lambda_+,\mu_+)$ is positive definite, $d(\lambda_+,\mu_+)$ is well defined by (3.4), and (λ_+,μ_+) is a local maximizer of $\Psi(\lambda,\mu)$ in $\mathcal{R}^2_+$. Therefore, we have that

$$\nabla\Psi(\lambda_+,\mu_+) \leq 0 \quad (3.15)$$

and

$$(\lambda_+,\mu_+)^T \nabla\Psi(\lambda_+,\mu_+) = 0. \quad (3.16)$$

This shows that $d(\lambda_+,\mu_+)$ is the global solution of (3.1)–(3.3) and (λ_+,μ_+) are the corresponding Lagrangian multipliers. □

Since $\arg\max \Psi(\lambda, \mu)$ is a convex set, it includes an interior point of Ω_0 if it strictly includes a segment. If $\arg\max \Psi(\lambda, \mu)$ on $\text{int}\Omega_0$ is a segment and there is a point of this segment in the interior of Ω_0, there exists a sufficient small $\epsilon >$ such that (3.14) holds. Thus it follows from Theorem 3.2 that there exists a global solution of (3.1)–(3.3) with a positive definite Hessian of the Lagrangian. Otherwise, the Hessian of the Lagrangian might not be positive semidefinite. By the theorems stated in section 2, the Hessian $H(\lambda, \mu)$ at the global solution of (3.1)–(3.3) has at most one negative eigenvalue, and the corresponding Lagrangian multipliers (λ, μ) may locate in the region

$$\Omega_1 = \{(\lambda, \mu) \in \mathcal{R}^2_+ \mid H(\lambda, \mu) \text{ has one negative eigenvalue}\}. \tag{3.17}$$

Next we investigate the structure of Ω_1. Let

$$\Omega_1 = \bigcup_{k \in \mathcal{K}} \Omega_{1k}, \tag{3.18}$$

where $\mathcal{K}$ is an index set; $\Omega_{1k}, k \in \mathcal{K}$, are connected sets; and Ω_{1k} and Ω_{1j} are disconnected for any $k \neq j$, $k, j \in \mathcal{K}$. Whether $\mathcal{K}$ is a finite index set makes no difference to our discussions by the location theorem given in section 5. By defining

$$u_{\lambda k} = \sup\nolimits_{(\lambda,\mu) \in \Omega_{1k}} \lambda, \tag{3.19}$$

$$u_{\mu k} = \sup\nolimits_{(\lambda,\mu) \in \Omega_{1k}} \mu, \tag{3.20}$$

$$l_{\lambda k} = \inf\nolimits_{(\lambda,\mu) \in \Omega_{1k}} \lambda, \tag{3.21}$$

and

$$l_{\mu k} = \inf\nolimits_{(\lambda,\mu) \in \Omega_{1k}} \mu, \tag{3.22}$$

we have the following lemma,

LEMMA 3.3. *If there are* $(\lambda_1, \mu_1) \in \Omega_{1k}$, $(\lambda_2, \mu_2) \in \Omega_{1j}$, *where* $k \neq j$ *and* $\lambda_1 > \lambda_2$, *then we have*

$$l_{\lambda k} \geq u_{\lambda j}, \tag{3.23}$$

$$u_{\mu k} \leq l_{\mu j}. \tag{3.24}$$

Moreover, (3.23) *and* (3.24) *are both equalities or strict inequalities.*

Proof. It is easy to show that the set $\{\lambda \mid \exists(\lambda, \mu) \in \Omega_{1k}\}$ is a segment for any fixed μ, as is $\{\mu \mid \exists(\lambda, \mu) \in \Omega_{1k}\}$ for any fixed λ. If $l_{\lambda k} < u_{\lambda j}$, there is a $\lambda_0 \in (l_{\lambda k}, u_{\lambda j})$. By the above definitions, there are $\mu_k^0, \mu_j^0 \in R$ such that $(\lambda_0, \mu_k^0) \in \Omega_{1k}$ and $(\lambda_0, \mu_j^0) \in \Omega_{1j}$. Without loss of generality, let $\mu_k^0 < \mu_j^0$. Denote $\rho_i(B)$ as the ith eigenvalue of B. We have

$$\rho_n(H(\lambda_0, \mu_k^0)) \leq \rho_n(H(\lambda_0, \mu)) \leq \rho_n(H(\lambda_0, \mu_j^0)) < 0, \tag{3.25}$$

$$0 \leq \rho_{n-1}(H(\lambda_0, \mu_k^0)) \leq \rho_{n-1}(H(\lambda_0, \mu)) \leq \rho_{n-1}(H(\lambda_0, \mu_j^0)) \tag{3.26}$$

for all $\mu \in (\mu_k^0, \mu_j^0)$. Thus $(\lambda_0, \mu) \in \Omega_1$ for all $\mu \in (\mu_k^0, \mu_j^0)$, which implies that Ω_{1k} and Ω_{1j} are connected. The contradiction proves (3.23), and (3.24) can be proved similarly.

If $l_{\lambda k} > u_{\lambda j}$ and $u_{\mu k} = l_{\mu j}$, $H(u_{\lambda j}, l_{\mu j})$ is positive semidefinite and $H(l_{\lambda k}, u_{\mu k}) = H(u_{\lambda j}, l_{\mu j}) + (l_{\lambda k} - u_{\lambda j})I$ is positive definite, contradicting the definitions (3.19)–(3.22).

Suppose that $l_{\lambda k} = u_{\lambda j}$ and $u_{\mu k} < l_{\mu j}$. Since for any sufficiently small $\varepsilon > 0$,

$$(l_{\lambda k} - \varepsilon, \mu) \notin \Omega_{1l} \tag{3.27}$$

for any $\mu \in (u_{\mu k}, l_{\mu j})$ and $l \in \mathcal{K}$, $H(l_{\lambda k} - \varepsilon, \mu)$ has at least two negative eigenvalues. It also can be shown that $H(l_{\lambda k} + \varepsilon, \mu)$ is positive definite. Thus taking $\varepsilon \to 0$, we obtain that $H(l_{\lambda k}, \mu)$ has zero eigenvalues with multiplicity at least two for $\mu \in (u_{\mu k}, l_{\mu j})$.

Since $\det(H(l_{\lambda k}, \mu))$ is a polynomial with zeros with multiplicity at least two for $\mu \in (u_{\mu k}, l_{\mu j})$, $\det(H(l_{\lambda k}, \mu))$ has zeros with multiplicity at least two for all $\mu \geq 0$, which means that the dimension of $\mathrm{Null}(H(l_{\lambda k}, \mu))$ is no less than two. Therefore, $H(u_{\lambda j} - \varepsilon, \mu)$ has at least two negative eigenvalues for all $\varepsilon > 0$ and $\mu > l_{\mu j}$, contradicting the definition of Ω_{1j}. □

In the following, we denote $\Omega_{1k} \succ \Omega_{1j}$ if (3.23)–(3.24) hold. Moreover, from the above proof, there is at most one segment in the intersection of any positive-slope straight line in $\mathcal{R}_+^2$ and any connected branch of Ω_1.

DEFINITION 3.4. *Two connected branches $\Omega_{1k} \succ \Omega_{1j}$ of Ω_1 are called consecutive connected branches if there is no other connected branch Ω_{1l} of Ω_1 such that*

$$\Omega_{1k} \succ \Omega_{1l} \succ \Omega_{1j}, \tag{3.28}$$

and they are called two adjoint connected branches if (3.23) *and* (3.24) *hold as equalities.*

The following lemma tells us the more detailed structure of the border of Ω_0.

LEMMA 3.5. *For any two adjoint connected branches Ω_{1k} and Ω_{1j} of Ω_1, $\Omega_{1k} \succ \Omega_{1j}$, $(l_{\lambda k}, u_{\mu k})$ is an extreme point of the convex set Ω_0.*

Proof. It suffices to prove that $\partial\Omega_0 \bigcap \partial(\Omega_{1k} \bigcup \Omega_{1j})$ is not a segment in any neighborhood of $(l_{\lambda k}, u_{\mu k})$. We prove this lemma by contradiction. Suppose

$$(l_{\lambda k}, u_{\mu k}) \in \overline{\Omega_{1k}} \bigcap \overline{\Omega_{1j}}, \tag{3.29}$$

$$(\bar{\lambda}_k, \bar{\mu}_k) \in \partial\Omega_0 \bigcap \overline{\Omega}_{1k}, \tag{3.30}$$

and

$$(\bar{\lambda}_j, \bar{\mu}_j) \in \partial\Omega_0 \bigcap \overline{\Omega}_{1j} \tag{3.31}$$

are in a straight line. Then there exists $0 < \delta < 1$, such that

$$H(l_{\lambda k}, u_{\mu k}) = \delta H(\bar{\lambda}_k, \bar{\mu}_k) + (1 - \delta) H(\bar{\lambda}_j, \bar{\mu}_j). \tag{3.32}$$

Since $(\bar{\lambda}_k - \varepsilon, \bar{\mu}_k - \varepsilon) \in \Omega_{1k}$ for $\varepsilon > 0$ sufficient small, $H(\bar{\lambda}_k - \varepsilon, \bar{\mu}_k - \varepsilon)$ has exactly one negative eigenvalue. Taking $\varepsilon \to 0+$, we can show that $H(\bar{\lambda}_k, \bar{\mu}_k)$ is positive semidefinite and has one multiple zero eigenvalue. The fact is also true for $H(\bar{\lambda}_j, \bar{\mu}_j)$. Suppose that

$$z_1 \in \ \mathrm{Null}(H(\bar{\lambda}_k, \bar{\mu}_k)) \tag{3.33}$$

and

$$z_2 \in \text{ Null}(H(\bar\lambda_j, \bar\mu_j)), \tag{3.34}$$

where Null($\cdot$) denotes the null space of a matrix. Since $H(l_{\lambda k}, u_{\mu k})$ is positive semidefinite, for any $\varepsilon > 0$, $H((1+\varepsilon)l_{\lambda k}, (1+\varepsilon)u_{\mu k})$ is positive definite. For any Ω_{1l} satisfying $\Omega_{1l} \succeq \Omega_{1k}$,

$$l_{\lambda l} > (1-\varepsilon)l_{\lambda k}, \tag{3.35}$$

while for any Ω_{1l} satisfying $\Omega_{1j} \succeq \Omega_{1l}$,

$$u_{\mu l} > (1-\varepsilon)u_{\mu k}, \tag{3.36}$$

for any $\varepsilon > 0$. So for any $\varepsilon > 0$, $((1-\varepsilon)l_{\lambda k}, (1-\varepsilon)u_{\mu k}) \notin \Omega_{1j}$ for all j, and hence $H((1-\varepsilon)l_{\lambda k}, (1-\varepsilon)u_{\mu k})$ has at least two negative eigenvalues. Let $\varepsilon \to 0+$; we can see that $H(l_{\lambda k}, u_{\mu k})$ is positive semidefinite and has zero eigenvalues with multiplicity at least two.

Suppose $\text{span}\{v_1, v_2\} \subset \text{Null}(H(l_{\lambda k}, u_{\mu k}))$. Then for all $z \in \text{span}\{v_1, v_2\}$,

$$z^T H(l_{\lambda k}, u_{\mu k})z = z^T(\delta H(\bar\lambda_k, \bar\mu_k) + (1-\delta)H(\bar\lambda_j, \bar\mu_j))z = 0. \tag{3.37}$$

The above equality implies that $z^T H(\bar\lambda_k, \bar\mu_k)z = 0$ and $z^T H(\bar\lambda_j, \bar\mu_j)z = 0$, so $z_1 = z_2$, and

$$\text{Null}(H(\bar\lambda_k, \bar\mu_k)) = \text{ Null}(H(\bar\lambda_j, \bar\mu_j)), \tag{3.38}$$

and the dimension of $\text{Null}(H(l_{\lambda k}, u_{\mu k}))$ is equal to 1, contradicting the fact that $H(l_{\lambda k}, u_{\mu k})$ has zero eigenvalues with multiplicity two. ∎

4. Definitions on the boundary. In this section, we deal with the boundary of Ω_0. First we define the dual function on the boundary of Ω_0 based on the definitions in $\text{int}\Omega_0$. Assuming that $(\bar\lambda, \bar\mu) \in \partial\Omega_0$ with $H(\bar\lambda, \bar\mu)$ singular, we define

$$\Psi(\bar\lambda, \bar\mu) = \lim_{\varepsilon \to 0+} \Psi(\bar\lambda + \varepsilon, \bar\mu). \tag{4.1}$$

In summary, $\Psi(\bar\lambda, \bar\mu)$ is the right limit of $\Psi(\cdot,\cdot)$ at the point $(\bar\lambda, \bar\mu)$ along the line $\mu = \bar\mu$. Because $\Psi(\cdot,\cdot)$ is continuous in $\text{int}\Omega_0$, definition (4.1) also holds for the interior point of Ω_0.

LEMMA 4.1. *Equation* (4.1) *is well defined.*

Proof. Equation (3.5) can be rewritten as

$$\Psi(\lambda, \mu) = -\frac{1}{2}(g + \mu Ac)^T H(\lambda, \mu)^{-1}(g + \mu Ac) - \frac{\lambda}{2}\Delta^2 - \frac{\mu}{2}(\xi^2 - \|c\|^2) \tag{4.2}$$

when $(\lambda, \mu) \in \text{int}\Omega_0$. Therefore,

$$\Psi(\lambda, \mu) \le -\frac{\lambda}{2}\Delta^2 - \frac{\mu}{2}(\xi^2 - \|c\|^2), \tag{4.3}$$

which shows that $\Psi(\lambda, \mu)$ is locally upper bounded in $\text{int}\Omega_0$.

If (2.1) is inconsistent, the right-hand side of (4.1) is $-\infty$. Otherwise, suppose that $g + \bar\mu Ac = H(\bar\lambda, \bar\mu)v$ for some $v \in \mathcal{R}^n$. It is easy to see that for any positive semidefinite matrix A,

$$\lim_{\varepsilon \to 0+} (A + \varepsilon I)^{-1} A = A^+ A. \tag{4.4}$$

Therefore, the following limit exists:

$$
\begin{aligned}
& \lim_{\varepsilon\to 0+} d(\bar\lambda+\varepsilon,\bar\mu) \\
= \;& \lim_{\varepsilon\to 0+} -H(\bar\lambda+\varepsilon,\bar\mu)^{-1}(g+\bar\mu Ac) \\
= \;& \lim_{\varepsilon\to 0+} -H(\bar\lambda+\varepsilon,\bar\mu)^{-1}H(\bar\lambda,\bar\mu)\bar g \\
= \;& -H(\bar\lambda,\bar\mu)^{+}H(\bar\lambda,\bar\mu)\bar g.
\end{aligned} \tag{4.5}
$$

$d(\bar\lambda+\varepsilon,\bar\mu)$ is uniformly bounded when $\varepsilon\to 0+$. Suppose there are two sequences $\{\lambda_{1k}\}$ and $\{\lambda_{2k}\}$ such that

$$\lim_{\lambda_{1k}\to\bar\lambda+}\Psi(\lambda_{1k},\bar\mu)\neq\lim_{\lambda_{2k}\to\bar\lambda+}\Psi(\lambda_{2k},\bar\mu). \tag{4.6}$$

By the mean value theorem, we have

$$\Psi(\lambda_{1k},\bar\mu)-\Psi(\lambda_{2k},\bar\mu)=\frac{1}{2}(\lambda_{1k}-\lambda_{2k})\|d(\lambda_m,\bar\mu)\|. \tag{4.7}$$

The right-hand side of (4.7) vanishes by the boundedness of $d(\cdot,\bar\mu)$, contradicting (4.6). This completes our proof. ∎

Thus $\Psi(\lambda,\mu)$ is defined on the closed set Ω_0 and can take a finite value or $-\infty$, but not $+\infty$. However,

$$\max_{(\lambda,\mu)\in\Omega_0}\Psi(\lambda,\mu) \tag{4.8}$$

may be $+\infty$, and

$$\arg\max_{(\lambda,\mu)\in\Omega_0}\Psi(\lambda,\mu) \tag{4.9}$$

may lie at the infinity on the dual space $\mathcal{R}^2_+$. In section 5, we will see that this case can be handled in the same way as the finite case. Since $\Psi(\cdot,\cdot)$ is well defined on the closed set Ω_0, we define the set

$$S=\left\{(\lambda,\mu)\in\Omega_0 \mid (\lambda,\mu)=\arg\max_{\Omega_0}\Psi(\lambda,\mu)\right\}. \tag{4.10}$$

Because $\Psi(\lambda,\mu)$ is concave in $\mathrm{int}\Omega_0$ and also concave on Ω_0 by Lemma 4.2 given below, its maxima point set may be a segment on the $\partial\Omega_0$. Then Lemma 3.5 implies that this segment belongs to only one connected branch.

LEMMA 4.2. *$\Psi(\lambda,\mu)$ is a concave function on the closed set Ω_0.*

Proof. Let (λ_1,μ_1) and (λ_2,μ_2) be two points in Ω_0. Then we have that $H(\lambda_1+\varepsilon,\mu_1)$ and $H(\lambda_2+\varepsilon,\mu_2)$ are positive definite for all $\varepsilon>0$, and hence so is $H(\frac{\lambda_1+\lambda_2+2\varepsilon}{2},\frac{\mu_1+\mu_2}{2})$. Since $\Psi(\cdot,\cdot)$ is concave in $\mathrm{int}\Omega_0$,

$$\Psi\left(\frac{\lambda_1+\lambda_2+2\varepsilon}{2},\frac{\mu_1+\mu_2}{2}\right)\geq\frac{1}{2}(\Psi(\lambda_1+\varepsilon,\mu_1)+\Psi(\lambda_2+\varepsilon,\mu_2)). \tag{4.11}$$

Taking limits on both sides of the above inequality we deduced that $\Psi(\lambda,\mu)$ is concave on Ω_0. ∎

Assuming $(\bar\lambda,\bar\mu)\in\partial\Omega_0$, we define

$$d(\bar\lambda,\bar\mu)=\lim_{\varepsilon\to 0+}d(\bar\lambda+\varepsilon,\bar\mu). \tag{4.12}$$

If (2.1) is inconsistent for $(\bar\lambda, \bar\mu)$, the right-hand side of (4.1) goes to $-\infty$. In this case $d(\bar\lambda, \bar\mu)$ is undefined in (4.12). Suppose (2.1) is consistent at $(\bar\lambda, \bar\mu)$. Then we can choose, for convenience, the minimum norm least square solution of (2.1). In the following, we will see that it is important that the limit (4.12) satisfies the property stated in Lemma 4.3 instead of the definition (4.12) itself.

LEMMA 4.3. *Assuming that* (2.1) *holds at* $(\bar\lambda, \bar\mu) \in \partial\Omega_0$, *we have the following property:*

$$\lim_{\varepsilon\to 0+} H(\bar\lambda, \bar\mu) d(\bar\lambda + \varepsilon, \bar\mu) = -(g + \bar\mu A c). \tag{4.13}$$

Proof. First we have

$$\begin{aligned} H(\bar\lambda, \bar\mu) d(\bar\lambda + \varepsilon, \bar\mu) &= -H(\bar\lambda, \bar\mu) H(\bar\lambda + \varepsilon, \bar\mu)^{-1}(g + \bar\mu A c) \\ &= -H(\bar\lambda, \bar\mu) \left(H(\bar\lambda, \bar\mu) + \varepsilon I\right)^{-1}(g + \bar\mu A c). \end{aligned} \tag{4.14}$$

It is easy to see that for any positive semidefinite matrix A,

$$\lim_{\varepsilon\to 0+} A(A + \varepsilon I)^{-1} = AA^+, \tag{4.15}$$

where A^+ is the Moore–Penrose generalized inverse of A. Thus, it follows from (4.14) and (4.15) that

$$\begin{aligned} \lim_{\varepsilon\to 0+} H(\bar\lambda, \bar\mu) d(\bar\lambda + \varepsilon, \bar\mu) &= -H(\bar\lambda, \bar\mu) H(\bar\lambda, \bar\mu)^+(g + \bar\mu A c) \\ &= -(g + \bar\mu A c), \end{aligned} \tag{4.16}$$

which gives (4.13). ∎

Since we have

$$\begin{aligned} & \lim_{\varepsilon\to 0+} (g + \bar\mu A c)^T H(\bar\lambda + \varepsilon, \bar\mu)^{-1}(g + \bar\mu A c) \\ = \;& \lim_{\varepsilon\to 0+} (g + \bar\mu A c)^T H(\bar\lambda + \varepsilon, \bar\mu)^{-1} H(\bar\lambda + \varepsilon, \bar\mu) H(\bar\lambda + \varepsilon, \bar\mu)^{-1}(g + \bar\mu A c) \\ = \;& \lim_{\varepsilon\to 0+} (g + \bar\mu A c)^T H(\bar\lambda + \varepsilon, \bar\mu)^{+} H(\bar\lambda + \varepsilon, \bar\mu) H(\bar\lambda + \varepsilon, \bar\mu)^{+}(g + \bar\mu A c) \\ = \;& (g + \bar\mu A c)^T H(\bar\lambda, \bar\mu)^+(g + \bar\mu A c), \end{aligned} \tag{4.17}$$

the following result follows from our extended definitions given in (4.1).

LEMMA 4.4. *For* $(\bar\lambda, \bar\mu) \in \Omega_0$,

$$\Psi(\bar\lambda, \bar\mu) = \begin{cases} -\infty & \textit{if } (2.1) \textit{ is inconsistent,} \\ -\frac{1}{2}(g + \bar\mu A c)^T H(\bar\lambda, \bar\mu)^+(g + \bar\mu A c) - \frac{\lambda}{2}\Delta^2 - \frac{\mu}{2}(\xi^2 - \|c\|^2) & \\ & \textit{otherwise} \end{cases} \tag{4.18}$$

and

$$d(\bar\lambda, \bar\mu) = \begin{cases} \textit{undefined} & \textit{if } (2.1) \textit{ is inconsistent,} \\ -H(\bar\lambda, \bar\mu)^+(g + \bar\mu A c) & \textit{otherwise.} \end{cases} \tag{4.19}$$

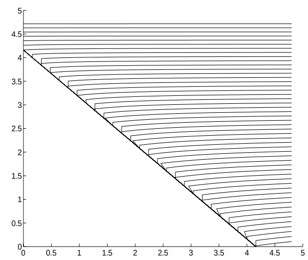

FIG. 5.1. *Example* 5.1.

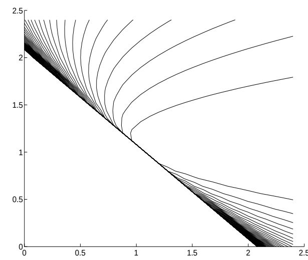

FIG. 5.2. *Example* 5.2.

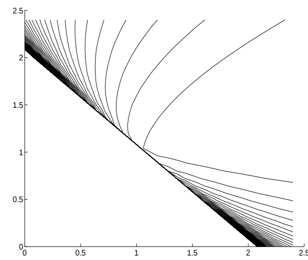

FIG. 5.3. *Example* 5.3.

5. Location of global solution. In this section, we study the relations between the set S defined by (4.10) and the Lagrangian multipliers $(\bar{\lambda}, \bar{\mu})$ at the global solutions of problem (1.1)–(1.3). First, we consider the following cases:

- There exists a $(\lambda_+, \mu_+) \in S$ satisfying (3.14). So, $d(\lambda_+, \mu_+)$ is a global solution of problem (3.1)–(3.3) due to Theorem 3.2. In this case, S may be a singleton or a segment.
- A segment of S lies in $\partial\Omega_{1k}$ for some $k \in \mathcal{K}$. (k is unique due to Lemma 3.5.) Theorem 5.4, given below, states that we have obtained the desired global solution.
- S is a singleton such that $S \subset \partial\Omega_1$. For this case, in this section we give the locating branches in which the global solution lies, which generally includes two or three connected branches of Ω_1. In this case, the Hessian of any global solution of problem (3.1)–(3.3) might not be positive semidefinite. And we still cannot determine in which connected branch the global solution lies. This is the hard case of subproblem (1.1)–(1.3).

The following examples show the last two cases.

Example 5.1. *A segment maxima of the dual function.* Let

$$B = \text{diag}(-4, -2), A = \text{diag}(1, 2), g = (0, 4)', c = (0, 3)', \Delta = 3, \xi = \sqrt{6}.$$

Then $d = (\pm\sqrt{5}, -2)'$ with the Lagrangian multiplier $(4 - \mu, \mu)$ and the Hessian of Lagrangian $\text{diag}(0, 2 + 3\mu)$, where $\mu \in [0, 4]$.

The following two examples show the hardness of the last case.

Example 5.2. Let

$$B = \text{diag}(-2, 2), A = \text{diag}(1, 1), g = (2, 0)', c = (-2, 0)', \Delta = 2, \xi = 1.$$

Then $d = (2, 0)'$ with the Lagrangian multiplier $(1, 0)$ and the Hessian of Lagrangian $\text{diag}(-1, 3)$.

Example 5.3. Let

$$B = \text{diag}(-1, -2), A = \text{diag}(1, 1), g = (-4, 6)', c = (0, -6)', \Delta = 5, \xi = 5.$$

Then $d = (4, 3)'$ with the Lagrangian multiplier $(1, 1)$ and the Hessian of Lagrangian $\text{diag}(1, 0)$.

The contours of the dual functions of the above examples are given in Figures 5.1, 5.2, and 5.3.

If $S \subset \Omega_{1k}$ is a segment for some $k \in \mathcal{K}$, we already get the solution by adding a null-space-step of the Hessian of the Lagrangian by Theorem 5.4.

THEOREM 5.4. *If $S \subset \partial\Omega_{1k}$ is a segment, then there exists a solution of* (2.1)–(2.3) *where the Hessian of the Lagrangian is positive semidefinite.*

Proof. Let $(\bar\lambda, \bar\mu)$ and $(\hat\lambda, \hat\mu)$ be two different points of S. By the definition of S, we have

$$\lim_{i\to+\infty} \Psi(\bar\lambda_i, \bar\mu) = \max_{\Omega_0} \Psi(\lambda, \mu), \tag{5.1}$$

$$\lim_{i\to+\infty} \Psi(\hat\lambda_i, \hat\mu) = \max_{\Omega_0} \Psi(\lambda, \mu), \tag{5.2}$$

where $\bar\lambda_i - \bar\lambda \to 0+$ and $\hat\lambda_i - \hat\lambda \to 0+$. Since $\Psi(\cdot,\cdot)$ approximates to a constant in a neighborhood of S, the above two relations give that

$$\lim_{i\to+\infty} \begin{pmatrix} \|d(\bar\lambda_i, \bar\mu)\|^2 - \Delta^2 \\ \|A^T d(\bar\lambda_i, \bar\mu) + c\|^2 - \xi^2 \end{pmatrix}^T \begin{pmatrix} \bar\lambda - \hat\lambda \\ \bar\mu - \hat\mu \end{pmatrix} = 0 \tag{5.3}$$

and

$$\lim_{i\to+\infty} d(\bar\lambda_i, \bar\mu)(\bar\lambda_i - \hat\lambda_i) - y(\bar\lambda_i, \bar\mu)(\bar\mu - \hat\mu) = 0. \tag{5.4}$$

Since $S \subset \partial\Omega_{1k}$, $H(\bar\lambda, \bar\mu)$, and $H(\hat\lambda, \hat\mu)$ are singular. By the concavity of $\Psi(\cdot,\cdot)$, we have that $H(\lambda, \mu)$ is singular and positive semidefinite for all (λ, μ) in the segment between $(\bar\lambda, \bar\mu)$ and $(\hat\lambda, \hat\mu)$. For any $0 \neq v \in \mathrm{Null}(H(\bar\lambda, \bar\mu))$ we have that

$$v^T H(\bar\lambda, \bar\mu) v = v^T H(\hat\lambda, \hat\mu) v = 0, \tag{5.5}$$

which implies that

$$\|v\|^2(\bar\lambda - \hat\lambda) + \|A^T v\|^2(\bar\mu - \hat\mu) = 0. \tag{5.6}$$

Therefore,

$$\lim_{i\to+\infty} v^T(d(\bar\lambda_i, \bar\mu)(\bar\lambda_i - \hat\lambda_i) - y(\bar\lambda_i, \bar\mu)(\bar\mu - \hat\mu)) = 0, \tag{5.7}$$

which, together with (5.4) and (5.6), gives that

$$\lim_{i\to+\infty} \begin{pmatrix} \|d(\bar\lambda_i, \bar\mu) + tv\|^2 - \Delta^2 \\ \|A^T(d(\bar\lambda_i, \bar\mu) + tv) + c\|^2 - \xi^2 \end{pmatrix}^T \begin{pmatrix} \bar\lambda - \hat\lambda \\ \bar\mu - \hat\mu \end{pmatrix} = 0 \tag{5.8}$$

for all $t \in \mathcal{R}$. Since the right-hand sides of (5.1) and (5.2) are not $-\infty$, it can be seen that (2.1) is consistent at $(\bar\lambda, \bar\mu)$ and $(\hat\lambda, \hat\mu)$. Thus

$$\lim_{i\to\infty} d(\bar\lambda_i, \bar\mu) = d(\bar\lambda, \bar\mu). \tag{5.9}$$

Since $(\bar\lambda, \bar\mu) \in S$, by the reasons mentioned in Theorem 3.2, we see that $\|d(\bar\lambda, \bar\mu)\| \le \Delta$. Therefore, we can choose t_i such that

$$\lim_{i\to+\infty} \|d(\lambda_i, \mu) + t_i v\| = \Delta. \tag{5.10}$$

Let $\bar{d}$ be any limit point of $\{d(\lambda_i, \mu) + t_i v\}$, so

$$\|\bar{d}\| = \Delta, \tag{5.11}$$

$$\|A^T\bar{d} + c\| = \xi, \tag{5.12}$$

and

$$H(\bar{\lambda}, \bar{\mu})\bar{d} = -(g + \bar{\mu}Ac), \tag{5.13}$$

which implies that $\bar{d}$ is a global solution, $(\bar{\lambda}, \bar{\mu})$ is the corresponding pair of Lagrangian multipliers, and $H(\bar{\lambda}, \bar{\mu})$ is positive semidefinite. ∎

6. Location of global solution: Hard case. We now consider the hard case of problem (1.1)–(1.3), in which S is a singleton on $\partial\Omega_0$ and the Hessian at the global solution may have one negative eigenvalue. In order to determine the region where the solution locates, we introduce the following definition. In all the following discussion we assume that $S = \{(\lambda_+, \mu_+)\} \subset \partial\Omega_0$ is a singleton.

DEFINITION 6.1. *Define two sets*

$$\mathcal{L} = \{\lambda_e < \lambda_+ \mid ri\{\{\lambda = \lambda_e\} \cap \Omega_1\} = \emptyset\} \cup \{0\} \tag{6.1}$$

and

$$\mathcal{M} = \{\mu_e < \mu_+ \mid ri\{\{\mu = \mu_e\} \cap \Omega_1\} = \emptyset\} \cup \{0\}. \tag{6.2}$$

For $\lambda_e \in \mathcal{L}$ and $\mu_e \in \mathcal{M}$, the set

$$\Omega(\lambda_e, \mu_e) = \left(\overline{\Omega}_1 \cup S\right) \cap \{\lambda \geq \lambda_e, \mu \geq \mu_e\} \tag{6.3}$$

is called a related region of S.

Because S might not be in Ω_1 or even $\overline{\Omega}_1$, the term $\overline{\Omega}_1 \cup S$ must occur in (6.3). Here, $ri\{\{\lambda = \lambda_e\} \cap \Omega_1\} = \emptyset$ implies that it is impossible for $H(\lambda_e, \mu)$ to have exactly one negative eigenvalue for any μ. In the latter case, for the λ-direction, we may have the following two cases:

(i) $\lambda_+ > 0$, and for all $0 \leq \lambda_e < \lambda_+$, $ri\{\{\lambda = \lambda_e\} \cap \Omega_1\} \neq \emptyset$;

(ii) $\lambda_+ = 0$.

Similarly, we have two cases for the μ-direction. If $\lambda_e = 0$ or $\mu_e = 0$, the related region is the same as $\overline{\Omega}_1 \cup S$ in the λ-direction or in the μ-direction. Therefore, there exists a Lagrange multiplier at the global solution lies in the related region in the λ-direction or in the μ-direction. Thus, we need only to consider the case when $\lambda_e \neq 0$ and $\mu_e \neq 0$. In this case we have that $\lambda_+ \neq 0$ and $\mu_+ \neq 0$.

First, we need the following lemma to prove our location theorem.

LEMMA 6.2. *Assume that S is a singleton. If the triple (λ^*, μ^*, d^*) satisfies KKT system* (2.1)–(2.3) *with $(\lambda^*, \mu^*) \in \Omega_0$, and if either of the statements*

(i) $\lambda^* \neq 0$,

(ii) $\mu^* \neq 0$ *and* $\det(B + \lambda^* I + \mu AA^T)$ *does not vanish identically for $\mu \geq \mu^*$*

holds, then

$$S = \{(\lambda^*, \mu^*)\}. \tag{6.4}$$

Proof. It suffices to prove, for any fixed $(\lambda, \mu) \in \text{int}\Omega_0$,

$$\Psi(\lambda^*, \mu^*) \geq \Psi(\lambda, \mu). \tag{6.5}$$

We assume, first, that $\lambda^* \neq 0$. For any $\varepsilon > 0$, it is easy to see that d^* is a global solution of the subproblem

$$\min_{d\in\mathcal{R}^n} \bar{\Phi}(d) = \frac{1}{2}d^TBd + \bar{g}^Td \tag{6.6}$$

subject to

$$\|d\| \leq \Delta, \tag{6.7}$$

$$\|A^Td + c\| \leq \xi, \tag{6.8}$$

where $\bar{g} = g - \varepsilon d^*$ and the corresponding pair of Lagrangian multipliers is $(\lambda^* + \varepsilon, \mu^*)$. Define that

$$\bar{d}(\lambda,\mu) = -H(\lambda,\mu)^{-1}(\bar{g} + \mu Ac) \tag{6.9}$$

and

$$\bar{\Psi}(\lambda,\mu) = -\frac{1}{2}(\bar{g} + \mu Ac)^T H(\lambda,\mu)^{-1}(\bar{g} + \mu Ac) - \frac{\lambda}{2}\Delta^2 - \frac{\mu}{2}(\xi^2 - \|c\|^2) \tag{6.10}$$

when $H(\lambda,\mu)$ is positive definite. If $H(\lambda,\mu)$ is singular, $\Psi(\lambda,\mu)$ can be defined as in (4.1). Then we have

$$\nabla\bar{\Psi}(\lambda^* + \varepsilon, \mu^*) \leq 0 \tag{6.11}$$

and

$$(\lambda^* + \varepsilon, \mu^*)^T \nabla\bar{\Psi}(\lambda^* + \varepsilon, \mu^*) = 0. \tag{6.12}$$

Since the Hessian $H(\lambda^* + \varepsilon, \mu^*)$ is positive definite, (6.11) and (6.12) imply that $(\lambda^* + \varepsilon, \mu^*)$ is a stationary point of $\bar{\Psi}(\lambda,\mu)$. Since $\bar{\Psi}(\lambda,\mu)$ is concave in Ω_0, $(\lambda^* + \varepsilon, \mu^*)$ is a global maximizer of $\bar{\Psi}(\lambda,\mu)$ on Ω_0, i.e.,

$$\bar{\Psi}(\lambda^* + \varepsilon, \mu^*) \geq \bar{\Psi}(\lambda,\mu) \text{ for all } (\lambda,\mu) \in \text{ int}\Omega_0. \tag{6.13}$$

For the left-hand side of (6.13), we have

$$\begin{aligned} & \lim_{\varepsilon\to 0+} \bar{\Psi}(\lambda^* + \varepsilon, \mu^*) \\ = \; & \lim_{\varepsilon\to 0+} -\tfrac{1}{2}d^{*T}H(\lambda^* + \varepsilon, \mu^*)d^* - \tfrac{\lambda^*+\varepsilon}{2}\Delta^2 - \tfrac{\mu^*}{2}(\xi^2 - \|c\|^2) \\ = \; & -\tfrac{1}{2}d^{*T}H(\lambda^*, \mu^*)d^* - \tfrac{\lambda^*}{2}\Delta^2 - \tfrac{\mu^*}{2}(\xi^2 - \|c\|^2). \end{aligned} \tag{6.14}$$

Because (λ^*, μ^*, d^*) satisfies the KKT system, it follows that

$$\begin{aligned} & -\tfrac{1}{2}d^{*T}H(\lambda^*, \mu^*)d^* \\ = \; & -\tfrac{1}{2}d^{*T}H(\lambda^*, \mu^*)H^+(\lambda^*, \mu^*)H(\lambda^*, \mu^*)d^* \\ = \; & -\tfrac{1}{2}(g + \mu^*Ac)^T H^+(\lambda^*, \mu^*)(g + \mu^*Ac). \end{aligned} \tag{6.15}$$

Equations (6.14), (6.15), and (4.18) imply that

$$\lim_{\varepsilon\to 0+} \bar{\Psi}(\lambda^* + \varepsilon, \mu^*) = \Psi(\lambda^*, \mu^*). \tag{6.16}$$

For any $(\lambda, \mu) \in \text{int}\Omega_0$, $H(\lambda, \mu)$ is positive definite. Thus, it is easy to see that

$$\lim_{\varepsilon \to 0+} \bar{\Psi}(\lambda + \varepsilon, \mu) = \Psi(\lambda, \mu) \quad \text{for all } (\lambda, \mu) \in \text{int}\Omega_0. \tag{6.17}$$

Therefore, (6.5) follows from (6.13), (6.16), and (6.17).

Now we consider the case that $\mu^* \neq 0$, and $\det(B + \lambda^* I + \mu AA^T) \not\equiv 0$ for $\mu \geq \mu^*$. Let $\bar{g} = g - \varepsilon AA^T d^*$, where d^* is a global solution of (6.6)–(6.8) with $(\lambda^*, \mu^* + \varepsilon)$ the corresponding pair of Lagrangian multipliers. Our assumption implies that there exists a small ε^* such that

$$B + \lambda^* I + (\mu^* + \varepsilon) AA^T \tag{6.18}$$

is positive definite for all $0 < \varepsilon < \varepsilon^*$. Then, we also have equalities (6.15) and (6.17), and, similarly, we have

$$\begin{aligned} &\lim_{\varepsilon \to 0+} \bar{\Psi}(\lambda^*, \mu^* + \varepsilon) \\ = \; & -\tfrac{1}{2} {d^*}^T H(\lambda^*, \mu^*) d^* - \tfrac{\lambda^*}{2} \Delta^2 - \tfrac{\mu^*}{2}(\xi^2 - \|c\|^2). \end{aligned} \tag{6.19}$$

Thus we can prove the same result. □

Moreover, with the additional assumption that d^* is feasible for both constraints, we can prove the result of Lemma 6.2 without the assumption that $\det(H(\lambda^*, \mu)) \not\equiv 0$ for $\mu > \mu^*$. Suppose the condition $\det(H(\lambda^*, \mu)) \not\equiv 0$ for $\mu > \mu^*$ fails, i.e., $\det(B + \lambda^* I + \mu AA^T) \equiv 0$ for $\mu \geq \mu^*$. Let μ_c be the minimal μ such that $(\lambda^*, \mu) \in \Omega_0$, i.e., $H(\lambda^*, \mu)$ is positive semidefinite for $\mu = \mu_c$ and is not positive semidefinite for $\mu < \mu_c$. By Lemma 4.2, $\Psi(\lambda^*, \mu)$ is a concave function on $\mu \in [\mu_c, +\infty)$. By the arguments of Theorem 3.2,

$$\frac{d\Psi(\lambda^*, \mu)}{d\mu} |_{\mu = \mu^*} \leq 0, \tag{6.20}$$

which holds as an inequality only if $\mu^* = \mu_c$. Hence μ^* is the maximizer of function $\Psi(\lambda^*, \cdot)$ on $[\mu_c, +\infty)$. Thus, by (3.15) and (3.16), (λ^*, μ^*) is the maximizer of $\Psi(\lambda, \mu)$.

In the following, we discuss some properties of the so-called shifted problem.

We consider the "shifted" problem $\widehat{P}$:

$$\min \quad \hat{\Phi}(\hat{d}) = \frac{1}{2} \hat{d}^T (B + \lambda_e I + \mu_e AA^T) \hat{d} + (g + \mu_e Ac)^T \hat{d} \tag{6.21}$$

subject to

$$\|\hat{d}\|^2 \leq \Delta^2, \tag{6.22}$$

$$\|A^T \hat{d} + c\|^2 \leq \xi^2. \tag{6.23}$$

Actually, except for a constant, the objective function $\hat{\Phi}(d)$ is the sum of the original objective function $\Phi(d)$ and a penalty term $\frac{1}{2}(\lambda_e \|d\|^2 + \mu_e \|A^T d + c\|^2)$.

The dual function of $\widehat{P}$ is

$$\hat{\Psi}(\hat{\lambda}, \hat{\mu}) = \hat{\Phi}(\hat{d}) + \frac{\hat{\lambda}}{2}(\|\hat{d}\|^2 - \Delta^2) + \frac{\hat{\mu}}{2}(\|A^T \hat{d} + c\|^2 - \xi^2), \tag{6.24}$$

where

$$\hat{d} = \hat{d}(\hat{\lambda}, \hat{\mu}) = -(\hat{B} + \hat{\lambda} I + \hat{\mu} AA^T)^{-1} (\hat{g} + \hat{\mu} Ac), \tag{6.25}$$

$\hat{B} = B + \lambda_e I + \mu_e AA^T$, $\hat{g} = g + \mu_e Ac$, and $\hat{\lambda} \geq 0, \hat{\mu} \geq 0$ are the multipliers of problem $\widehat{P}$. The Hessian of the Lagrangian is $\hat{H}(\hat{\lambda}, \hat{\mu}) = \hat{B} + \hat{\lambda} I + \hat{\mu} AA^T$. We also define the regions,

$$\widehat{\Omega}_0 = \{(\hat{\lambda}, \hat{\mu}) \in \mathcal{R}^2_+ \mid \hat{H}(\hat{\lambda}, \hat{\mu}) \text{ is positive semidefinite}\}, \tag{6.26}$$

and

$$\widehat{\Omega}_1 = \{(\hat{\lambda}, \hat{\mu}) \in \mathcal{R}^2_+ \mid \hat{H}(\hat{\lambda}, \hat{\mu}) \text{ has one negative eigenvalue}\}. \tag{6.27}$$

It is easy to show that

$$\widehat{\Omega}_0 = (\Omega_0 \cap \{\lambda \geq \lambda_e, \mu \geq \mu_e\}) - (\lambda_e, \mu_e) \tag{6.28}$$

and

$$\widehat{\Omega}_1 = (\Omega_1 \cap \{\lambda \geq \lambda_e, \mu \geq \mu_e\}) - (\lambda_e, \mu_e). \tag{6.29}$$

The statement (6.29) also holds for the connected branches $\widehat{\Omega}_{1j}$ and Ω_{1j} of the regions $\widehat{\Omega}_1$ and Ω_1, respectively, if $\Omega_{1j} \cup \{\lambda \geq \lambda_e, \mu \geq \mu_e\} \neq \emptyset$. We also have that two connected branches $\widehat{\Omega}_{1j}$ and $\widehat{\Omega}_{1k}$ of $\widehat{\Omega}_1$ are consecutive or adjoint if and only if their counterparts, Ω_{ij} and Ω_{1k}, the connected branches of Ω_1, are consecutive or adjoint (assuming that all the connected branches are well defined). In other words, the translated dual plane holds all these properties of the dual plane of original problem if the dual variables satisfy $\lambda \geq \lambda_e$ and $\mu \geq \mu_e$. Similarly to (3.7),

$$\nabla\hat{\Psi}(\hat{\lambda}, \hat{\mu}) = \frac{1}{2}\begin{pmatrix} \|\hat{d}(\hat{\lambda}, \hat{\mu})\|^2 - \Delta^2 \\ \|A^T\hat{d}(\hat{\lambda}, \hat{\mu}) + c\|^2 - \xi^2 \end{pmatrix}. \tag{6.30}$$

From the KKT conditions of the original problem and those of problem $\widehat{P}$, we have

$$d(\lambda + \lambda_e, \mu + \mu_e) = \hat{d}(\lambda, \mu). \tag{6.31}$$

So, by (6.30) and (3.7), the following equation holds:

$$\nabla\Psi(\lambda + \lambda_e, \mu + \mu_e) = \nabla\hat{\Psi}(\lambda, \mu). \tag{6.32}$$

Moreover, the difference between the dual functions of these two problems is only a constant depending on λ_e and μ_e:

$$\Psi(\lambda + \lambda_e, \mu + \mu_e) = \hat{\Psi}(\lambda, \mu) + \frac{\lambda_e}{2}\Delta^2 + \frac{\mu_e}{2}(\xi^2 - \|c\|^2). \tag{6.33}$$

By definition (4.1), the equality (6.33) holds also for $(\lambda, \mu) \in \partial\Omega_0$. Hence,

$$(\lambda_+ - \lambda_e, \mu_+ - \mu_e) = \arg \max_{(\hat{\lambda}, \hat{\mu}) \in \widehat{\Omega}_0} \widehat{\Psi}(\hat{\lambda}, \hat{\mu}). \tag{6.34}$$

Now we are ready to prove the main result of this section.

THEOREM 6.3. *If S is a singleton, there exist multipliers (λ, μ) in the related region of S such that (λ, μ) are the corresponding Lagrangian multipliers of a global minimizer of* (1.1)–(1.3).

Proof. If there is a feasible point d_* such that the triple (λ_+, μ_+, d_*) solves (2.1)–(2.3), then the global solution of problem (3.1)–(3.3) is d_*.

Suppose $\hat{d}_g$ is a global solution of problem (6.21)–(6.23), and $(\hat{\lambda}_g, \hat{\mu}_g) \in \widehat{\Omega}_1$ are the corresponding Lagrangian multipliers. If we suppose $\lambda_e \neq 0$, then we have $\hat{\lambda}_g \neq 0$. Otherwise, since

$$ri\{\{\hat{\lambda} = \hat{\lambda}_g\} \cap \Omega_1\} = \emptyset, \tag{6.35}$$

$\widehat{H}(\hat{\lambda}_g, \hat{\mu}_g)$ is positive semidefinite, and then $\widehat{\Psi}$ reaches its maximum at two points $(\hat{\lambda}_g, \hat{\mu}_g)$ and $(\lambda_+ - \lambda_e, \mu_+ - \mu_e)$, which means that Ψ also reaches its maximum at two points. This contradicts the assumption that S is a singleton. If $\hat{\lambda}_g = 0$, then $\lambda_e = 0$ and $\lambda_e + \hat{\lambda}_g = 0$. Then we have

$$(\hat{\lambda}_g + \lambda_e)(\|d(\hat{\lambda}_g + \lambda_e, \hat{\mu}_g + \mu_e)\| - \Delta) = \hat{\lambda}_g(\|\hat{d}(\hat{\lambda}_g, \hat{\mu}_g)\| - \Delta) = 0 \tag{6.36}$$

and, similarly,

$$(\hat{\mu}_g + \mu_e)(\|A^T d(\hat{\lambda}_g + \lambda_e, \hat{\mu}_g + \mu_e) + c\| - \xi) = \hat{\mu}_g(\|A^T \hat{d}(\hat{\lambda}_g, \hat{\mu}_g) + c\| - \xi) = 0. \tag{6.37}$$

This implies that

$$\hat{\Phi}(\hat{d}_g) = \Phi(\hat{d}_g) + \frac{\lambda_e}{2}\Delta^2 + \frac{\mu_e}{2}(\xi^2 - \|c\|^2). \tag{6.38}$$

Moreover,

$$d(\lambda_g, \mu_g) = \hat{d}(\hat{\lambda}_g, \hat{\mu}_g) \tag{6.39}$$

follows (6.31) with (λ_g, μ_g) the corresponding Lagrangian multipliers. Furthermore, for any feasible $d \in \mathcal{R}^n$ of the original problem, we have that

$$\hat{\Phi}(\hat{d}_g) \leq \hat{\Phi}(d). \tag{6.40}$$

Expressions (6.38) and (6.40) imply that

$$\begin{aligned} \Phi(d(\lambda_g, \mu_g)) &\leq \Phi(d) + \lambda_e(\|d\|^2 - \Delta^2) + \mu_e(\|A^T d + c\|^2 - \xi^2) \\ &\leq \Phi(d). \end{aligned} \tag{6.41}$$

The above inequality indicates that $d(\lambda_g, \mu_g)$ is a global solution of the original problem. This completes our proof. □

The above theorem illustrates the relation between the location of the Lagrangian multipliers and the maxima of the dual function on the region where the Hessian of the Lagrangian is positive semidefinite. From Definition 6.1, we can see that the choices of λ_e and μ_e are independent of each other. It can be seen that the larger λ_e and μ_e are, the smaller the related region is. Now we choose the minimal related region of S, i.e., we find the maxima (or supremum) of $\mathcal{L}$ and $\mathcal{M}$.

First, we consider the set $\mathcal{M}$. For $\varepsilon > 0$, we consider all the indices j such that

$$\Omega_{1j} \cap \{\mu_+ - \varepsilon < \mu < \mu_+\} \neq \emptyset. \tag{6.42}$$

By Lemma 3.3, this set can be divided into three cases.

Case 1. For sufficiently small $\varepsilon > 0$, there is no such j. Then, we may choose $\mu_e = \mu_+ - \varepsilon$ for sufficiently small $\varepsilon > 0$. Actually, the global solution lies in $\{\mu \geq \mu_+\}$.

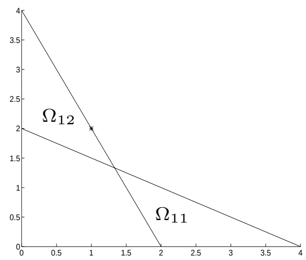

FIG. 6.1. (a) *of Example* 6.4.

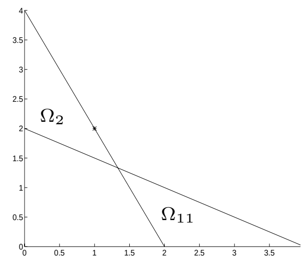

FIG. 6.2. (b) *of Example* 6.4.

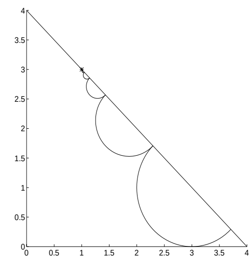

FIG. 6.3. (c) *of Example* 6.4.

Case 2. For sufficiently small $\varepsilon > 0$, there is exactly one such j. Then we easily see that the best choice of μ_e is $\mu_e = l_{\mu j}$.

Case 3. For any small $\varepsilon > 0$, there are infinitely many j such that (6.42) holds. Then μ_e can be chosen as any $l_{\mu j}$, which implies, actually, that the global solution lies in $\{\mu \geq l_{\mu j}\}$ for all j. Since the supremum of all these indices is μ_+, then the global solution lies in $\{\mu \geq \mu_+\}$ as in the first case.

These three cases show that the number of existent connected branches in $\{\mu < \mu_+\}$ is at least one. See Figures 6.1, 6.2, and 6.3.

Example 6.4. (a) Let

$$B = \mathrm{diag}(-2, -4), \qquad A = \mathrm{diag}\left(\frac{\sqrt{2}}{2}, \sqrt{2}\right).$$

The related region of $(4/3, 4/3)$ is $\Omega_{11} \cup \Omega_{12}$.

(b) Let

$$B = \mathrm{diag}(-2, -2, -4), \qquad A = \mathrm{diag}\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \sqrt{2}\right).$$

The related region of $(4/3, 4/3)$ is Ω_{11}, and the Hessian with Lagrangian multiplier in Ω_2 has two negative eigenvalues.

(c) We would like to show an example where there are infinitely many connected branches near one point. However, we could not find such an example or prove its nonexistence.

For the λ-direction, the case is slightly different. Considering Case 2, if there exists k such that $\Omega_{1k} \cap \{\lambda_+ - \varepsilon < \lambda < \lambda_+\} \neq \emptyset$, then the following statement may fail:

$$ri\{\{\lambda = \lambda_{\lambda k}\} \cap \Omega_1\} = \emptyset. \tag{6.43}$$

That is to say, $B + l_{\lambda k} I + \mu A A^T$ can be singular and have one negative eigenvalue in an interval $(\underline{\mu_s}, \overline{\mu_s})$. Therefore, we cannot set $\lambda_e = l_{\lambda k}$. In this case,

$$\det(B + l_{\lambda k} I + \mu A A^T) = 0 \tag{6.44}$$

for $\mu \in (\underline{\mu_s}, \overline{\mu_s})$. Since $\det(B + l_{\lambda k} I + \mu A A^T)$ is a polynomial of μ, the above relation implies that

$$\det(B + l_{\lambda k} I + \mu A A^T) = 0 \quad \text{for all}\, \mu \geq 0. \tag{6.45}$$

Thus for any $\lambda < \lambda_{\lambda k}$, $B + \lambda I + \mu AA^T$ has at least one negative eigenvalue for all μ, which implies that $l_{\lambda k'} = 0$ and $u_{\mu k'} = +\infty$ if there is a connected branch $\Omega_{1k'}$ such that $\Omega_{1k} \succ \Omega_{1k'}$. Therefore, there is at most one $\Omega_{1k'}$ such that $\Omega_{1k} \succ \Omega_{1k'}$. So we can set $\lambda_e = 0$ and there are at most two connected branches in the region $\{0 \le \lambda \le \lambda_+\}$.

For example, let

$$B = \operatorname{diag}(-5, -8, -3), \quad A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}; \tag{6.46}$$

then $\lambda = 3$ is the singular line and $(\underline{\mu_s}, \overline{\mu_s}) = (\frac{5}{4}, 2)$. In section 7, we will present an example to show that the Hessian at the global solution may have one negative eigenvalue and can be singular.

From the above analyses, there are at most three connected branches of Ω_1 in the minimal related region of S. In the case that there are three connected branches in the minimal related region of S, we will have that (λ_+, μ_+) is an adjoint point of two connected branches of Ω_1. Moreover, there must be a singular line of the Lagrangian Hessian in the related region. Thus there are indices $k, k' \in \mathcal{K}$ such that

$$l_{\lambda k} = \lambda_+ = u_{\lambda k'} \tag{6.47}$$

$$u_{\mu k} = \mu_+ = l_{\mu k'}, \tag{6.48}$$

and

$$\det(B + l_{\lambda k'} I + \mu AA^T) \equiv 0 \quad \text{for all}\, \mu \ge 0. \tag{6.49}$$

Here, it follows from (6.47)–(6.48) that Ω_{1k} and $\Omega_{1k'}$ are two adjoint connected branches.

7. Distribution of local solutions. In this section we show that, for the CDT problem, there may exist two local solutions whose corresponding Lagrangian multipliers lie in the same connected branch Ω_{1k} defined by (3.18) of the region where the Hessian of the Lagrangian possesses exactly one negative eigenvalue. It is also possible that the Hessian of the Lagrangian can have one negative eigenvalue and a zero eigenvalue. The following example shows that there may exist two local solutions in one connected branch of Ω_1.

Example 7.1. Let

$$B = \begin{pmatrix} -34/9 & \\ & -3 \end{pmatrix}, \quad A = \begin{pmatrix} 4/3 & \\ & 1 \end{pmatrix}, \quad g = \begin{pmatrix} 24 \\ 27 \end{pmatrix}, \quad c = \begin{pmatrix} -10 \\ 13 \end{pmatrix}, \tag{7.1}$$

$\Delta = 13$, and $\xi = 10$.

The global and local-nonglobal solutions of problem (3.1)–(3.3) are

$$d_g = \begin{pmatrix} 0 \\ -13 \end{pmatrix}, \quad d_l = \begin{pmatrix} 12 \\ -5 \end{pmatrix}, \tag{7.2}$$

with the Lagrangian multipliers $(\lambda_g, \mu_g) = (12/13, 9/5)$ and $(\lambda_l, \mu_l) = (28/51, 94/51)$, respectively. Then we easily see, from Figures 7.1 and 7.2, that the two points

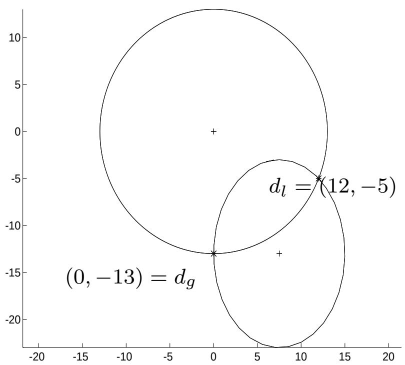

FIG. 7.1. *Primal space.*

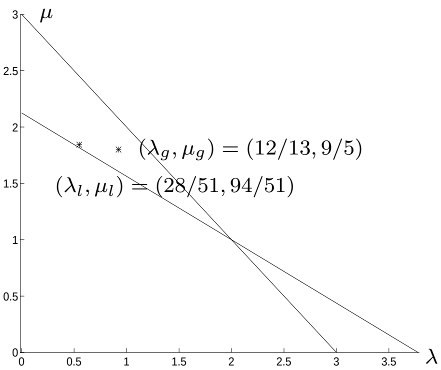

FIG. 7.2. *Dual space.*

$(12/13, 9/5)$, $(28/51, 94/51)$ are in the same connected branch. For the single-ball-constrained quadratic minimization, there is at most one local-nonglobal solution; see Martínez [10].

To show the distribution of local solutions, first we need a lemma.

LEMMA 7.2. *Assume that*

(a) $d(\lambda_1^*, \mu_1^*)$ *and* $d(\lambda_2^*, \mu_2^*)$ *are two stationary points of problem* (3.1)–(3.3), *i.e., satisfy the KKT system,* (2.1)–(2.3);

(b) $\lambda_2^* \geq \lambda_1^* \geq 0$ *and* $\mu_2^* \geq \mu_1^* \geq 0$;

then

$$\Phi(d(\lambda_2^*, \mu_2^*)) \leq \Phi(d(\lambda_1^*, \mu_1^*)). \tag{7.3}$$

The equality in (7.3) *holds if and only if*

$$\lambda_2^* = \lambda_1^* \tag{7.4}$$

and

$$(\mu_2^* - \mu_1^*)A^T(d(\lambda_2^*, \mu_2^*) - d(\lambda_1^*, \mu_1^*)) = 0. \tag{7.5}$$

Proof. For simplicity, we use the notations

$$H_1 = H(\lambda_1^*, \mu_1^*), \qquad H_2 = H(\lambda_2^*, \mu_2^*), \tag{7.6}$$

and

$$g_1 = g + \mu_1^* Ac, \qquad g_2 = g + \mu_2^* Ac. \tag{7.7}$$

Let $d_i = d(\lambda_i^*, \mu_i^*)$, $i = 1, 2$. Then we get

$$H_1 d_1 = -g_1, \qquad H_2 d_2 = -g_2. \tag{7.8}$$

Using (7.8), we have

$$\Phi(d_i) = -\frac{1}{2} d_i{}^T H_i d_i - \frac{\lambda_i^*}{2}\Delta^2 - \frac{\mu_i^*}{2}(\xi^2 - \|c\|^2) \tag{7.9}$$

for $i = 1, 2$. Hence

$$\begin{aligned}
&\Phi(d_1) - \Phi(d_2)\\
&= \tfrac{1}{2}d_2^T H_2 d_2 - \tfrac{1}{2}d_1^T H_1 d_1 + \tfrac{1}{2}(\lambda_2^* - \lambda_1^*)\Delta^2 + \tfrac{1}{2}(\mu_2^* - \mu_1^*)(\xi^2 - \|c\|^2)\\
&\geq \tfrac{1}{2}d_1^T g_1 - \tfrac{1}{2}d_2^T g_2 + \tfrac{1}{2}(\lambda_2^* - \lambda_1^*)d_1^T d_2\\
&\qquad + \tfrac{1}{2}(\mu_2^* - \mu_1^*)((A^T d_1 + c)^T (A^T d_2 + c) - \|c\|^2) \qquad (7.10)\\
&= \tfrac{1}{2}d_1^T g_1 - \tfrac{1}{2}d_2^T g_2 + \tfrac{1}{2}d_1^T (H_2 - H_1) d_2 + \tfrac{1}{2}(\mu_2^* - \mu_1^*)(d_1 + d_2)^T Ac\\
&= \tfrac{1}{2}d_1^T g_1 - \tfrac{1}{2}d_2^T g_2 - \tfrac{1}{2}d_1^T g_2 + \tfrac{1}{2}d_2^T g_1 + \tfrac{1}{2}(d_1 + d_2)^T (g_2 - g_1)\\
&= 0.
\end{aligned}$$

The equality in (7.3) holds if and only if the equality in (7.10) holds, which is equivalent to

$$(\lambda_2^* - \lambda_1^*)(\Delta^2 - {d_1}^T d_2) = 0 \tag{7.11}$$

and

$$(\mu_2^* - \mu_1^*)(\xi^2 - (A^T d_1 + c)^T (A^T d_2 + c)) = 0. \tag{7.12}$$

If $\lambda_2^* > \lambda_1^*$, (7.11) gives

$${d_1}^T d_2 = \Delta^2 = {d_1}^T d_1 = {d_2}^T d_2, \tag{7.13}$$

which implies that $d_1 = d_2$. Then,

$$(\lambda_2^* - \lambda_1^*)d_1 + (\mu_2^* - \mu_1^*)A(A^T d_1 + c) = 0, \tag{7.14}$$

and any triple (λ, μ, d), with (λ, μ) in the straight line joining (λ_2^*, μ_2^*) and (λ_1^*, μ_1^*), is also a solution to the KKT system. Similar to equation (2.48) of Yuan [16],

$${d_1}^T A(A^T d_1 + c) \geq 0. \tag{7.15}$$

Relations (7.14), (7.15) and $\mu_2^* \geq \mu_1^*$ give

$$(\lambda_2^* - \lambda_1^*)\|d_1\|^2 \leq 0, \tag{7.16}$$

which contradicts $\lambda_2^* > \lambda_1^*$ and (7.13). Therefore, (7.11) is true if and only if $\lambda_2^* = \lambda_1^*$.

If $\mu_2^* > \mu_1^*$, the equality in (7.12) gives

$$(A^T d_1 + c)^T (A^T d_2 + c) = \xi^2 = \|A^T d_1 + c\|^2 = \|A^T d_2 + c\|^2, \tag{7.17}$$

which implies $A^T d_1 + c = A^T d_2 + c$, i.e., $A^T(d_1 - d_2) = 0$. On the other hand, if $A(d_1 - d_2) = 0$, we have $A^T d_1 + c = A^T d_2 + c$, which implies

$$(A^T d_1 + c)^T (A^T d_2 + c) = \|A^T d_1 + c\|^2 = \|A^T d_2 + c\|^2. \tag{7.18}$$

Since $\mu_2^* > \mu_1^* \geq 0$, we have

$$\|A^T d_2 + c\| = \xi; \tag{7.19}$$

therefore $(A^T d_1 + c)^T (A^T d_2 + c) = \xi^2$. Thus we see that (7.12) is equivalent to (7.5). □

In Theorem 7.4, we have two points satisfying (b) of Lemma 7.2. Furthermore, Theorem 7.4 states that these two points are not local solutions at the same time while both the Hessians are not singular. This assumption cannot be moved, as the following example shows. In this example we show that there may exist a global solution of (3.1)–(3.3) with its Hessian having one negative eigenvalue and being singular.

Example 7.3. In this example, (3.1)–(3.3) have the global solution (λ_g, μ_g, d_g) satisfying $\lambda_g = 0$, both constraints are active at d_g and the Hessian $H(\lambda_g, \mu_g)$ has one negative eigenvalue.

$$B = \begin{pmatrix} -1 & \\ & -4 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}, \quad g = \begin{pmatrix} 2 \\ -10 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \tag{7.20}$$

$\Delta = \sqrt{2}$ and $\xi = \sqrt{5}$; then $d_g = (-1, -1)^T$ with its Lagrangian multipliers $(\lambda_g, \mu_g) = (0, 3)$ is the global solution of (3.1)–(3.3). Then the problem is

$$\min_{\bar{d} \in \mathcal{R}^3} \quad \frac{1}{2} \bar{d}^T \bar{B} \bar{d} + \bar{g}^T \bar{d} \tag{7.21}$$

subject to

$$\|\bar{d}\|^2 \leq \Delta^2, \tag{7.22}$$

$$\|\bar{A}^T \bar{d} + \bar{c}\|^2 \leq \xi^2, \tag{7.23}$$

where $\bar{d} = (d_1, d_2, d_3)^T$, $\bar{B} = \text{diag}(B, 0)$, $\bar{A}^T = (A^T, 0)$, $\bar{g}^T = (g^T, 0)$, and $\bar{c}^T = (c^T, 0)$. Its global solution is $\bar{d}_g = (-1, -1, 0)^T$, with the Lagrangian multipliers $(\lambda_g, \mu_g) = (0, 3)$ and the Hessian $H_g = \text{diag}(2, -1, 0)$. So the Hessian H_g at the global solution has one negative eigenvalue and is singular. However, for the single-ball-constrained quadratic minimization, there is no local solution where the Hessian has one negative eigenvalue and is singular (see Martínez [10]).

THEOREM 7.4. *It is not possible for two local solutions $d(\lambda_1^*, \mu_1^*)$ and $d(\lambda_2^*, \mu_2^*)$ of problem* (3.1)–(3.3) *to satisfy $\lambda_1^* > \lambda_2^* > 0$, $\mu_1^* \geq \mu_2^* > 0$, with $H(\lambda_i^*, \mu_i^*), i = 1, 2$, being a nonsingular matrix with exactly one negative eigenvalue.*

Proof. Suppose that (λ_1^*, μ_1^*) and (λ_2^*, μ_2^*) are the Lagrangian multipliers satisfying the above assumption.

Consider the following problem $(\tilde{P})$:

$$\min_{\tilde{d} \in \mathcal{R}^n} \frac{1}{2} \tilde{d}^T \tilde{B} \tilde{d} + \tilde{g}^T \tilde{d} \tag{7.24}$$

subject to

$$\|\tilde{D}^T \tilde{d} + \tilde{c}\| \leq \tilde{\xi}, \tag{7.25}$$

where

$$\begin{aligned} \tilde{B} &= B + \lambda_2^* I + \mu_2^* A A^T, \\ \tilde{g} &= g + \mu_2^* A c, \\ \tilde{D} &= (\tau_1 I + \tau_2 A A^T)^{\frac{1}{2}}, \\ \tilde{c} &= \tau_2 \tilde{D}^{-1} A c, \\ \tilde{\xi} &= (\tilde{c}^T \tilde{c} - \tau_2 \xi^2 + \tau_2 c^T c + \tau_1 \Delta^2)^{\frac{1}{2}}, \end{aligned} \tag{7.26}$$

and $\tau_1 = \lambda_1^* - \lambda_2^*$ and $\tau_2 = \mu_1^* - \mu_2^*$.

Since $\lambda_1^* > \lambda_2^* > 0$ and $\mu_1^* \geq \mu_2^* > 0$, $\tilde{D}$ and $\tilde{c}$ in (7.26) are well defined. Suppose $\tilde{t}$ is the Lagrangian multiplier of $(\tilde{P})$; then the KKT system of $(\tilde{P})$ is

$$(\tilde{B} + \tilde{t}\tilde{D}\tilde{D}^T)\tilde{d}^* = -(\tilde{g} + \tilde{t}\tilde{D}\tilde{c}) \tag{7.27}$$

and

$$\tilde{t}(\|\tilde{D}^T\tilde{d}^* + \tilde{c}\| - \tilde{\xi}) = 0. \tag{7.28}$$

It can be verified that $d(\lambda_1^*, \mu_1^*)$ and $d(\lambda_2^*, \mu_2^*)$ are two stationary points of $(\tilde{P})$ with multipliers $\tilde{t} = 1$ and $\tilde{t} = 0$, respectively. Set

$$\phi(\tilde{t}) = \frac{1}{2}(\|\tilde{D}^T\tilde{d}(\tilde{t}) + \tilde{c}\|^2 - \tilde{\xi}^2). \tag{7.29}$$

By direct calculations and from the result given by Martínez [10], we have

$$\begin{aligned}
&\phi'(\tilde{t})|_{\tilde{t}=0} \\
&= -(\tilde{g} - \tilde{B}\tilde{D}^{-T}\tilde{c})^T\tilde{H}^{-1}\tilde{D}\tilde{D}^T\tilde{H}^{-1}\tilde{D}\tilde{D}^T\tilde{H}^{-1}(\tilde{g} - \tilde{B}\tilde{D}^{-T}\tilde{c})|_{\tilde{t}=0} \\
&= -(\tau_1 d_1^* + \tau_2 y_1^*)^T H(\lambda_1^*, \mu_1^*)^{-1}(\tau_1 d_1^* + \tau_2 y_1^*) \\
&= (\tau_1, \tau_2)\nabla^2\Psi(\lambda_1^*, \mu_1^*)\begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \\
&< 0,
\end{aligned} \tag{7.30}$$

where $\tilde{H} = \tilde{B} + \tilde{t}\tilde{D}\tilde{D}^T$, $d_1^* = d(\lambda_1^*, \mu_1^*)$, and $y_1^* = A(A^T d_1^* + c)$. Since there are two stationary points with the Hessian one negative eigenvalue, from Lemma 4.3 in Martínez [10], (7.30) is a strict inequality.

Let $\delta d = d_2^* - d_1^*$, where $d_2^* = d(\lambda_2^*, \mu_2^*)$; then

$$H(\lambda_1^*, \mu_1^*)\delta d = \tau_1 d_1^* + \tau_2 y_1^* \tag{7.31}$$

follows by direct calculations. Hence

$$\begin{aligned}
\delta d^T H(\lambda_1^*, \mu_1^*)\delta d &= \delta d^T H(\lambda_1^*, \mu_1^*) H(\lambda_1^*, \mu_1^*)^{-1} H(\lambda_1^*, \mu_1^*)\delta d \\
&= -(\tau_1, \tau_2)\nabla^2\Psi(\lambda_1^*, \mu_1^*)\begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \\
&> 0.
\end{aligned} \tag{7.32}$$

Define

$$S^1 = \left\{ d \mid \langle \frac{d}{\|d\|}, \frac{\delta d}{\|\delta d\|} \rangle \geq 1 - \varepsilon \right\} \cup \{0\} \tag{7.33}$$

and

$$S^2 = \left\{ d \mid \langle \frac{d}{\|d\|}, \frac{\delta d}{\|\delta d\|} \rangle \leq 1 - \frac{1}{2}\varepsilon \right\} \cup \{0\}, \tag{7.34}$$

where

$$\varepsilon^{\frac{1}{2}} = \frac{\delta d^T H_* \delta d}{8\|\delta d\|^2 \|H_*\|} \tag{7.35}$$

and $H_* = H(\lambda_1^*, \mu_1^*)$. Since H_* has one negative eigenvalue and $\delta d \neq 0$, (7.35) is well defined. It is easy to verify that S^1 and S^2 are closed sets. Moreover, if $d \in S^1$,

$$\left\| \frac{d}{\|d\|} - \frac{\delta d}{\|\delta d\|} \right\| \leq 2\varepsilon \tag{7.36}$$

and

$$\begin{aligned} &\tfrac{1}{\|d\|^2} d^T H_* d \\ &= \left(\tfrac{d}{\|d\|} - \tfrac{\delta d}{\|\delta d\|} + \tfrac{\delta d}{\|\delta d\|} \right)^T H_* \left(\tfrac{d}{\|d\|} - \tfrac{\delta d}{\|\delta d\|} + \tfrac{\delta d}{\|\delta d\|} \right) \\ &\geq \tfrac{\delta d^T H_* \delta d}{\|\delta d\|^2} - 2\sqrt{2\varepsilon}\|H_*\| - 2\varepsilon\|H_*\| \\ &\geq 0, \end{aligned} \tag{7.37}$$

while if $d \in S^2$,

$$d^T \left(I - \frac{\delta d \delta d^T}{\|\delta d\|^2} \right) d \geq 0. \tag{7.38}$$

Let

$$\theta_1 = \min\{d^T H_* d \mid \|d\| = 1, d \in S^1\}, \tag{7.39}$$

$$\theta_2 = \min \left\{ d^T \left(I - \frac{\delta d \delta d^T}{\|\delta d\|^2} \right) d \mid \|d\| = 1, d \in S^2 \right\}, \tag{7.40}$$

and $\theta = \min\{\theta_1, \theta_2\}$. It can be verified that $\theta > 0$ since S^1, S^2, and $\{d \mid \|d\| = 1\}$ are all closed sets. From Lemma 2.3 of Yuan [16], there are $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = 1$ such that

$$\alpha_1(H_* - \theta I) + \alpha_2 \left(I - \theta I - \frac{\delta d \delta d^T}{\|\delta d\|^2} \right) \tag{7.41}$$

is positive semidefinite. Since neither $(H_* - \theta I)$ nor $(I - \theta I - \frac{\delta d \delta d^T}{\|\delta d\|^2})$ is positive semidefinite, $\alpha_1 \neq 0$ and $\alpha_2 \neq 0$. Let $m_0 = \frac{\alpha_2}{\alpha_1}$; the matrix $H_* + m_0(I - \frac{\delta d \delta d^T}{\|\delta d\|^2})$ is positive semidefinite. Now the problem (P_p)

$$\min_d \quad \frac{1}{2} d^T B_p d + g_p^T d \tag{7.42}$$

subject to

$$\|d\| \leq \Delta, \tag{7.43}$$

$$\|A^T d + c\| \leq \xi, \tag{7.44}$$

where $B_p = B + m_0(I - \frac{\delta d \delta d^T}{\|\delta d\|^2})$ and $g_p = g - (B_p - B)d_1^*$, possesses two global solutions

$$(\lambda_1^*, \mu_1^*, d_1^*) \text{ and } (\lambda_2^*, \mu_2^*, d_2^*), \tag{7.45}$$

both with positive semidefinite Hessian. That their objective function value must be the same contradicts Lemma 7.2. ∎

Remark. From Theorem 7.4, all the local solutions with the multipliers in $\mathrm{int}\Omega_1$ are permuted in the way the connected branches of Ω_1 are. As Example 7.1 shows, in one connected branch of Ω_1, there may exist a global and a local solution simultaneously. It is important to give the characteristic of the global solution and hence construct an algorithm with which to find the global solution instead of the local solution.

8. Conclusions and future work. We investigated the dual plane of the CDT subproblem which is related to a matrix pencil with two parameters. We also extended the general Lagrangian dual function from the region where the Lagrangian Hessian is positive definite to its closure. The location and permutation of the Lagrangian multipliers were studied and the differences between the CDT subproblem and the trust region subproblem were presented.

We have given various results on the locations of the corresponding Lagrange multipliers. These results may be used in the construction of numerical methods for the CDT subproblem based on identifying the multipliers. The main result shows that the Lagrangian multipliers corresponding to a global minimizer of the CDT problem locates in finitely many, often two or three, connected branches of Ω_1 if there is no global minimizer with the Hessian of the Lagrangian positive semidefinite. Roughly speaking, if we define the degree of the nonpositive definite of a symmetric matrix as the number of its negative eigenvalues, an important property of the CDT problem is that its complexity is not related to the nonpositive degree of the Hessian.

For some trust region methods, the trial step can be any sufficient descent feasible direction instead of the global minimizer or a local minimizer. Thus, it is interesting to search for efficient algorithms to compute approximate global minimizers of the CDT subproblem in the primal space.

Acknowledgment. We would like to thank two anonymous referees for their valuable comments which improved the paper.

REFERENCES

[1] R. H. Byrd and R. B. Schnabel, *Approximate solution of the trust region problem by minimization over two-dimensional subspace*, Math. Programming Ser. A, 40 (1988), pp. 247–263.
[2] 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.
[3] Z. Chen, *Some algorithms for a class of CDT subproblems*, in Operations Research and Its Applications, Lecture Notes in Oper. Res. 1, D. Z. Du, X. Zhang, and W. Wang, eds., World Publishing Corporation, Beijing, 1996, pp. 108–114.
[4] J. E. Dennis and K. A. Williamson, *A robust trust region algorithm for nonlinear programming*, Technical Report, Math. Sci. Dept., Rice University, Houston, TX, 1991.
[5] J. E. Dennis, Jr., M. El-Alem, and M. C. Maciel, *A global convergence theory for general trust-region-based algorithms for equality constrained optimization*, SIAM J. Optim., 7 (1997), pp. 177–207.
[6] J. G. Ecker and R. D. Niemi, *A dual method for quadratic programs with quadratic constraints*, SIAM J. Appl. Math., 28 (1975), pp. 568–576.
[7] M. M. El-Alem and R. A. Tapia, *Numerical experience with a polyhedral-norm CDT trust region algorithm*, J. Optim. Theory Appl., 85 (1995), pp. 575–591.
[8] M. Y. Fu, Z. Q. Luo, and Y. Y. Ye, *Approximation Algorithms for Quadratic Programming*, University of Iowa, http://dollar.biz.uiowa.edu/col/ye/paper.html (1996).
[9] M. Heinkenschloss, *On the solution of a two ball trust region subproblem*, Math. Programming, 64 (1994), pp. 249–276.
[10] J. M. Martínez, *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optim., 4 (1994), pp. 159–176.
[11] J. M. Martínez and S. A. Santos, *A trust region strategy for minimization on arbitrary domains*, Math. Programming, 68 (1995) pp. 267–302.
[12] S. Mehrotra and J. Sun, *A method of analytic centers for quadratically constrained convex quadratic programs*, SIAM J. Numer. Anal., 28 (1991), pp. 529–544.
[13] J. M. Peng and Y. Yuan, *Optimality conditions for the minimization of a quadratic with two quadratic constraints*, SIAM J. Optim., 7 (1997), pp. 579–594.

[14] E. Phan-huy-Hao, *Quadratically constrained quadratic programming: Some applications and a method for solution*, Zeitschrift Oper. Res., 26 (1982), pp. 105–119.

[15] M. J. D. Powell and Y. Yuan, *A trust region algorithm for equality constrained optimization*, Math. Programming, 49 (1991), pp. 189–211.

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

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

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