MATHEMATICAL ENGINEERING
TECHNICAL REPORTS

Computing the Signed Distance between Overlapping Ellipsoids

Satoru IWATA, Yuji NAKATSUKASA
and Akiko TAKEDA

METR 2014–20 July 2014

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

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

The METR technical reports are published as a means to ensure timely dissemination of scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the authors or by other copyright holders, notwithstanding that they have offered their works here electronically. It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may not be reposted without the explicit permission of the copyright holder.

COMPUTING THE SIGNED DISTANCE BETWEEN OVERLAPPING ELLIPSOIDS*

SATORU IWATA , YUJI NAKATSUKASA , AND AKIKO TAKEDA †

Abstract. Computing the signed distance between two ellipsoids is a convex optimization problem when the two ellipsoids have no intersection, but becomes non-convex when the ellipsoids overlap. Efficient algorithms for convex optimization problems are thus not guaranteed to find the correct signed distance between overlapping ellipsoids. In this paper, we first show that computing the signed distance is equivalent to minimizing the norm along the boundary of the Minkowski difference. We then derive a polynomial-time $O(n^6)$ algorithm, where n is the dimension of the ellipsoids, that obtains a global minimizer on the boundary of the Minkowski difference and hence provides the exact signed distance. The algorithm first finds all the points that satisfy the Karush-Kuhn-Tucker (KKT) conditions, and then identifies a relevant KKT point with the smallest signed distance. The primary difficulty in computing the KKT points is that they are the solutions of two bivariate rational equations, whose poles are not known explicitly. Our key step is to convert the rational equations into polynomial equations, which we do by constructing certain bivariate matrix pencils whose zeros of determinants are the zeros of the rational functions. This reduces the problem to a two-parameter quadratic eigenvalue problem, which can be solved via a single-parameter linear eigenvalue problem of larger (squared) size, for which reliable algorithms are available. Thus we provide the first polynomial-time algorithm for computing the signed distance between overlapping ellipsoids.

Key words. ellipsoids, signed distance, non-convex optimization, KKT conditions, Minkowski difference, two-parameter eigenvalue problem, quadratic eigenvalue problem

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

1. Introduction.

An ellipsoid E lying in an n-dimensional space centered at $b \in \mathbb{R}^n$ can be represented as the set of points x satisfying

$$(x-b)^\top A^{-1}(x-b) \le 1,$$

where $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. We denote such an ellipsoid by $E(b, A)$. The signed distance between two ellipsoids $E_1 = E(b_1, A_1)$ and $E_2 = E(b_2, A_2)$ is defined by (see e.g. [19])

$$\operatorname{dist}(E_1, E_2) = \max_{\|w\|=1} \left(\langle w, b_1\rangle - \sqrt{\langle w, A_1 w\rangle} - \langle w, b_2\rangle - \sqrt{\langle w, A_2 w\rangle} \right). \tag{1.1}$$

Here $\langle \cdot, \cdot \rangle$ denotes the inner product, which for simplicity we take to be $\langle x, y\rangle = x^\top y$, and $\|\cdot\|$ designates the norm defined by $\|x\| = \sqrt{\langle x, x\rangle}$.

Computing the signed distance between ellipsoids arises in various engineering problems such as collision detection in graphics, motion planning in robotics, and binary classification in machine learning [3, 25, 29].

When the two ellipsoids do not overlap, the signed distance $\operatorname{dist}(E_1, E_2)$ coincides with the perhaps more intuitive formulation of the distance

$$\begin{aligned} &\text{minimize} && \|x_1 - x_2\| \\ &\text{subject to} && x_1 \in E_1, \quad x_2 \in E_2. \end{aligned} \tag{1.2}$$

*Version of July 28, 2014. This research is supported by Grants-in-Aid for Scientific Research (Challenging Exploring Rearch, No. 26540007) from Japan Society for Promotion of Science.

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan (iwata@mist.i.u-tokyo.ac.jp, nakatsukasa@mist.i.u-tokyo.ac.jp, takeda@mist.i.u-tokyo.ac.jp).

The equivalence can be verified by noting that for a fixed vector w, the quantity inside the parenthesis of (1.1) represents the distance between the two parallel hyperplanes that are tangent to the ellipsoids, see Figure 1.1 for an illustration. More specifically, $\langle w, b_1\rangle - \langle w, b_2\rangle$ is the distance between two parallel hyperplanes with normal vector w that pass through the centers b_1 and b_2, and $\sqrt{\langle w, A_1 w\rangle}$ and $\sqrt{\langle w, A_2 w\rangle}$ are the distance between the centers b_1, b_2 and the hyperplanes tangent to the ellipsoids and have gradient parallel to w. To see this, note that the gradient at x_1 on the boundary of E_1 is $2A_1^{-1}(x_1 - b_1)$, so its normalized vector $w = \frac{A_1^{-1}(x_1-b_1)}{\|A_1^{-1}(x_1-b_1)\|}$ satisfies $w^\top(x_1 - b_1) = \sqrt{\langle w, A_1 w\rangle}$, which is the distance between the two hyperplanes that pass through b_1 and x_1. The term $\sqrt{\langle w, A_2 w\rangle}$ is analogous. Therefore the maximization problem (1.1) over w corresponds to finding parallel tangent hyperplanes of E_1, E_2 with the smallest distance.

When the ellipsoids overlap, much of the above interpretation remains valid; one difference is that now $\mathrm{dist}(E_1, E_2)$ is negative, whose absolute value quantifies the amount of overlap. In either case, $\mathrm{dist}(E_1, E_2)$ is the smallest signed distance between two parallel hyperplanes that are tangent to E_1, E_2 such that the two ellipsoids lie on opposite sides of each hyperplane. The optimal vector w in (1.1) is in fact parallel to the gradients of both E_1 and E_2 at the points x_1 and x_2.

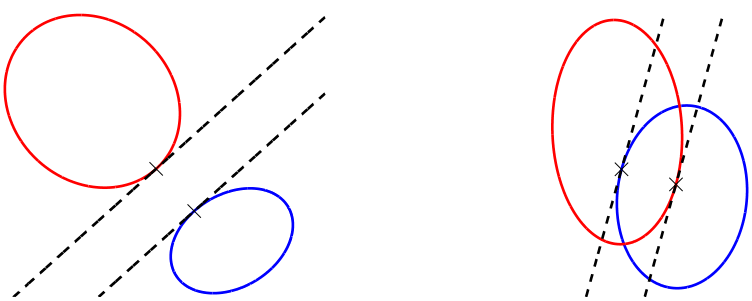

FIG. 1.1. *For non-overlapping ellipsoids E_1 and E_2 (left), the signed distance dist(E_1, E_2) is the distance between the two tangent planes indicated by the dashed lines. For overlapping ellipsoids (right) dist(E_1, E_2) is negative: the distance between the tangent planes times -1.*

The two cases of overlapping and non-overlapping ellipsoids as illustrated in Figure 1.1 lead to nontrivial mathematical differences. First, when the two ellipsoids overlap the equivalence between (1.1) and (1.2) is lost: the signed distance (1.1) takes a negative value, whereas the minimum value of (1.2) is zero. In this case, the signed distance provides information about how deeply the two ellipsoids cross with each other.

The second, and computationally more important, difference is that the problem of computing the signed distance becomes a *non-convex* optimization problem with many local minima. By contrast, in the non-overlapping case, finding the distance can be recast as a convex optimization problem and can be computed efficiently by solving second order cone programming (SOCP). A more direct geometric method has been proposed by Lin and Han [21], which computes the (standard) distance between two non-overlapping ellipsoids. This algorithm, which crucially relies on the convexity of the problem, is used by the Ellipsoidal Toolbox [19]. Indeed the signed distance is not supported by the toolbox: it returns the distance zero when the ellipsoids overlap.

Solving the non-convex optimization problem to compute the signed distance in the overlapping case is nontrivial, and to our knowledge there is no method that is guaranteed to provide a globally optimal solution. The goal of this paper is to develop such an algorithm. Our approach is to introduce a dual problem of (1.1) with zero duality gap, and find all the points (x_1, x_2) that satisfy the Karush-Kuhn-Tucker (KKT) conditions for its relaxation. We then identify a relevant KKT point with the smallest dual objective value. The Lagrange multipliers corresponding to KKT points are characterized as the common zeros of two bivariate rational functions, whose poles are not known explicitly. A key step is to convert the rational equations to polynomial matrix equations by considering the determinants of certain matrix pencils. This results in solving a two-parameter quadratic eigenvalue problem, which can be done via solving single-parameter linear eigenvalue problems of squared size, for which reliable algorithms exist [16, 18, 23]. The main cost of our algorithm is in solving a $4n^2 \times 4n^2$ (generalized) eigenvalue problem, and the overall complexity is $O(n^6)$.

Since computing the signed distance is fundamentally more difficult in the overlapping case, it is recommended in practice to check in advance whether the two ellipsoids overlap or not. This can be done via a quadratically constrained quadratic programming problem. In the non-overlapping case, a convex optimization solver or a more specified geometric method [21] are more efficient than the method we describe and hence recommended. The focus of this paper is the non-convex case, where such efficient algorithms may fail to provide the global solution.

This paper is organized as follows. In Section 2, we introduce an equivalent dual problem of (1.1). Section 3 derives the KKT conditions for the primal (1.1) and its relaxed dual problem to give an outline of our algorithm. Section 4 discusses the solution of bivariate determinantal equations resulting from the KKT conditions and shows that our algorithm works in generic cases with $O(n^6)$ time. In Section 5, we analyze the degenerate case and show that our algorithm still works by employing certain preprocessing techniques. In Section 6, we describe how to check the local optimality of KKT points in the primal problem. Finally, we provide numerical experiments in Section 7 to illustrate the properties and performance of our algorithm.

2. Dual problem. Computing the signed distance by solving the optimization problem (1.1) directly via finding all the KKT points appears to be difficult due to the presence of the square roots in the objective functions. In this section, we show by taking the dual that computing the signed distance is equivalent to

$$\begin{aligned} \underset{x_1,x_2}{\text{minimize}} \quad & \|x_1 - x_2\| \\ \text{subject to} \quad & x_1 \in E_1, \quad x_2 \in E_2, \\ & x_1 - x_2 \in \operatorname{bd}(E_1 \ominus E_2). \end{aligned} \tag{2.1}$$

Here $\ominus$ designates the Minkowski difference and $\operatorname{bd}(U)$ indicates the boundary of a set U. For simplicity, defining $U = E_1 \ominus E_2$, we write (2.1) as

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \|x\| \\ \text{subject to} \quad & x \in \operatorname{bd}(U). \end{aligned} \tag{2.2}$$

Note that U is a convex set because E_1 and E_2 are convex.

The two ellipsoids are said to *overlap* if they share a common interior point. Testing whether E_1 and E_2 overlap or not can be done by solving the optimization

problem:

$$\begin{aligned} &\text{minimize} && \eta \\ &\text{subject to} && (x-b_1)^\top A_1^{-1}(x-b_1) \le \eta, \\ & && (x-b_2)^\top A_2^{-1}(x-b_2) \le \eta. \end{aligned} \tag{2.3}$$

The optimal value satisfying $\eta_* \ge 1$ indicates the non-overlapping case and $\eta_* < 1$ indicates the overlapping case. This is a convex quadratic optimization problem, and can be solved efficiently by an interior-point method [7].

When the ellipsoids E_1, E_2 have no overlap, U does not contain the origin in its interior. In this case, even if $\mathrm{bd}(U)$ is replaced by U in (2.2), the resulting problem (1.2) has an optimal solution on $\mathrm{bd}(U)$. Therefore solving the convex problem (1.2) is equivalent to computing the signed distance (1.1) when the two ellipsoids do not overlap. The non-overlapping case of (2.2) includes the special case where the ellipsoids E_1, E_2 are touching externally but do not overlap. In such a case, (1.2) has an unique optimal solution satisfying $x_1 = x_2$.

We mainly focus on the case where the ellipsoids E_1 and E_2 overlap. In this case, the origin lies inside the convex set U and (1.2) does not have an optimal solution on $\mathrm{bd}(U)$. Therefore, we need to deal with a non-convex optimization problem; the problem (2.2) is equivalent to finding the minimum distance from a point in a convex set U to the complement of U, for which a dual problem with strong duality is formulated by Briec [8], building upon the work of Nirenberg [26]. Following Briec [8], we introduce the dual problem of (2.1) as

$$\begin{aligned} &\underset{w}{\text{maximize}} && \min_{x \in U} x^\top w \\ &\text{subject to} && \|w\| = 1. \end{aligned} \tag{2.4}$$

Since $E_1 = \{b_1 + A_1^{1/2} u_1 : \|u_1\| \le 1\}$ and $E_2 = \{b_2 + A_2^{1/2} u_2 : \|u_2\| \le 1\}$, the optimal objective value of (2.4) is equivalent to

$$\max_{w:\|w\|=1} \min_{\|u_1\|,\|u_2\| \le 1} \left(w^\top (b_1 - b_2) + w^\top (A_1^{1/2} u_1 - A_2^{1/2} u_2) \right). \tag{2.5}$$

Regarding the minimization with respect to u_1 and u_2, the first term $w^\top(b_1 - b_2)$ is independent of u_1 and u_2, so the minimum is clearly attained when u_1 and u_2 are given by

$$u_1 = -\frac{A_1^{1/2} w}{\|A_1^{1/2} w\|}, \quad u_2 = \frac{A_2^{1/2} w}{\|A_2^{1/2} w\|}.$$

We substitute this into (2.5) to obtain (1.1). Thus (2.1) is the dual problem of computing the signed distance (1.1).

3. Finding the KKT points. Solving the dual problem (2.1) directly is highly nontrivial. In fact, it is not an easy task to write down the KKT conditions, because no explicit formula for $\mathrm{bd}(U)$ is available.

We overcome this difficulty by introducing the following relaxation problem:

$$\begin{aligned} &\underset{x_1, x_2}{\text{minimize}} && \|x_1 - x_2\| \\ &\text{subject to} && x_1 \in \mathrm{bd}(E_1), \quad x_2 \in \mathrm{bd}(E_2). \end{aligned} \tag{3.1}$$

We find all the KKT points of (3.1), as described in Section 3.1. From among the obtained KKT points (x_1, x_2), we select those having $x_1 - x_2$ on the boundary of U as we explain in Section 3.2. It will be shown in Section 3.2 that these selected KKT points of (3.1) correspond to the KKT points of the primal problem (1.1). We can then compute the signed distance by comparing the objective values.

3.1. KKT conditions for relaxation. In this section, we present how to find the KKT points of the relaxation problem (3.1), which can be recast as

$$\begin{aligned} &\underset{x_1,x_2}{\text{minimize}} \quad \|x_1 - x_2\|^2 \\ &\text{subject to} \\ &(x_1 - b_1)^\top A_1^{-1}(x_1 - b_1) = 1, \qquad (3.2) \\ &(x_2 - b_2)^\top A_2^{-1}(x_2 - b_2) = 1. \qquad (3.3) \end{aligned}$$

The KKT conditions for the problem consist of these feasibility constraints and

$$x_1 - x_2 = \lambda A_1^{-1}(x_1 - b_1), \qquad (3.4)$$

$$x_2 - x_1 = \mu A_2^{-1}(x_2 - b_2), \qquad (3.5)$$

where λ and μ are Lagrange multipliers. These conditions imply

$$\mu(x_2 - b_2) = -\lambda A_2 {A_1}^{-1}(x_1 - b_1), \qquad (3.6)$$

$$\lambda(x_1 - b_1) = -\mu A_1 {A_2}^{-1}(x_2 - b_2). \qquad (3.7)$$

Combining (3.4) and (3.6), we obtain

$$\mu x_1 - \mu b_2 = \lambda\mu {A_1}^{-1}(x_1 - b_1) - \lambda A_2 {A_1}^{-1}(x_1 - b_1).$$

Similarly, from (3.5) and (3.7), we obtain

$$\lambda x_2 - \lambda b_1 = \lambda\mu {A_2}^{-1}(x_2 - b_2) - \mu A_1 {A_2}^{-1}(x_2 - b_2).$$

Defining

$$N(\lambda, \mu) = \lambda\mu I - \mu A_1 - \lambda A_2,$$

we can write the two equalities as

$$\mu(b_1 - b_2) = N(\lambda, \mu){A_1}^{-1}(x_1 - b_1), \qquad (3.8)$$

$$\lambda(b_2 - b_1) = N(\lambda, \mu){A_2}^{-1}(x_2 - b_2). \qquad (3.9)$$

The two equations (3.8) and (3.9) can be solved for x_1, x_2 to express them in terms of λ, μ, which can be substituted into (3.2), (3.3) to yield two bivariate rational equations with respect to λ, μ. However, solving them for λ, μ is difficult as the poles of the rational equations are unknown.

Instead, we formulate a pair of matrix equations that provide the appropriate Lagrange multipliers: we introduce a pair of matrices $M_1(\lambda, \mu)$ and $M_2(\lambda, \mu)$ defined by

$$M_1(\lambda, \mu) = \begin{bmatrix} \mu A_1 & N(\lambda, \mu) \\ N(\lambda, \mu) & \mu D \end{bmatrix}, \quad M_2(\lambda, \mu) = \begin{bmatrix} \lambda A_2 & N(\lambda, \mu) \\ N(\lambda, \mu) & \lambda D \end{bmatrix}, \qquad (3.10)$$

where $D = dd^\top$, $d := b_1 - b_2$.

LEMMA 3.1. *For every pair of* $x_1 \in \mathrm{bd}(E_1)$ *and* $x_2 \in \mathrm{bd}(E_2)$ *that satisfies the KKT conditions with multipliers* λ *and* μ*, we have* $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$.

Proof. By (3.2)–(3.5), if one of λ or μ is zero, then so is the other. In this case, $M_1(\lambda,\mu) = M_2(\lambda,\mu) = O$. Hence we may assume that neither λ nor μ is zero.

By the conditions (3.8) and (3.9), we see that $d = b_1 - b_2$ must belong to $\mathrm{Im}\, N(\lambda,\mu)$. Therefore, if $N(\lambda,\mu)$ is singular, then we have $\mathrm{rank}\begin{bmatrix} D & N(\lambda,\mu)\end{bmatrix} = \mathrm{rank}\begin{bmatrix} d & N(\lambda,\mu)\end{bmatrix} < n$, which implies that both $M_1(\lambda,\mu)$ and $M_2(\lambda,\mu)$ are singular.

Now suppose that $N(\lambda,\mu)$ is nonsingular. Note that

$$M_1'(\lambda,\mu) = \begin{bmatrix} \mu A_1 & N(\lambda,\mu) & O \\ N(\lambda,\mu) & O & \mu d \\ O & \mu d^\top & -\mu \end{bmatrix}, \quad M_2'(\lambda,\mu) = \begin{bmatrix} \lambda A_2 & N(\lambda,\mu) & O \\ N(\lambda,\mu) & O & \lambda d \\ O & \lambda d^\top & -\lambda \end{bmatrix}$$

satisfy $\det M_1'(\lambda,\mu) = -\mu \det M_1(\lambda,\mu)$ and $\det M_2'(\lambda,\mu) = -\lambda \det M_2(\lambda,\mu)$. Since

$$\begin{bmatrix} \mu A_1 & N(\lambda,\mu) \\ N(\lambda,\mu) & O \end{bmatrix}^{-1} = \begin{bmatrix} O & N(\lambda,\mu)^{-1} \\ N(\lambda,\mu)^{-1} & -\mu N(\lambda,\mu)^{-1} A_1 N(\lambda,\mu)^{-1} \end{bmatrix},$$

we also have

$$\det M_1'(\lambda,\mu) = (-1)^n \det N(\lambda,\mu)^2 \{\mu^3 d^\top N(\lambda,\mu)^{-1} A_1 N(\lambda,\mu)^{-1} d - \mu\}.$$

Thus we obtain

$$\begin{aligned} \det M_1(\lambda,\mu) &= (-1)^{n-1} \det N(\lambda,\mu)^2 \{\mu^2 d^\top N(\lambda,\mu)^{-1} A_1 N(\lambda,\mu)^{-1} d - 1\} \\ &= (-1)^{n-1} \det N(\lambda,\mu)^2 \{(x_1 - b_1)^\top {A_1}^{-1} (x_1 - b_1) - 1\}, \end{aligned}$$

where the last equality follows from (3.8). Similarly, using (3.9), we obtain

$$\begin{aligned} \det M_2(\lambda,\mu) &= (-1)^{n-1} \det N(\lambda,\mu)^2 \{\lambda^2 d^\top N(\lambda,\mu)^{-1} A_2 N(\lambda,\mu)^{-1} - 1\} \\ &= (-1)^{n-1} \det N(\lambda,\mu)^2 \{(x_2 - b_2)^\top {A_2}^{-1} (x_2 - b_2) - 1\}. \end{aligned}$$

It then follows from (3.2) and (3.3) that $\det M_1(\lambda,\mu) = 0$ and $\det M_2(\lambda,\mu) = 0$. □

Lemma 3.1 suggests computing all possible pairs of Lagrange multipliers λ and μ for the KKT points by solving the bivariate determinantal equations

$$\det(M_1(\lambda,\mu)) = \det(M_2(\lambda,\mu)) = 0. \tag{3.11}$$

We will discuss how to solve (3.11) in Section 4. In practice, the computed pairs (λ,μ) may contain nonreal pairs and those at infinity. These can immediately be removed from consideration because (λ,μ) must be real at the KKT points.

For each pair of λ and μ thus obtained, one can compute x_1 and x_2 by solving the system of linear equations (3.8) and (3.9). If $N(\lambda,\mu)$ is nonsingular, then x_1 and x_2 are uniquely determined, and they naturally satisfy the feasibility conditions (3.2) and (3.3). If $N(\lambda,\mu)$ is singular, among all the solutions of (3.8) and (3.9), we select those that satisfy (3.2) and (3.3), and verify that (3.4) and (3.5) hold. This uniquely determines x_1, x_2 if the null space of $N(\lambda,\mu)$ is 1-dimensional. If $\dim \mathrm{Ker}\, N(\lambda,\mu) > 1$, then we have infinitely many solutions of x_1, x_2 satisfying (3.2) and (3.3). Specifically, denoting by N_0 the null space of $N(\lambda,\mu)$, they can be written as $x_1 = x_{1,*} + N_0 v_1$, $x_2 = x_{2,*} + N_0 v_2$, where $x_{1,*}, x_{2,*}$ are any vectors satisfying (3.8) and (3.9); for example the least-squares solution. We then solve (3.4) and (3.5) for v_1, v_2.

3.2. KKT points for the primal problem. The KKT conditions for the primal problem (1.1) consist of the feasibility constraint $w^\top w = 1$ and

$$-(b_1 - b_2) + \frac{A_1 w}{\sqrt{w^\top A_1 w}} + \frac{A_2 w}{\sqrt{w^\top A_2 w}} = \varphi w, \tag{3.12}$$

where φ is a Lagrange multiplier. Then the objective value of (1.1) at a KKT point w coincides with

$$w^\top (b_1 - b_2) - \sqrt{w^\top A_1 w} - \sqrt{w^\top A_2 w} = -\varphi. \tag{3.13}$$

Given a KKT point of (1.1), one can obtain a KKT point of (3.1) as follows.

LEMMA 3.2. *Suppose that w is a KKT point of* (1.1) *with a Lagrange multiplier φ. Then defining*

$$x_1 := b_1 - \frac{A_1 w}{\sqrt{w^\top A_1 w}} \quad \lambda := \varphi\sqrt{w^\top A_1 w}, \tag{3.14}$$

$$x_2 := b_2 + \frac{A_2 w}{\sqrt{w^\top A_2 w}}, \quad \mu := \varphi\sqrt{w^\top A_2 w}, \tag{3.15}$$

we obtain a KKT point (x_1, x_2) of (3.1) *with Lagrange multipliers λ, μ.*

Proof. It is easy to see from (3.14) and (3.15) that x_1 and x_2 satisfy the feasibility constraints (3.2) and (3.3), respectively. By (3.14), we have

$$\lambda A_1^{-1}(x_1 - b_1) = -\frac{\lambda w}{\sqrt{w^\top A_1 w}} = -\varphi w.$$

Similarly, by (3.15), we also have

$$\mu A_2^{-1}(x_2 - b_2) = \frac{\mu w}{\sqrt{w^\top A_2 w}} = \varphi w.$$

It follows from (3.12), (3.14) and (3.15) that $x_1 - x_2 = -\varphi w$ holds. Hence x_1 and x_2 satisfy (3.4) and (3.5). ∎

Note that the Lagrange multipliers λ and μ thus obtained must both have the same sign as φ. Intuitively, such a KKT point (x_1, x_2) with Lagrange multipliers having the same sign places $x_1 - x_2$ on the boundary of $E_1 \ominus E_2$, because at such points the gradients $2A_1^{-1}(x_1 - b_1)$ of (3.2) and $2A_2^{-1}(x_2 - b_2)$ of (3.3) point towards exactly the opposite directions.

We now intend to claim conversely that any KKT point of (3.1) with nonzero Lagrange multipliers with the same signs leads to a KKT point of (1.1).

LEMMA 3.3. *Suppose (x_1, x_2) is a KKT point of the relaxed dual problem* (3.1) *with nonzero Lagrange multipliers λ, μ having the same sign, say $\sigma \in \{+, -\}$. Then defining*

$$w := -\frac{\sigma(x_1 - x_2)}{\|x_1 - x_2\|}, \quad \varphi := \sigma\|x_1 - x_2\|, \tag{3.16}$$

we obtain a KKT point w of (1.1) *with a Lagrange multiplier φ.*

Proof. The definitions (3.16) of w and φ imply that $x_1 - x_2 = -\varphi w$. By (3.4) and (3.2), we have

$$\begin{aligned}\frac{A_1 w}{\sqrt{w^\top A_1 w}} &= -\frac{\sigma A_1(x_1 - x_2)}{\sqrt{(x_1 - x_2)^\top A_1 (x_1 - x_2)}} \\ &= -\frac{\sigma\lambda(x_1 - b_1)}{|\lambda|\sqrt{(x_1 - b_1)^\top A_1^{-1}(x_1 - b_1)}} = -(x_1 - b_1).\end{aligned}$$

Similarly, by (3.5) and (3.3), we also have

$$\begin{aligned}\frac{A_2 w}{\sqrt{w^\top A_2 w}} &= \frac{\sigma A_2(x_2 - x_1)}{\sqrt{(x_2 - x_1)^\top A_2 (x_2 - x_1)}} \\ &= \frac{\sigma\mu(x_2 - b_2)}{|\mu|\sqrt{(x_2 - b_2)^\top A_2^{-1}(x_2 - b_2)}} = x_2 - b_2.\end{aligned}$$

Combining these equalities, one can easily verify that w and φ satisfy (3.12). □

It will be demonstrated later in Section 7 that (3.1) admits KKT points with Lagrange multipliers of different signs. Such KKT points correspond to interior points of $E_1 \ominus E_2$, which are infeasible in the original dual problem (2.1).

For a KKT point (x_1, x_2) with Lagrange multipliers λ, μ, if one of λ or μ is zero, then so is the other by (3.6) and (3.7). In this case, $x_1 = x_2$ must hold. Conversely, any point $x \in \mathrm{bd}(E_1) \cap \mathrm{bd}(E_2)$ provides a KKT point $(x_1, x_2) := (x, x)$ with Lagrange multipliers $\lambda, \mu := 0$. Thus, in the overlapping case, the relaxed dual problem (3.1) has infinitely many KKT points with zero Lagrange multipliers, but they are irrelevant to the signed distance problem. Fortunately, such KKT points are removed automatically in the algorithm we describe in Section 4.1.

If (3.1) has a KKT point (x_1, x_2) with Lagrange multipliers λ, μ both negative, then the objective value at the corresponding KKT point w of (3.1) must be positive, which means the signed distance is positive. This can happen only if the two ellipsoids are disjoint.

Conversely, since the signed distance between two non-overlapping ellipsoids must be nonnegative, (3.13) and Lemma 3.2 imply that there is a KKT point (x_1, x_2) for (3.1) with Lagrange multipliers $\lambda, \mu \leq 0$. Note that (x_1, x_2) remains to be a KKT point with the same Lagrange multipliers if we replace the feasibility constraints by $x_1 \in E_1$ and $x_2 \in E_2$ (including the interior). The resulting convex optimization problem (1.2) has a unique optimal solution. Since each KKT point in a convex optimization problem provides an optimal solution, this establishes the uniqueness of the KKT point for (3.1) with nonpositive Lagrange multipliers with the same signs in the non-overlapping case, and the signed distance is obtained by $\mathrm{dist}(E_1, E_2) := \|x_1 - x_2\|$.

3.3. Outline of the algorithm. For computing the signed distance between overlapping ellipsoids, Lemmas 3.2 and 3.3 together with (3.13) suggest finding all the KKT points (x_1, x_2) for (3.1) with positive Lagrange multipliers and choosing the one that attains the minimum value of $\|x_1 - x_2\|$ among them. In fact, Lemma 3.3 and (3.13) ensure that the selected (x_1, x_2) provides a KKT point for (1.1) with the objective value equal to $-\|x_1 - x_2\|$. On the other hand, it follows from Lemma 3.2 and (3.13) that any KKT point for (1.1) with the objective value $-\varphi$ yields a KKT point (x_1, x_2) for (3.1) with positive Lagrange multipliers such that $\|x_1 - x_2\| = \varphi$.

This means if there were a better KKT point for the primal problem, then there should have been a better one in the relaxed dual than the selected one. Thus the algorithm computes the signed distance correctly in the overlapping case.

We now summarize the whole algorithm for computing the signed distance by a pseudocode.

Algorithm 3.1 Compute the signed distance between two ellipsoids E_1 and E_2.

1: Test whether E_1 and E_2 overlap via the problem (2.3). If they do not overlap, compute the distance as a convex optimization problem [21].
2: Solve the bivariate determinantal equations (3.11).
3: Among the pairs with $\lambda > 0, \mu > 0$, find the corresponding x_1, x_2.
4: For each pair obtained in Step 3, compute and compare $\|x_1 - x_2\|$. The pair with the smallest one gives the signed distance by $\operatorname{dist}(E_1, E_2) := -\|x_1 - x_2\|$.

4. Solving the bivariate determinantal equations. We now discuss how to solve the bivariate determinantal equations (3.11) for λ and μ. The algorithm is shown to run in $O(n^6)$ time, which is the dominant cost in Algorithm 3.1.

4.1. Reduction to univariate linear eigenvalue problems. Observe that (3.11) is a quadratic two-parameter eigenvalue problem, for which recent studies describe numerical solutions by reducing it to a single-parameter eigenvalue problem [16, 18, 23]. In our case, (3.11) has a special form that can be expressed as

$$\det(F_{00} + \lambda F_{10} + \mu F_{01} + \lambda\mu F_{11}) = 0, \tag{4.1}$$

$$\det(G_{00} + \lambda G_{10} + \mu G_{01} + \lambda\mu G_{11}) = 0. \tag{4.2}$$

In the general quadratic problem there are two additional terms $F_{20}\lambda^2 + F_{02}\mu^2$ in (4.1) and $G_{20}\lambda^2 + G_{02}\mu^2$ in (4.2), but in our case we do not have them. We can take advantage of this structure to reduce the computational cost: following [18], the $2n \times 2n$ quadratic two-parameter eigenvalue problem (4.1), (4.2) can be solved via two $(2n)^2 \times (2n)^2$ quadratic single-parameter eigenvalue problems

$$\begin{aligned} \det Q_1(\lambda) := \det\big[&\lambda^2(F_{11} \otimes G_{10} - F_{10} \otimes G_{11}) + (F_{01} \otimes G_{00} - F_{00} \otimes G_{01}) \\ &+ \lambda(F_{11} \otimes G_{00} - F_{00} \otimes G_{11} - F_{10} \otimes G_{01} + F_{01} \otimes G_{10})\big] = 0, \end{aligned} \tag{4.3}$$

$$\begin{aligned} \det Q_2(\mu) := \det\big[&\mu^2(F_{11} \otimes G_{01} - F_{01} \otimes G_{11}) + (F_{10} \otimes G_{00} - F_{00} \otimes G_{10}) \\ &+ \mu(F_{11} \otimes G_{00} - F_{00} \otimes G_{11} + F_{10} \otimes G_{01} - F_{01} \otimes G_{10})\big] = 0. \end{aligned} \tag{4.4}$$

The solutions (λ, μ) for (3.11) match those of (4.3) and (4.4). Moreover, if the eigenvectors of (3.11) are v_1 and v_2, i.e., $M_1(\lambda, \mu)v_1 = 0$ and $M_2(\lambda, \mu)v_2 = 0$, then those of (4.3) and (4.4) are $v_1 \otimes v_2$, with eigenvalues λ and μ, respectively [18]. That is, each solution of (4.3) is paired with one solution of (4.4). In our case, defining $J = \left[\begin{smallmatrix} O & I \\ I & O \end{smallmatrix}\right]$, we have

$$F_{00} = O,\ F_{10} = \begin{bmatrix} O & -A_2 \\ -A_2 & O \end{bmatrix},\ F_{01} = \begin{bmatrix} A_1 & -A_1 \\ -A_1 & D \end{bmatrix},\ F_{11} = J, \tag{4.5}$$

$$G_{00} = O,\ G_{10} = \begin{bmatrix} A_2 & -A_2 \\ -A_2 & D \end{bmatrix},\ G_{01} = \begin{bmatrix} O & -A_1 \\ -A_1 & O \end{bmatrix},\ G_{11} = J. \tag{4.6}$$

Generally, we can solve the single-parameter quadratic eigenvalue problem (4.3), (4.4) by linearization, which results in a linear (generalized) eigenvalue problem of doubled size. However, here we have $F_{00} = G_{00} = O$, which can be taken advantage of. All the terms involving $F_{00,}G_{00}$ in (4.3) and (4.4) are zero, and in particular the constant terms $(F_{01} \otimes G_{00} - F_{00} \otimes G_{01})$ and $(F_{10} \otimes G_{00} - F_{00} \otimes G_{10})$ are both zero. It follows that instead of solving the quadratic eigenvalue problems, we can simply solve the $4n^2 \times 4n^2$ linear generalized eigenvalue problems

$$\det L_1(\lambda) := \det\big[\lambda(F_{11} \otimes G_{10} - F_{10} \otimes G_{11}) + (F_{01} \otimes G_{10} - F_{10} \otimes G_{01})\big] = 0, \tag{4.7}$$

$$\det L_2(\mu) := \det\big[\mu(F_{11} \otimes G_{01} - F_{01} \otimes G_{11}) + (F_{10} \otimes G_{01} - F_{01} \otimes G_{10})\big] = 0. \tag{4.8}$$

This problem is twice smaller than the $8n^2 \times 8n^2$ eigenvalue problem one obtains by reducing the quadratic eigenvalue problems (4.3), (4.4) by the standard approach of linearization [30]. This size reduction has the effect of removing the "trivial" solutions $\lambda = \mu = 0$ for $\det M_1 = \det M_2 = 0$.

4.2. Connection with Bézoutians. Here we mention a connection of the above process to Bézoutians, which is also used later. Forming Q_1, Q_2 in (4.3), (4.4) from M_1, M_2 is related to constructing the Bézoutians for matrix polynomials [28]. Specifically, this conversion from a two-parameter eigenvalue problem to a single-parameter eigenvalue problem is equivalent to taking the Bézoutian for the two matrix polynomials $I_{2n} \otimes M_1(\lambda, \mu)$ and $M_2(\lambda, \mu) \otimes I_{2n}$, which are of size $4n^2 \times 4n^2$. Here the Kronecker products are taken to achieve commutativity, which makes it convenient for forming the Bézoutian for matrix polynomials [20].

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

$$\mathcal{B}(s,t) = \frac{P_1(s)P_2(t) - P_2(s)P_1(t)}{s-t} = \sum_{i,j=0}^{k-1,k-1} B_{ij} s^i t^j \tag{4.9}$$

in s and t. Here B_{ij} 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 4.1 ([20, Theorem 1.1]). *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 [13, 20].

We now claim that $\det(Q_1(\lambda)) = 0$ in (4.3) solves the equation $\det(B(\lambda)) = 0$ for λ, where $B(\lambda)$ is the Bézout matrix between $P_1(\mu) := I_{2n} \otimes M_2(\lambda, \mu)$ and $P_2(\mu) := M_1(\lambda, \mu) \otimes I_{2n}$ with respect to μ. More specifically, define the Bézoutian

$$\mathcal{B}(s,t) = \frac{1}{s-t}\big((I_{2n} \otimes M_2(\lambda, s))(M_1(\lambda, t) \otimes I_{2n}) - (M_1(\lambda, s) \otimes I_{2n})(I_{2n} \otimes M_2(\lambda, t))\big).$$

Then writing $\mathcal{B}(s,t) = \sum_{i=0}^{1}\sum_{j=0}^{1} s^i t^j B_{ij}$ we have $Q_1(\lambda) = B_{10} = B_{01}$ (which is a matrix polynomial in λ), the coefficient matrix of $\mathcal{B}(s,t)$ with linear terms in s, or equivalently t. Noting that $B_{11} = O$, the Bézout matrix is

$$B(\lambda) = \begin{bmatrix} B_{00} & Q_1(\lambda) \\ Q_1(\lambda) & O \end{bmatrix}, \tag{4.10}$$

so solving $\det Q_1(\lambda) = 0$ is equivalent to finding the values of $\lambda = \lambda_*$ for which the Bézout matrix is singular. By Lemma 4.1, this is in turn equivalent to finding λ_* such that $(I_{2n} \otimes M_2(\lambda_*, \mu_*))v = (M_1(\lambda_*, \mu_*) \otimes I_{2n})v = 0$ for some μ_* and $v \neq 0$. The discussion for $Q_2(\mu)$ is completely analogous, in which we start with the bivariate matrix polynomial

$$\mathcal{B}(s,t) = \frac{1}{s-t}\big((I_{2n} \otimes M_2(s,\mu))(M_1(t,\mu) \otimes I_{2n}) - (M_1(s,\mu) \otimes I_{2n})(I_{2n} \otimes M_2(t,\mu))\big).$$

This process is a generalization from scalars to matrices of solving a pair of bivariate polynomial equations via Bézoutians [5, 24].

4.3. The number of solutions is finite in the generic case. For the above process to correctly find the KKT points, we need to show that the solution to (3.11) is zero-dimensional, that is, the number of pairs (λ, μ) satisfying (3.11) is finite. In fact, if we can show it is finite, then from the linear eigenvalue problems (4.7) and (4.8) (or by Bézout's theorem), we see that it is in fact bounded by $4n^2$.

Lemma 3.1 shows that every pair (λ_*, μ_*) corresponding to a KKT point satisfies $\det M_1(\lambda_*, \mu_*) = \det M_2(\lambda_*, \mu_*) = 0$. In addition, if $M_1(\lambda, \mu)v_1 = 0$ and $M_2(\lambda, \mu)v_2 = 0$, then $L_1(\lambda)(v_1 \otimes v_2) = L_2(\mu)(v_1 \otimes v_2) = 0$ ([16], see also Section 5.1). Thus it suffices to show that the matrix pencils $L_1(\lambda), L_2(\mu)$ are regular. We now claim that generically this property holds. First we make precise the meaning of "generic" in our problem.

Definition 4.2. *We call the signed distance problem with* (A_1, A_2, b_1, b_2) generic *if at least one of* A_1 *and* A_2 *has no eigenvector orthogonal to* $d = b_1 - b_2$.

Calling this generic is justified because the property holds with probability 1 for randomly generated positive definite matrices A_1, A_2 as long as $d \neq 0$.

We will need the following lemma.

Lemma 4.3. *Let* X *be a real matrix in the form*

$$X = \begin{bmatrix} \Lambda & \Psi \\ \Psi & dd^\top \end{bmatrix}, \tag{4.11}$$

where d *is a vector,* Λ, Ψ *are* $n \times n$ *diagonal,* Λ *is nonsingular and* Ψ *is singular. Then* X *is nonsingular if and only if there is no null vector of* Ψ *that is orthogonal to* d.

Proof. Defining

$$\hat{X} = \begin{bmatrix} \Lambda & \Psi & O \\ \Psi & O & d \\ O & d^\top & -1 \end{bmatrix},$$

we have $\det \hat{X} = -\det X$, so it suffices to examine when $\hat{X}$ is singular. Recalling that Λ and Ψ are both diagonal with Λ nonsingular, by the determinant expansion we easily see that $\hat{X}$ is singular if $\Psi_{ii} = d_i = 0$ for some i. Similarly, $\hat{X}$ is singular if Ψ has two or more diagonal elements equal to zero. It remains to treat the case where Ψ

has exactly one zero diagonal element $\Psi_{ii} = 0$ with the corresponding element $d_i \neq 0$. In this case there is only one nonzero term in the Leibniz determinant formula, which is indeed nonzero, hence $\det \hat{X} \neq 0$. □

We now show that L_1, L_2 are regular pencils for generic (A_1, A_2, b_1, b_2).

THEOREM 4.4. *The matrix pencils $L_1(\lambda), L_2(\mu)$ in (4.7), (4.8) are regular, i.e., $\det L_1(\lambda) \neq 0$ and $\det L_2(\mu) \neq 0$ for some (λ, μ), if at least one of A_1 and A_2 has no eigenvector orthogonal to d.*

Proof. We write $L_1(\lambda) = \lambda X_1 + Y_1$, $L_2(\mu) = \mu X_2 + Y_2$ and proceed by examining the linear terms X_1, X_2, whose nonsingularity is a sufficient condition for $L_1(\lambda), L_2(\mu)$ to be regular. We have

$$X_1 = (F_{11} \otimes G_{10} - F_{10} \otimes G_{11}) = J \otimes \begin{bmatrix} A_2 & -A_2 \\ -A_2 & D \end{bmatrix} - \begin{bmatrix} O & -A_2 \\ -A_2 & O \end{bmatrix} \otimes J.$$

To simplify the expression, we reformulate the problem in the coordinate system determined by the eigendecomposition $A_2 = Q\Lambda_2 Q^\top$, in which $\widetilde{A}_1 := Q^\top A_1 Q$, $\widetilde{b}_1 = Q^\top b_1$, $\widetilde{b}_2 = Q^\top b_2$, $\tilde{d} = \widetilde{b}_1 - \widetilde{b}_2$, and $\widetilde{D} := Q^\top D Q = \tilde{d}\tilde{d}^\top$. Then we define the matrix $\widetilde{X}_1$ unitarily similar to X_1 by

$$\widetilde{X}_1 = (\widetilde{F}_{11} \otimes \widetilde{G}_{10} - \widetilde{F}_{10} \otimes \widetilde{G}_{11}) = J \otimes \begin{bmatrix} \Lambda_2 & -\Lambda_2 \\ -\Lambda_2 & \widetilde{D} \end{bmatrix} - \begin{bmatrix} O & -\Lambda_2 \\ -\Lambda_2 & O \end{bmatrix} \otimes J.$$

We can write this as

$$\widetilde{X}_1 = \begin{bmatrix} O_{2n^2 \times 2n^2} & \widetilde{X} \\ \widetilde{X} & O_{2n^2 \times 2n^2} \end{bmatrix}, \quad \widetilde{X} = \mathrm{diag}(\widetilde{G}_{10} + \lambda_1 J, \widetilde{G}_{10} + \lambda_2 J, \ldots, \widetilde{G}_{10} + \lambda_n J),$$

where λ_i is the ith eigenvalue of A_2. It follows that $\widetilde{X}_1$ is singular if and only if $\widetilde{G}_{10} + \lambda_i J$ is singular for some i. For each i we have

$$\widetilde{G}_{10} + \lambda_i J = \begin{bmatrix} \Lambda_2 & \lambda_i I - \Lambda_2 \\ \lambda_i I - \Lambda_2 & \widetilde{d}\widetilde{d}^\top \end{bmatrix}.$$

Note that the off-diagonal blocks have a zero diagonal in the (i, i) entry.

By Lemma 4.3, $(\widetilde{G}_{10} + \lambda_i J)$ is singular if and only if $\widetilde{d}_i = 0$. Since the argument holds for every i, we conclude that X_1 is singular if and only if $\widetilde{d}_i = 0$ for some i. In the original coordinate system, this condition asserts that A_2 has an eigenvector that is orthogonal to d. An analogous argument shows that X_2 is singular if and only if A_1 has an eigenvector that is orthogonal to d.

When just one of A_1 and A_2 has an eigenvector v orthogonal to d, one of L_1 or L_2 has a nonsingular linear coefficient, so we can compute $4n^2$ values of λ (or μ) and solve $\det M_1 = \det M_2 = 0$ for the other variable. Note that for λ fixed to any value, $\det M_1$ and $\det M_2$ are not identically zero as μ varies. Hence $\det M_1 = \det M_2 = 0$ yields a pair of regular eigenvalue problems. Hence for each λ there are finitely many solutions. It follows that $\det M_1 = \det M_2 = 0$ has a finite number of solutions, and hence L_1, L_2 are both regular. □

We have thus proved that L_1 and L_2 are regular matrix pencils in the generic case. It remains to consider the situation where both A_1 and A_2 have eigenvectors that are orthogonal to d.

Note that indeed L_1 and L_2 can be singular. Suppose that A_1 and A_2 share an eigenvector v orthogonal to d. For simplicity, we consider $\hat{M}_1 := \left[\begin{smallmatrix} Q & O \\ O & Q \end{smallmatrix}\right]^\top M_1 \left[\begin{smallmatrix} Q & O \\ O & Q \end{smallmatrix}\right]$, $\hat{M}_2 := \left[\begin{smallmatrix} Q & O \\ O & Q \end{smallmatrix}\right]^\top M_2 \left[\begin{smallmatrix} Q & O \\ O & Q \end{smallmatrix}\right]$, where Q is an orthogonal matrix whose last column is v. Then we see that for every λ, there exists μ such that $\hat{M}_1, \hat{M}_2$ both have a zero column in the second block columns, which shows they are clearly singular. In such cases some remedy is needed for the algorithm to compute the signed distance correctly, as we describe in Section 5.

4.4. Alternative formulation for M_1, M_2 that avoids the cross term $\lambda\mu$. We briefly discuss another formulation of bivariate matrix pencils to obtain the KKT points. In the above derivation, the presence of the cross term $\lambda\mu$ made the problem quadratic instead of linear. In order to avoid the cross term $\lambda\mu$, one can construct a pair of larger matrices

$$K_1(\lambda,\mu) = \begin{bmatrix} A_1 & O & \lambda I - A_1 & \lambda I \\ O & O & A_2 & \mu I \\ \lambda I - A_1 & A_2 & D & O \\ \lambda I & \mu I & O & O \end{bmatrix}, \tag{4.12}$$

$$K_2(\lambda,\mu) = \begin{bmatrix} A_2 & O & \mu I - A_2 & \mu I \\ O & O & A_1 & \lambda I \\ \mu I - A_2 & A_1 & D & O \\ \mu I & \lambda I & O & O \end{bmatrix}. \tag{4.13}$$

Then $\det K_j(\lambda,\mu) = \det M_j(\lambda,\mu)$ holds for $j = 1, 2$. The crux is that solving the equation $\det K_1(\lambda,\mu) = \det K_2(\lambda,\mu) = 0$ for λ, μ is a *linear* two-parameter eigenvalue problem, as opposed to the quadratic problem (4.3)–(4.4) with (4.5)–(4.6).

4.5. Complexity analysis. We have presented two possible methods for solving $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$; one solves $\det L_1(\lambda) = 0, \det L_2(\mu) = 0$, and the other $\det K_1(\lambda,\mu) = 0, \det K_2(\lambda,\mu) = 0$. Let us compare the computational costs in terms of the size of the resulting linear generalized eigenvalue problems.

1. The first approach requires to solve the linear generalized eigenvalue problems $\det L_1(\lambda) = 0$, $\det L_2(\mu) = 0$ of size $(2n)^2 = 4n^2$. 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 [14, §7.7.7], the computational cost is about $30(4n^2)^3 \approx (1.9 \times 10^3)n^6$ flops.
   Since a ballpark estimate of a feasible arithmetic cost for a current standard desktop machine is about 10^{13} flops, a realistic limit for this method is $n \lesssim 50$.
2. Solving an $N \times N$ linear two-parameter eigenvalue problem can be done by reducing the problem to an $N^2 \times N^2$ linear generalized eigenvalue problem formed as Kronecker products of the coefficient matrices [2]. Since the matrices in $\det K_1(\lambda,\mu) = 0, \det K_2(\lambda,\mu) = 0$ in (4.12), (4.13) are of size $4n$, the resulting linear generalized eigenvalue problems are of size $(4n)^2 = 16n^2$. With this approach a realistic limit is $n \lesssim 25$.

We conclude that the first approach of solving the quadratic problem (4.5), (4.6) via $\det L_1(\lambda) = 0, \det L_2(\mu) = 0$ is more efficient. Moreover, one can show that the linear generalized eigenvalue problems resulting from (4.12), (4.13) are generically singular, unlike (4.7) and (4.8) which are generically regular. For these reasons, we henceforth focus only on solving (3.11) via (4.7) and (4.8).

5. Degenerate case analysis. Algorithm 3.1 successfully computes the signed distance for almost all pairs of E_1 and E_2, including virtually any randomly generated example. However, there remain cases where the matrix pencils L_1, L_2 are singular, and so the algorithm may fail to compute all the KKT points. In this section, we analyze such situations and show that Algorithm 3.1 can still be used with appropriate treatments to compute the signed distance reliably.

5.1. Relation between M_j and L_j. We first prove that the regularity of L_1, L_2 is precisely the condition for $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$ to have a finite number of solutions.

LEMMA 5.1. *Let $M_1(\lambda,\mu), M_2(\lambda,\mu)$ and $L_1(\lambda), L_2(\mu)$ be as defined in* (3.10), (4.7), (4.8). *Then the pencils $L_1(\lambda)$ and $L_2(\mu)$ are regular if and only if the number of solutions (λ,μ) for which $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$ is finite.*

Moreover, if $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$ has infinitely many solutions, then for every fixed λ (or μ) there exists a value of μ (or λ) for which $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$.

Proof. The key is to use the connection to Bézoutians. As seen in Section 4.2, values of λ for which $\det Q_1(\lambda) = 0$ are those for which the Bézout matrix $B(\lambda)$ is singular between $P_1(\mu) = I_{2n} \otimes M_2(\lambda,\mu)$ and $P_2(\mu) = M_1(\lambda,\mu) \otimes I_{2n}$ viewed as matrix polynomials in μ, regarding λ as a hidden variable. Hence by Lemma 4.1, if $P_1(\mu), P_2(\mu)$ are regular matrix polynomials at $\lambda = \lambda_0$, they share an eigenpair (μ_0, x), that is, $P_1(\mu_0)x = P_2(\mu_0)x = 0$ for a nonzero $x \in \mathbb{R}^{4n^2}$, and in this case $L_1(\lambda_0)$ has the same null vector x. The discussion for L_2 is entirely analogous.

To invoke this result it is crucial to show that L_1, L_2 correspond to the Bézout matrices of two *regular* matrix polynomials. Now note that for any fixed value of λ, both $M_1(\lambda,\mu)$ and $M_2(\lambda,\mu)$ are regular matrix polynomials in μ, in that there exist values of μ for which they are nonsingular. It follows that $Q_1(\lambda)$ is always the off-diagonal part of the Bézout matrix (4.10) between two regular matrix polynomials $P_1(\mu), P_2(\mu)$, and since $L_1(\lambda) = \frac{Q(\lambda)}{\lambda}$, $L_1(\lambda_0)$ is a singular matrix if and only if $P_1(\mu), P_2(\mu)$ share an eigenpair (μ_0, x) at λ_0.

This condition is equivalent to $\det M_1(\lambda_0,\mu_0) = \det M_2(\lambda_0,\mu_0) = 0$. To see this, note that $P_1(\mu_0)$ is singular if and only if $M_2(\lambda_0,\mu_0)$ is singular, and if $M_2 v_2 = 0$ then $P_1(\mu)$ has null space $w \otimes v_2$ for any vector $w \in \mathbb{R}^{2n}$. Indeed, we have $P_1(\mu)(w \otimes v_2) = (I_{2n} \otimes M_2)(w \otimes v_2) = w \otimes M_2 v_2 = 0$. Similarly, $P_2(\mu)$ is singular if and only if $M_1(\lambda_0,\mu_0)$ is singular, and if $M_1 v_1 = 0$ then $P_2(\mu)$ has null space $v_1 \otimes w$ for any $w \in \mathbb{R}^{2n}$. Hence taking $x := v_1 \otimes v_2$ we have $P_1(\mu_0)x = P_2(\mu_0)x = 0$, or equivalently $L_1(\mu_0)x = L_2(\mu_0)x = 0$.

To establish the last statement of the lemma we see that if $L_1(\lambda)$ is singular then for every fixed $\lambda = \lambda_0$ the matrix $L_1(\lambda_0)$ has a null vector, and by the Bézoutian connection this means the matrix pencils $M_1(\lambda_0,\mu)$ and $M_2(\lambda_0,\mu)$ share an eigenvalue μ. The same argument proves the counterpart for any fixed $\mu = \mu_0$. ∎

5.2. Conditions for L_1, L_2 to be singular. We have seen in Section 4.3 that a necessary condition for L_1, L_2 to be singular is that both A_1 and A_2 have an eigenvector that is orthogonal to d. We next present a sufficient condition for L_1, L_2 to be singular. Recall that a linear subspace W is called an invariant subspace of a square matrix A if $AW \subseteq W$. By the same notation, we may express an orthonormal basis of the subspace.

PROPOSITION 5.2. *Suppose that A_1 and A_2 have a shared invariant subspace W that is orthogonal to d. Then $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$ has infinitely many*

solutions with $N(\lambda,\mu)$ singular, and $L_1(\lambda)$, $L_2(\mu)$ are singular matrix pencils.

Proof. Let $k = \dim W$. By the assumption, $W^\top d = 0$ and there exist $k \times k$ matrices X, Y such that $A_1 W = WX$, $A_2 W = WY$. Hence

$$N(\lambda,\mu)W = \lambda\mu W - \mu A_1 W - \lambda A_2 W = W(\lambda\mu I - \mu X - \lambda Y),$$

and for each fixed value of μ, there exists λ such that $\mu X x = \lambda(\mu I - Y)x$ with $x \neq 0$, i.e., λ is a generalized eigenvalue of the matrix pencil $\mu X - \lambda(\mu I - Y)$. Hence for the nonzero vector $v := Wx \in W$ we have $N(\lambda,\mu)v = 0$ and $v^\top d = 0$, so M_1 and M_2 have a null vector $\left[\begin{smallmatrix} 0 \\ v \end{smallmatrix}\right]$, hence $\det M_1 = \det M_2 = 0$. It follows that $\det L_1(\lambda) = 0$ for every value of λ, and hence, by Lemma 5.1, L_1 is singular. The discussion for L_2 is analogous. ∎

Note that most of the solutions (λ,μ) for $\det M_1 = \det M_2 = 0$ with N singular as in the above proposition are irrelevant for our problem, as they do not satisfy the KKT conditions; indeed $|\lambda|, |\mu|$ are necessarily bounded at the KKT points, see Section 6.2. Below we discuss remedies for this issue of having infinitely many solutions.

Although we have derived both a necessary condition and a sufficient condition for L_1, L_2 to be singular (and we derive another sufficient condition in Section 5.5.2), identifying the precise necessary and sufficient condition appears to be highly non-trivial. We leave this as an open problem.

5.3. Regularity test for L_1, L_2. We describe a process to test, and guarantee when successful, that the matrix pencils L_1, L_2 are regular. One direct method is to use the GUPTRI algorithm [10] for L_1 or L_2, but this can cost $O(n^8)$ operations in the worst case as the matrices have size $O(n^2)$. We look for cheaper alternatives.

By the necessary condition, the first test to guarantee regularity of L_1, L_2 is to see if both A_1, A_2 have an eigenvector orthogonal to d. This can be done with $O(n^3)$ cost. Generic cases pass this test, in which we proceed with solving (4.7), (4.8).

If this first test fails, we employ another test: fix λ to an arbitrary number and check whether $\det M_1(\lambda,\mu) = \det M_2(\lambda,\mu) = 0$ has a solution μ by solving two linear eigenvalue problems of size $2n$, which can be done with $O(n^3)$ cost. If there is no solution μ, then we can safely conclude that L_1, L_2 are regular. On the other hand, if the test fails for more than $4n^2$ distinct values of λ, we may assert that L_1, L_2 are singular. In this case, we take the additional preprocessing steps described below before invoking Algorithm 3.1. Thus the test for regularity requires $O(n^5)$ cost in the worst case.

5.4. Preprocessing by dimension reduction. Our next goal is to rectify Algorithm 3.1 so that it can deal with pathological cases where L_1, L_2 are singular.

Proposition 5.2 implies that a sufficient condition for L_1, L_2 to be singular is for A_1 and A_2 to have a common eigenvector orthogonal to d. Such cases may arise naturally in practice, an extreme example being $A_1 = A_2 = I$. In general, if the intersection of eigenspaces of A_1 and A_2 is of dimension two or larger, then there must be a common eigenvector orthogonal to d.

Fortunately, there is a remedy in such cases. Let W_1 and W_2 be the eigenspaces of A_1 and A_2 corresponding to the eigenvalues ρ_1 and ρ_2, respectively, and let $W := W_1 \cap W_2$ be nonzero. We now define subspaces $\hat{W}$ and $\bar{W}$ as follows. If $W^\top d \neq 0$, let $\bar{W}$ be the set of vectors in W orthogonal to d. If $W^\top d = 0$, we take $\bar{W}$ to be an arbitrary subspace of W with one dimension less than W. Then we define $\hat{W}$ as the orthogonal complement of $\bar{W}$ in $\mathbb{R}^n$. Using the same notation, we may designate

orthonormal bases of these subspaces. We can simultaneously block-diagonalize A_1 and A_2 by the orthogonal matrix $[\hat{W}\ \bar{W}]$, so that

$$[\hat{W}\ \bar{W}]^\top A_1[\hat{W}\ \bar{W}] = \begin{bmatrix} \hat{A}_1 & O \\ O & \bar{A}_1 \end{bmatrix}, \quad [\hat{W}\ \bar{W}]^\top A_2[\hat{W}\ \bar{W}] = \begin{bmatrix} \hat{A}_2 & O \\ O & \bar{A}_2 \end{bmatrix}. \tag{5.1}$$

We then solve the problem of reduced size with $d := \hat{W}^\top d$, $A_1 := \hat{A}_1$, $A_2 := \hat{A}_2$. We can repeat the reduction process recursively until no intersection of eigenspaces is of dimension two or larger.

Algorithm 5.1 Preprocessing by dimension reduction.

1: Compute the eigenvalue decompositions of A_1, A_2, and look for the intersection W of eigenspaces with $\dim W \geq 2$.
2: If such W exists, then take $\hat{W}$ as in (5.1) and call Algorithm 5.1 with $b_1 := \hat{W}^\top b_1$, $A_1 := \hat{W}^\top A_1 \hat{W}$, $b_2 := \hat{W}^\top b_2$, $A_2 := \hat{W}^\top A_2 \hat{W}$.
3: If there exists no such W, then return (E_1, E_2) with $E_1 := E(b_1, A_1)$ and $E_2 := E(b_2, A_2)$.

Below we verify that this process yields a pair of ellipsoids having the same signed distance as the original one. Recalling the definition (1.1) and writing $w = \left[\begin{smallmatrix} \hat{w} \\ \bar{w} \end{smallmatrix}\right]$ with $\hat{w} \in \hat{W}$ and $\bar{w} \in \bar{W}$, we have

$$\mathrm{dist}(E_1, E_2) = \max_{\|\hat{w}\|^2 + \|\bar{w}\|^2 = 1} f(\hat{w}, \bar{w})$$

where using $\bar{W}^\top d = 0$ we can write

$$f(\hat{w}, \bar{w}) = \langle \hat{w}, \hat{W}^\top d \rangle - \sqrt{\langle \hat{w}, \hat{A}_1 \hat{w} \rangle + \langle \bar{w}, \bar{A}_1 \bar{w} \rangle} - \sqrt{\langle \hat{w}, \hat{A}_2 \hat{w} \rangle + \langle \bar{w}, \bar{A}_2 \bar{w} \rangle}.$$

The following result shows that the dimension reduction employed in Algorithm 5.1 results in an equivalent signed distance problem of reduced size.

PROPOSITION 5.3. *Under the above notation, for every pair $(\hat{w}, \bar{w})$ with $\|\hat{w}\|^2 + \|\bar{w}\|^2 = 1$ and $\bar{w} \neq 0$, there exists a vector $\hat{w}'$ with $\|\hat{w}'\| = 1$ such that $f(\hat{w}', 0) \geq f(\hat{w}, \bar{w})$.*

Proof. Note that there exists a common eigenvector $\hat{v}$ such that $\hat{A}_1 \hat{v} = \rho_1 \hat{v}$ and $\hat{A}_2 \hat{v} = \rho_2 \hat{v}$. We choose $\hat{v}$ so that $\langle \hat{v}, W^\top d \rangle \geq 0$ and $\|\hat{v}\| = 1$. Put $\hat{c} = \hat{v}^\top \hat{w}$ and $\hat{c}' = \sqrt{\hat{c}^2 + \|\bar{w}\|^2}$. Then $\hat{w}' = \hat{w} + (\hat{c}' - \hat{c})\hat{v}$ satisfies $\|\hat{w}'\|^2 = \|\hat{w}\|^2 + \|\bar{w}\|^2 = 1$. In addition, we have

$$\langle \hat{w}', \hat{A}_j \hat{w}' \rangle = \langle \hat{w}, \hat{A}_j \hat{w} \rangle + \langle \bar{w}, \bar{A}_j \bar{w} \rangle$$

for $j = 1, 2$, and hence

$$f(\hat{w}', 0) - f(\hat{w}, \bar{w}) = \langle \hat{w}', \bar{W}^\top d \rangle - \langle \hat{w}, \bar{W}^\top d \rangle = (\hat{c}' - \hat{c}) \langle \hat{v}, \bar{W}^\top d \rangle \geq 0,$$

as required. □

The preprocessing step in Algorithm 5.1 results in an equivalent signed distance problem with lower-dimensional A_j, b_j. The reduced problem belongs to the generic case (no eigenvector is orthogonal to d) for some instances including those with $A_1 = \alpha_1 I$, $A_2 = \alpha_2 I$, and $d \neq 0$ in the original setting.

5.5. Perturbing the ellipsoids to avoid singular L_1, L_2. There are situations in which even with the preprocessing step just described, Algorithm 3.1 does not result in L_1, L_2 regular. An example is when $W^\top d = 0$ in the above setting.

Another extreme example is when $b_1 = b_2$, i.e., when the center of the ellipsoids are the same. In this case $d = 0$, and it is easy to see that (3.11) holds for every pair (λ, μ) such that $N(\lambda, \mu)$ is singular. Even worse, one can show that even the Kronecker canonical form of the linear problems L_1, L_2 do not provide the KKT points λ, μ. Handling this case analytically appears to be surprisingly difficult.

To illustrate the distinct cases, Figure 5.1 shows the representative configuration of the ellipsoids in the plane $n = 2$: (i) the generic case, (ii) the degenerate case treated in Section 5.4, and (iii) the degenerate case treated in this subsection.

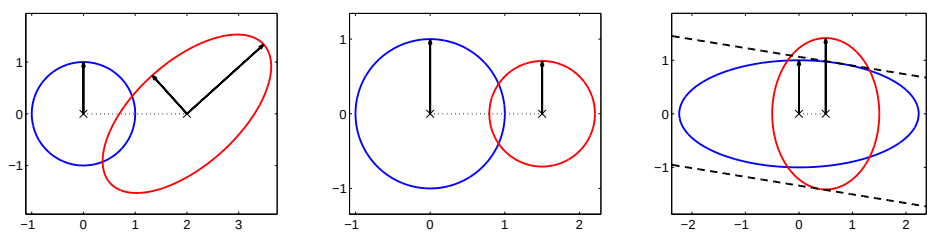

FIG. 5.1. *Two-dimensional illustration of generic and degenerate cases. Left: generic case where at least one of A_1 or A_2 has no eigenvector (shown by arrows) orthogonal to d, shown by the dotted line. Center: degenerate case of shared invariant subspace caused by multiple eigenvalues of A_1 and A_2. This case can be resolved by the process of dimension reduction as described in Section 5.4. Right: degenerate case of shared eigenspace caused by eigenvectors for distinct eigenvalues being orthogonal to d. In this case dimension reduction cannot be employed; indeed the tangent planes corresponding to the signed distance (dashed black lines) are not perpendicular to the horizontal axis. In this case our solution is to slightly perturb the center b_1.*

5.5.1. Perturbing a center. A simple remedy for the difficulty arising from $d = 0$, and also the case where $W := W_1 \cap W_2$ as above satisfies $W^\top d = 0$, is to introduce a small perturbation in the center b_j, thus perturbing d.

Specifically, the following procedure constructs a perturbed problem such that eigenvectors of A_1 corresponding to simple eigenvalues all become non-orthogonal to d. The perturbed problem belongs to the generic case for example when the original A_j had only simple eigenvalues.

Algorithm 5.2 Preprocessing by perturbing b_j.

1: Let $A_j = Q\Lambda_j Q^\top$ be the eigenvalue decomposition with an orthogonal matrix Q.
2: For each column vector q_i of Q corresponding to a *simple* eigenvalue, if $q_i^\top d = 0$, then set $b_j := b_j + \epsilon_i q_i$ with an arbitrary nonzero number ϵ_i of magnitude $O(\epsilon)$, where ϵ is the machine precision.

5.5.2. Perturbing a matrix. It is sometimes necessary to also perturb the matrices A_1, A_2. This happens when for every fixed λ, (even after the dimension reduction) there exists μ such that $N(\lambda, \mu) = \lambda\mu I - \mu A_1 - \lambda A_2$ has a null space of

dimension ≥ 2. An $n = 4$ example is

$$A_1 = 2I + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_2 = 2I + \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

In such cases, L_1, L_2 are singular regardless of d.

PROPOSITION 5.4. *Suppose that for every fixed $\lambda \in \mathbb{R}$, there exists $\mu \in \mathbb{R}$ such that $N(\lambda, \mu) = \lambda\mu I - \mu A_1 - \lambda A_2$ has a null space of dimension ≥ 2. Then L_1, L_2 in* (4.7), (4.8) *are singular matrix pencils.*

Proof. Let $\lambda \in \mathbb{R}$ any fixed nonzero finite number. By the assumption, there exists a finite $\mu \in \mathbb{R}$ such that $N(\lambda, \mu)$ has a null space of dimension two or larger, which means the null space contains a nonzero vector x with $x^\top d = 0$. Then we easily see that the vector $\left[\begin{smallmatrix} 0 \\ x \end{smallmatrix}\right] \in \mathbb{R}^{2n}$ is a null vector of both $M_1(\lambda, \mu)$ and $M_2(\lambda, \mu)$, so (λ, μ) is a solution of $\det M_1(\lambda, \mu) = \det M_2(\lambda, \mu) = 0$. Hence by Lemma 5.1 we conclude that L_1, L_2 are singular. $\square$

This is another sufficient condition for L_1, L_2 to be singular, which is neither a subset nor a superset of the condition in Proposition 5.2.

Since the statement of the proposition makes no reference to d, this result indicates that perturbing the ellipsoid centers is sometimes not enough. If we do not pass the regularity test in Section 5.3 after having perturbed the center b_j as in Algorithm 5.2, then as a last resort we perturb the matrix A_j as well.

Algorithm 5.3 Preprocessing by perturbing A_j and b_j.

1: Let $A_j = Q\Lambda_j Q^\top$ be the eigenvalue decomposition with an orthogonal matrix Q.
2: For each column vector q_i of Q corresponding to a *multiple* eigenvalue, if $q_i^\top d = 0$, then set $b_j := b_j + \epsilon_i q_i$ with an arbitrary nonzero number ϵ_i of magnitude $O(\epsilon)$.
3: Update $A_j := Q(\Lambda_j + \Delta)Q^\top$, where Δ is a diagonal matrix with nonzero distinct diagonals of magnitude $O(\epsilon)$.

The matrix A_j thus obtained has only simple eigenvalues, and no eigenvector is orthogonal to d, resulting in a generic situation of Definition 4.2.

5.5.3. Sensitivity analysis under perturbation. Perturbing the inputs clearly perturbs the output $\mathrm{dist}(E_1, E_2)$, which is undesirable if the change in the output is significantly magnified. We now examine the sensitivity of the signed distance $\mathrm{dist}(E_1, E_2)$ under perturbation in b_j and A_j.

Recalling (1.1), we first examine the term $\langle w, b_j \rangle$, the only term that depends on b_j. Defining $\hat{b}_j = b_j + \delta b_j$, for any vector w with $\langle w, w \rangle = 1$, we have

$$|\langle w, \hat{b}_j \rangle - \langle w, b_j \rangle| \leq \max_{\langle w, w \rangle = 1} |\langle w, \delta b_j \rangle| = \|\delta b_j\|.$$

Therefore, perturbing b_j by ϵ can only perturb the signed distance by at most ϵ, that is, the signed distance is a well-conditioned function of b_j, so slightly perturbing the ellipsoid centers is always harmless.

Perturbation in A_j is a bit more subtle. Defining $\hat{A}_j = A_j + \delta A_j$ and assuming the perturbation is small enough so that $\|\delta A_j\| \leq \sigma_{\min}(A_j)$, in which we take the

spectral norm $\|A\| := \sigma_{\max}(A)$, we have

$$\begin{aligned}\left|\sqrt{\langle w, \hat{A}_j w\rangle} - \sqrt{\langle w, A_j w\rangle}\right| &\le \left|\sqrt{\langle w, A_j w\rangle - \|\delta A_j\|} - \sqrt{\langle w, A_j w\rangle}\right| \\ &= \frac{\|\delta A_j\|}{\sqrt{\langle w, A_j w\rangle - \|\delta A_j\|} + \sqrt{\langle w, A_j w\rangle}} \\ &\lesssim \frac{\|\delta A_j\|}{2\sqrt{\langle w, A_j w\rangle}} \le \frac{\|\delta A_j\|}{2\sqrt{\sigma_{\min}(A_j)}}.\end{aligned}$$

Strict inequality holds here up to $O(\|\delta A_j\|^2)$. Using these bounds we obtain

$$|\mathrm{dist}(E_1, E_2) - \mathrm{dist}(\hat{E}_1, \hat{E}_2)| \lesssim \|\delta b_1\| + \|\delta b_2\| + \frac{\|\delta A_1\|}{2\sqrt{\sigma_{\min}(A_1)}} + \frac{\|\delta A_2\|}{2\sqrt{\sigma_{\min}(A_2)}}. \tag{5.2}$$

Note that equality can be attained in (5.2) up to $O(\|\delta A_j\|^2)$: denoting by v the eigenvector for $\sigma_{\min}(A_j)$, suppose $w = v$ and take $\delta A_j = \epsilon v v^\top$. The above argument shows that the conditioning with respect to perturbation in A_j is proportional to the smallest eigenvalue of A_j, so care is needed when A_j is ill-conditioned. For a matrix A, we denote by $\kappa(A)$ its condition number, i.e., $\kappa(A) := \sigma_{\max}(A)/\sigma_{\min}(A)$. Between A_1 and A_2, it is advised to perturb A_j with smaller condition number. If $\kappa(A_j) \gg 1$, then the change in the solution may be unacceptably large, and one may be advised to use higher precision arithmetic. If $\kappa(A_j) = O(1)$, which is easy to verify, then perturbing A_j can be done safely.

We note that while the conditioning of the signed distance can be analyzed as above, that of the corresponding points (x_1, x_2) on the ellipsoids cannot be bounded in general. For an extreme example, let E_1, E_2 both be unit spheres centered at the origin. Then the signed distance is clearly -2, attained at $(u, -u)$ for an arbitrary point u with $\|u\| = 1$. However, by a small ϵ perturbation in b_1, the corresponding points become uniquely determined with signed distance $\epsilon - 2$, illustrating the sensitivity of the points (x_1, x_2) to perturbation in b_j.

This does not necessarily indicate the weakness of the algorithm, but rather the difficulty of the problem of finding x_j. Indeed, small perturbation in A_j, b_j can be regarded as a small backward error [15, §1.8], so the backward stability is not lost.

In practice, perturbing b_1 to introduce a small nonzero d still results in $\det M_1(\lambda, \mu)$, $\det M_2(\lambda, \mu) \approx 0$ whenever $\det N = 0$, indicating numerical instability may be an issue. One remedy for this is to apply a diagonal congruence scaling $\mathrm{diag}(\sqrt{c}I, \frac{1}{\sqrt{c}}I)$ for some $c > 0$ to balance the norms of the diagonal blocks: we redefine

$$M_1(\lambda, \mu) := \begin{bmatrix} c_1 \mu A_1 & N(\lambda, \mu) \\ N(\lambda, \mu) & \frac{\mu}{c_1} D \end{bmatrix}, \quad M_2(\lambda, \mu) = \begin{bmatrix} c_2 \lambda A_2 & N(\lambda, \mu) \\ N(\lambda, \mu) & \frac{\lambda}{c_2} D \end{bmatrix}.$$

An appropriate choice is $c_1 \approx \sqrt{\frac{\|D\|}{\|A_1\|}}$, $c_2 \approx \sqrt{\frac{\|D\|}{\|A_2\|}}$. Clearly their determinants are the same as those in (3.10), but the numerical behavior is improved significantly. This scaling is recommended whenever the norms of $\|D\|$ and $\|A_j\|$ differ, so we employ it by default in our algorithm.

5.6. The linearization is regular. In summary, our overall algorithm proceeds as follows:

The rationale behind the choice of j is that if A_j has only simple eigenvalues then Algorithm 5.2 alone is sufficient, and otherwise we choose j so that perturbing

Algorithm 5.4 Overall algorithm to compute the signed distance between ellipsoids.

1: If the regularity test is not passed, then reduce dimension by Algorithm 5.1.
2: If the regularity test is not passed, then select $j \in \{1, 2\}$ as follows: set j if A_j has only simple eigenvalues. Otherwise, take j with smaller condition number $\kappa(A_j)$, and perturb b_j by Algorithm 5.2.
3: If the regularity test is not passed, then perturb A_j, b_j by Algorithm 5.3.
4: Invoke Algorithm 3.1 to compute the signed distance.

A_j by Algorithm 5.3 is more stable. After the preprocessing steps in Algorithms 5.1, 5.2 and 5.3, L_1, L_2 are regular matrix pencils, and $\det M_1 = \det M_2 = 0$ has a finite number of solutions. As a result, one can reliably compute the eigenvalues of L_1, L_2, which give a finite number $4n^2$ of candidate points for the KKT conditions, and hence Algorithm 3.1 computes the signed distance between overlapping ellipsoids with $O(n^6)$ cost.

6. KKT points and local optimality. We now turn to analyzing the KKT points; in particular we discuss how to check the local optimality of the KKT points. Although this is not directly necessary for Algorithm 3.1, which simply compares the objective values at the KKT points, investigating the local optima reveals the nonconvex nature of the problem and deepens our understanding.

6.1. Checking local optimality of the primal KKT points. For a KKT point w of the primal problem (1.1) obtained through (3.16), we would like to identify whether it is a local optimum or a saddle point. For simplicity, we introduce the equivalent problem

$$\underset{\|w\|=1}{\text{minimize}} \left(-\langle w, b_1 \rangle + \sqrt{\langle w, A_1 w \rangle} + \langle w, b_2 \rangle + \sqrt{\langle w, A_2 w \rangle} \right) \tag{6.1}$$

and its Lagrangian function

$$\mathcal{L}(w, \varphi) = -\langle w, b_1 \rangle + \sqrt{\langle w, A_1 w \rangle} + \langle w, b_2 \rangle + \sqrt{\langle w, A_2 w \rangle} + \frac{\varphi}{2}(1 - \|w\|^2). \tag{6.2}$$

Classical results in constrained optimization [27, §12.5] implies that if a KKT point $\hat{w}$ with a Lagrange multiplier $\hat{\varphi}$ is a local minimum then the projected Hessian of $\mathcal{L}$ onto the tangent plane $\mathcal{T}(\hat{w}) = \{v : \hat{w}^\top v = 0\}$ at $\hat{w}$ is positive semidefinite. Conversely, $\hat{w}$ is a strict local minimum if the projected Hessian of $\mathcal{L}$ onto $\mathcal{T}(\hat{w})$ is positive definite. Direct computation shows that the Hessian is

$$\frac{\partial^2 \mathcal{L}}{\partial w^2} = \frac{1}{\sqrt{w^\top A_1 w}} \left(A_1 - \frac{A_1 w w^\top A_1}{w^\top A_1 w} \right) + \frac{1}{\sqrt{w^\top A_2 w}} \left(A_2 - \frac{A_2 w w^\top A_2}{w^\top A_2 w} \right) - \varphi I. \tag{6.3}$$

Recall that our algorithm obtains the KKT point w for (1.1) through (3.16), and the corresponding KKT point (x_1, x_2) and Lagrange multipliers λ, μ for (3.1) are already available. Using these we can rewrite the right-hand side of (6.3) as

$$\frac{\partial^2 \mathcal{L}}{\partial w^2} = \frac{\varphi}{\lambda} \left(A_1 - (x_1 - b_1)(x_1 - b_1)^\top \right) + \frac{\varphi}{\mu} \left(A_2 - (x_2 - b_2)(x_2 - b_2)^\top \right) - \varphi I.$$

We can then test the local optimality by checking if the matrix $Q_w^\top \frac{\partial^2 \mathcal{L}}{\partial w^2} Q_w$ is positive definite, where Q_w is the $n \times (n-1)$ matrix such that $[\hat{w}, Q_w]$ is orthogonal. This is

a basic problem in linear algebra requiring $O(n^3)$ operations, which can be done by either attempting the Cholesky factorization or computing the eigenvalues.

There is some subtlety in testing local optimality by this process: when the projected Hessian $Q_w^\top \frac{\partial^2 \mathcal{L}}{\partial w^2} Q_w$ is positive semidefinite but not definite. In this case we can neither guarantee that the KKT point is locally optimal, nor exclude it as a saddle point. Fortunately, this rarely happens. Under the nondegeneracy assumption that the projected Hessian is nonsingular at every KKT point, we can enumerate all the locally optimal solutions.

6.2. Bounds for Lagrange multipliers. We show that at the KKT points the Lagrange multipliers λ, μ are bounded in absolute value.

PROPOSITION 6.1. *At every KKT point* (x_1, x_2) *of* (2.1) *with Lagrange multipliers* λ *and* μ *satisfying* (3.2)–(3.5)*, the absolute values of* λ *and* μ *are bounded by*

$$|\lambda| \le \frac{\|A_1\|(\|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|})}{\sqrt{\sigma_{\min}(A_1)}}, \tag{6.4}$$

$$|\mu| \le \frac{\|A_2\|(\|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|})}{\sqrt{\sigma_{\min}(A_2)}}. \tag{6.5}$$

Proof. By (3.2) and (3.3) we have

$$\sqrt{\sigma_{\min}(A_i)} \le \|x_i - b_i\| \le \sqrt{\sigma_{\max}(A_i)}, \quad i \in \{1, 2\}. \tag{6.6}$$

These also imply bounds on $\|x_1\|, \|x_2\|$: we have $\|x_i\| \le \|b_i\| + \sqrt{\|A_i\|}$. Now by (3.4) and (3.5) we have

$$|\lambda| \|A_1^{-1}(x_1 - b_1)\| = \|x_1 - x_2\| \le \|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|}. \tag{6.7}$$

Since $\|A_1^{-1}(x_1 - b_1)\| \ge \sigma_{\min}(A_1^{-1})\|x_1 - b_1\|$, from (6.7) we obtain

$$\begin{aligned} |\lambda| &\le \frac{\|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|}}{\|A_1^{-1}(x_1 - b_1)\|} \le \frac{\|A_1\|(\|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|})}{\|x_1 - b_1\|} \\ &\le \frac{\|A_1\|(\|b_1\| + \|b_2\| + \sqrt{\|A_1\|} + \sqrt{\|A_2\|})}{\sqrt{\sigma_{\min}(A_1)}}, \end{aligned}$$

where we used the lower bound in (6.6) for the last inequality. This proves (6.4). The bound (6.5) can be obtained entirely analogously. ∎

The bounds (6.4), (6.5) can be used to exclude some of the computed solutions for $\det L_1(\lambda) = 0$, $\det L_2(\mu) = 0$ that violate them.

7. Numerical experiments. This section presents numerical experiments to illustrate Algorithm 3.1. All the experiments were conducted in MATLAB 2013A. Unless otherwise mentioned, we generate the ellipsoids by setting $A_i = X_i^\top X_i$ with X_i being $n \times n$ random matrices and by taking random vectors as b_i, using the standard normal distribution `randn`.

7.1. Illustration in two dimensions. For ease of visualization we first consider the two-dimensional case $n = 2$.

Recall that in order to compute the signed distance our algorithm finds the KKT points for the relaxed dual problem (3.1). As shown in Section 3.2, the computed

KKT points contain those of the primal problem (1.1) (those with $\lambda\mu > 0$) and more (those with $\lambda\mu < 0$).

We illustrate the situation by visualizing the set $x_1 - x_2$ with the conditions $x_1 \in \mathrm{bd}(E_1)$, $x_2 \in \mathrm{bd}(E_2)$, then plotting the KKT points of (3.1) in the left figures below. Plotting the set $\{x_1 - x_2 \mid x_1 \in \mathrm{bd}(E_1), x_2 \in \mathrm{bd}(E_2)\}$ directly is nontrivial, so we fix $x_1 \in \mathrm{bd}(E_1)$ on a boundary of E_1 and plot $x_1 - x_2$ where x_2 moves along $x_2 \in \mathrm{bd}(E_2)$, then repeat this for many points on $x_1 \in \mathrm{bd}(E_1)$ (black curves). The convex set $U = E_1 \ominus E_2$ in Section 2 is the interior of the outer boundary of the black region $\{x_1 - x_2 \mid x_1 \in \mathrm{bd}(E_1), x_2 \in \mathrm{bd}(E_2)\}$. The KKT points computed by Algorithm 3.1 are indicated by the dots (green, blue and red); those on the boundary $\mathrm{bd}(E_1 \ominus E_2)$ correspond to the KKT points for the primal problem (1.1). In particular, the blue dots correspond to the local solutions as verified by the procedure in Section 6.1, and the red dot to the global solution.

In the center figures, we draw the ellipsoids together with the KKT points, with each pair of KKT points connected by black lines: the solid thick line corresponds to the solution, the dashed line connects x_1, x_2 at the KKT points with $\lambda\mu > 0$ (critical points for the primal problem (1.1)), and the dot-dashed lines show the other KKT points with $\lambda\mu < 0$ (critical points for the relaxation problem (3.1), with $x_1 - x_2$ in the interior of $U = E_1 \ominus E_2$).

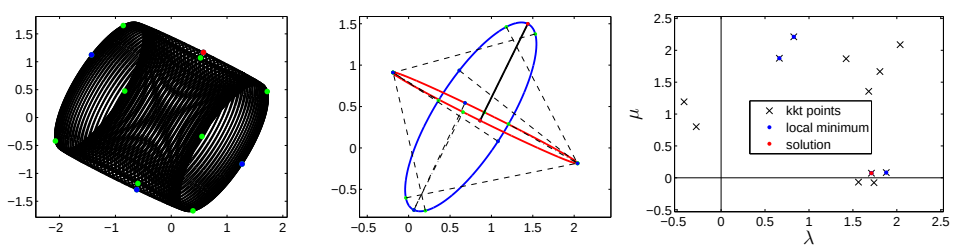

FIG. 7.1. *Overlapping case with four local minima. For the figures in this subsection, the three plots illustrate the following: Left: The Minkowski difference* $\{x_1 - x_2 | x_1 \in \mathrm{bd}(E_1), x_2 \in \mathrm{bd}(E_2)\}$ *and the KKT points. Center: ellipsoids and the KKT points. Right:* (λ, μ)*-values satisfying the KKT conditions.*

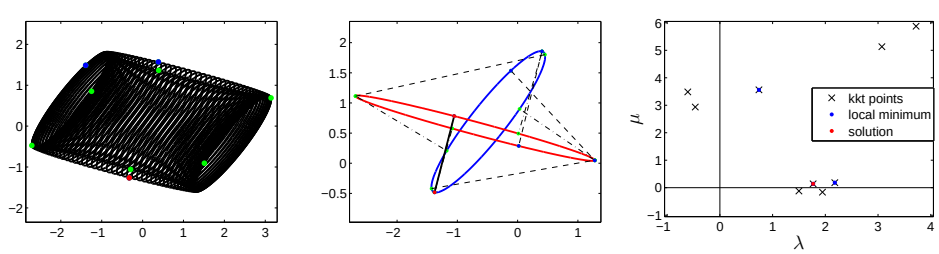

FIG. 7.2. *Overlapping case with three local minima.*

Observe from Figures 7.1–7.4 that the number of local minima for (1.1) in the overlapping case differs depending on the relative positioning of E_1 and E_2. Our experiments suggest that the number of local minima is $1, 2, 3$ or 4. Moreover, we observe that each local minimum for (1.1) is also a local minimum for the dual problem (2.1). The total number of KKT points for the relaxation problem (3.1) also depends largely on the relative positioning of the ellipsoids.

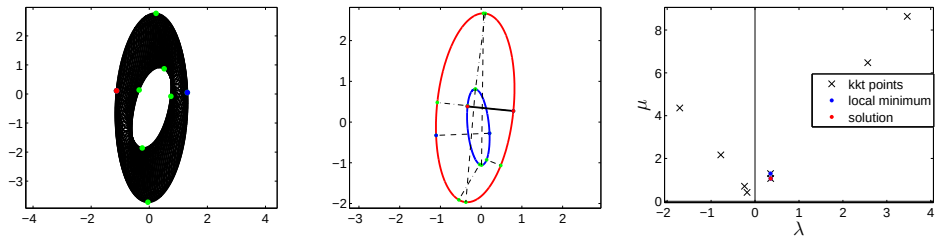

FIG. 7.3. *Overlapping case with two local minima.*

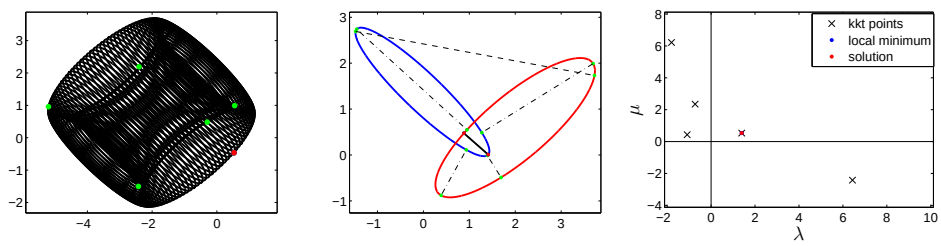

FIG. 7.4. *Overlapping case with one local minimum equal to the global solution.*

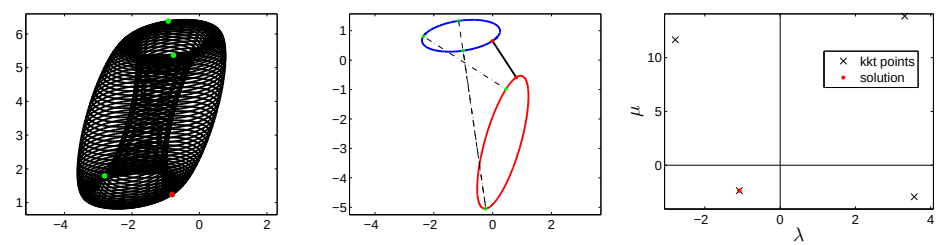

FIG. 7.5. *Non-overlapping case. There is just one local minimum equal to the global solution.*

Figure 7.5 is a typical illustration of a non-overlapping (convex) case, for which more efficient algorithms are available but Algorithm 3.1 is still applicable.

7.2. KKT points for larger dimension. We next run Algorithm 3.1 for larger dimension n and illustrate by plotting the Lagrange multipliers (λ, μ) corresponding to the KKT points for the relaxation problem (3.1). As before, the KKT points for the primal problem are those for which λ and μ have the same sign; see Figure 7.6.

In the non-overlapping case, we generally observe that there is at least one point in each quadrant of the (λ, μ)-plane and the unique point with $\lambda, \mu < 0$ corresponds to the global solution, verifying our results in Section 3.2.

In the overlapping case there are generally many more KKT points, reflecting the non-convexity of the problem. We also observe that the solution in the overlapping case generally corresponds to one of the KKT points closest to the origin with $\lambda, \mu > 0$. This is analogous to the point-ellipsoid distance problem, in which the solution corresponds to the KKT point with the smallest Lagrange multiplier [12, 17]. However, the right plot of Figure 7.6 suggests that the solution may not have the smallest value of λ or μ. Experiments suggest it may hold that the pair of Lagrange multipliers (λ_*, μ_*) for the global solution admits no other positive Lagrange multipliers (λ, μ)

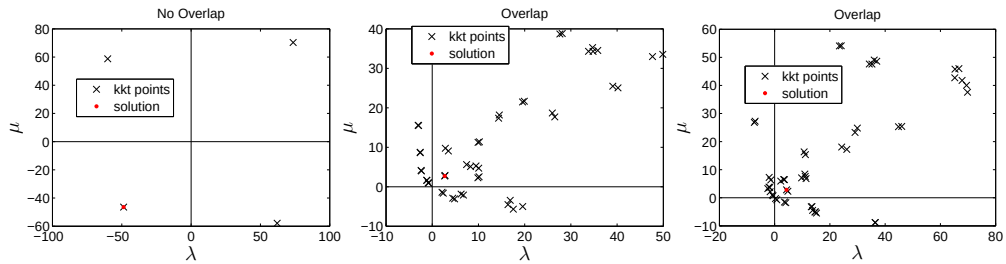

FIG. 7.6. *Values of* (λ, μ) *for the KKT points, non-overlapping case (left) and overlapping case (center and right).*

for KKT points with $\lambda > \lambda_*$ and $\mu > \mu_*$; a precise characterization appears to be nontrivial and we leave it as an open problem.

We also used the process in Section 6.1 to examine the number of local optima among the KKT points. Fortunately, the projected Hessian was always either strictly indefinite or positive definite, so we accurately counted the number of local optima. Experiments suggest that in the overlapping case there are at most *four* local optima, regardless of the dimension n. Note that for the special case of point-ellipsoid distance problem it is known that there can be at most *two* local minima [12, 22], and our experiments indicate that an analogous result may hold in our ellipsoid-ellipsoid problem. Making this observation precise is also an open problem.

7.3. Runtime with increasing dimension. The dominant cost of our algorithm is in computing the eigenvalues of a $4n^2 \times 4n^2$ linear generalized eigenvalue problem, for which the complexity is $O(n^6)$. To verify this numerically we generated random examples of varying dimension $n = 2, 3, \ldots, 40$ and examined the runtime.

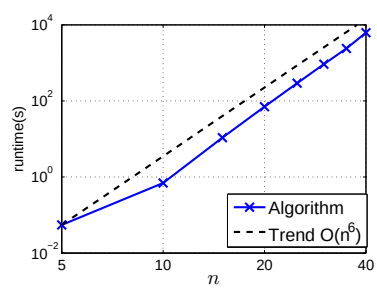

FIG. 7.7. *Runtime as the dimension* n *varies.*

Figure 7.7 confirms that the runtime scales as $O(n^6)$. In all cases with $n \geq 10$, at least 90% of the runtime was spent on the eigensolver.

8. Conclusion and discussion. We have proposed the first polynomial-time algorithm for computing the signed distance between two ellipsoids. The algorithm computes the KKT points for a relaxation of the dual problem, for which we analyze connections with the primal KKT points. A crucial component is to convert the KKT conditions into a two-parameter eigenvalue problem, which is then reduced to linear generalized eigenvalue problems of size $O(n^2)$. We analyze when the conversion results in a regular pencil, and propose remedies for the pathological cases. We present numerical experiments, which, in addition to illustrating the ellipsoids and

the algorithm performance, pose several open problems.

Let us conclude with remarks on possible future work. First, the $O(n^6)$ complexity can be a bottleneck when n is large, so a natural question is to design a more efficient algorithm. In addition, our algorithm perturbs the ellipsoids in the pathological cases, and it would be nice to have an approach that does not require such treatments.

Finally, an active area of research in non-convex optimization is the Celis-Dennis-Tapia (CDT) problem, which arises as a subproblem of a nonlinear optimization problem [9, 11]. In the CDT problem one is asked to minimize a non-convex quadratic function over the intersection of two ellipsoids. A number of recent studies (e.g. [1, 6]) examine and analyze the CDT problem, but to our knowledge there is no established deterministic polynomial-time algorithm for solving it, except for the one in the very recent preprint [4] that addresses this question and describes a polynomial-time algorithm for the CDT problem under some solvability assumptions on a system of quadratic equations.

This paper has focused on the signed distance problem, but we suspect some of the ideas that we have introduced may be useful for the CDT problem. Specifically, both problems have two ellipsoidal constraints, which lead to KKT conditions involving two Lagrange multipliers. In this paper we have computed the KKT points for the signed distance problem by solving an eigenvalue problem, and we conceive that this framework may be applicable also to the CDT and other optimization problems.

Acknowledgements. We are very grateful to Vanni Noferini for showing us the connection between the linearizations (4.7), (4.8) and Bézoutians, and we thank Martin Lotz for a fruitful discussion.

REFERENCES

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

[2] F. V. Atkinson, *Multiparameter Eigenvalue Problems*, Academic Press, NY, USA, 1972.

[3] C. Bhattacharyya, *Second order cone programming formulations for feature selection*, J. Mach. Learn. Res., 5 (2004), pp. 1417–1433.

[4] D. Bienstock, *A note on polynomial solvability of the CDT problem*, ArXiv 1406.6429, (2014).

[5] D. A. Bini and A. Marco, *Computing curve intersection by means of simultaneous iterations*, Numerical Algorithms, 43 (2006), pp. 151–175.

[6] I. M. Bomze and M. L. Overton, *Narrowing the difficulty gap for the celis-dennis-tapia problem*, Preprint, Department of Statistics and Operations Research, Universitaet Wien, Vienna, Austia, (2013).

[7] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[8] W. Briec, *Minimum distance to the complement of a convex set: duality result*, J. Optim. Theory Appl., 93 (1997), pp. 301–319.

[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] J. Demmel and B. Kågström, *The generalized Schur decomposition of an arbitrary pencil A-zB: robust software with error bounds and applications. Part I: theory and algorithms*, ACM Trans. Math. Soft., 19 (1993), pp. 160–174.

[11] M. El-Alem, *A global convergence theory for the celis-dennis-tapia trust-region algorithm for constrained optimization*, SIAM J. Numer. Anal., 28 (1991), pp. 266–290.

[12] G. E. Forsythe and G. H. Golub, *On the stationary values of a second-degree polynomial on the unit sphere*, J. SIAM, 13 (1965), pp. 1050–1068.

[13] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, USA, 2009. Unabridged republication of book first published by Academic Press in 1982.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 4th ed., 2012.

[15] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, USA, 2nd ed., 2002.

[16] M. E. Hochstenbach, A. Muhič, and B. Plestenjak, *On linearizations of the quadratic two-parameter eigenvalue problem*, Linear Algebra Appl., 436 (2012), pp. 2725–2743.

[17] S. Iwata, Y. Nakatsukasa, and A. Takeda, *Global optimization methods for extended Fisher discriminant analysis*, in Proc. Seventh AISTATS, JMLR W&CP 33, 2014, pp. 411–419.

[18] E. Jarlebring and M. E. Hochstenbach, *Polynomial two-parameter eigenvalue problems and matrix pencil methods for stability of delay-differential equations*, Linear Algebra Appl., 431 (2009), pp. 369–380.

[19] A. A. Kurzhanskiy and P. Varaiya, *Ellipsoidal toolbox*, Tech. Report EECS-2006-46, EECS, UC Berkeley, 2006.

[20] L. Lerer and M. Tismenetsky, *The Bézoutian and the eigenvalue-separation problem for matrix polynomials*, Integral Equations and Operator Theory, 5 (1982), pp. 386–445.

[21] A. Lin and S.-P. Han, *On the distance between two ellipsoids*, SIAM J. Optim., 13 (2002), pp. 298–308.

[22] J. M. Martínez, *Local minimizers of quadratic functions on Euclidean balls and spheres*, SIAM J. Optim., 4 (1994), pp. 159–176.

[23] A. Muhič and B. Plestenjak, *On the quadratic two-parameter eigenvalue problem and its linearization*, Linear Algebra Appl., 432 (2010), pp. 2529–2542.

[24] Y. Nakatsukasa, V. Noferini, and A. Townsend, *Computing the common zeros of two bivariate functions via Bézout resultants*, Numer. Math., (2014), pp. 1–29.

[25] J. S. Nath and C. Bhattacharyya, *Maximum margin classifiers with specified false positive and false negative error rates.*, in Proc. Seventh SDM, SIAM, 2007, pp. 35–46.

[26] L. Nirenberg, *Functional Analysis*, Academic Press, 1961.

[27] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer New York, 2nd ed., 1999.

[28] V. Noferini. private communication, 2013.

[29] A. Takeda, H. Mitsugi, and T. Kanamori, *A unified classification model based on robust optimization*, Neural Computation, 25 (2013), pp. 759–804.

[30] F. Tisseur and K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.