

# Approximation of $n$ -Dimensional Data Using Spherical and Ellipsoidal Primitives

Giuseppe Calafiore

**Abstract**—This paper discusses the problem of approximating data points in  $n$ -dimensional Euclidean space using spherical and ellipsoidal surfaces. A closed form solution is provided for spherical approximation, while an efficient, globally optimal solution for the ellipsoidal problem is proposed in terms of semidefinite programming (SDP). In addition, the paper presents a result for robust fitting in presence of outliers, and illustrates the theory with several numerical examples. A brief survey is also presented on the solutions to other relevant geometric approximation problems, such as ellipsoidal covering of convex hulls and pattern separation.

**Index Terms**—Geometric approximation, semidefinite programming, spherical and ellipsoidal fitting, outliers.

## I. INTRODUCTION

THE problem of approximating observed data with simple geometric primitives is, since the times of Gauss, of fundamental importance in many scientific endeavors. To this purpose, the most commonly used primitives are lines, planes (and their multidimensional extensions i.e., affine sets), cubes and polyhedra, and quadratic sets or surfaces, such as spheres and ellipsoids. The areas of application of geometric approximation techniques range from computer graphics and vision [2], [19], pattern recognition [21], [27] and representation of objects [17], to statistics [11], [23] and estimation [5]. The fitting process allows for reduction and simplification of the data, to the benefit of further data manipulation stages.

Despite the importance of the problem however, the results available in the literature on this topic suffer from various limitations and drawbacks. First, most of the available results propose solutions to the problem in the plane, i.e., only with data of dimension two, see for instance [8], [9] and references therein (see also [14] for a discussion of the three-dimensional [3-D] case). Second, most of the existing methods are based on the minimization of “algebraic distance” (or equation error) criteria, [2], [8], [22]. These criteria may not have a clear geometric interpretation, since they are not directly related to the geometric Euclidean distances from the points to the curve, as it is discussed in depth in [9]. Third, most of the available results (with the exception of the work in [8], for which however the previous two points hold) are based on iterative nonlinear optimization schemes, such as modifications of the Gauss-Newton algorithm,

see [9], [22], which may be prone to convergence issues and local solutions. Theoretical issues related to nonlinear LS problems encountered in computer vision are also discussed in [13].

In this paper, we present the solution of spherical and ellipsoidal fitting based on a “difference-of-squares” (DOS) geometric error criterion for points in  $n$ -dimensional space. The main novelties of the proposed methods are that a) the algorithms are either closed form, or guaranteed to converge to the *global* optimum in a finite number of iterations and b) they work without restriction in spaces of generic dimension  $n$ . In particular, we propose a closed-form solution for the spherical fitting problem, and a solution for ellipsoidal fitting based on a particular class of *convex* optimization algorithms, i.e., semidefinite programs (SDP, see [1], [16], [30], [33] and references therein). In practice, this implies that the *global* solution of the ellipsoidal fitting problem in  $\mathbb{R}^n$  can be determined efficiently using recently developed interior-point algorithms for SDPs, [6], [30], [32].

The paper is structured as follows: in Section II the closed-form solution to the spherical fitting problem is presented, the results for ellipsoidal approximation and robust ellipsoidal approximation are discussed in Sections III, and IV illustrates the theory with several numerical examples. Finally, in Section V, we briefly review some other geometric problems encountered in the literature related to ellipsoidal approximations. Conclusions are then drawn in Section VI.

## II. SPHERICAL FITTING

Let the given data points  $x_i \in \mathbb{R}^n, i = 1, \dots, N$  be grouped in the data matrix

$$X = [x_1 \ x_2 \ \dots \ x_N] \in \mathbb{R}^{n,N}.$$

The goal is to determine a sphere that approximates the data points, according to an appropriate optimality criterion. Let  $c \in \mathbb{R}^n$  represent the center of the sphere, and  $r \in \mathbb{R}^+$  the respective radius. The criterion we shall consider is based on the squared difference between the distance from the point to the center ( $d_i = \|x_i - c\|$ ) and the distance from the projection of the point on the sphere and the center ( $\|\tilde{x}_i - c\| = r$ ), see Fig. 1. This fitting criterion will be referred to as the DOS criterion in the sequel, and it is expressed as

$$f_{\text{DOS}}(c, r) = \sum_{i=1}^N (\|x_i - c\|^2 - r^2)^2. \quad (1)$$

Manuscript received January 30, 2001; revised March 14, 2002. This paper was recommended by Associate Editor A. Ollero.

The author is with the Dipartimento di Automatica e Informatica Politecnico di Torino, 24-10129 Torino, Italy (e-mail: calafiore@polito.it).

Publisher Item Identifier S 1083-4427(02)06006-X.

*Remark:* Other criteria have been proposed in the literature for surface fitting. From a geometric point of view, the most significant criterion is the Euclidean geometric criterion (EG)

$$f_{\text{EG}}(c, r) = \sum_{i=1}^N (\|x_i - c\| - r)^2 \quad (2)$$

which is based on the sum of squared geometric distances from the points to the surface, measured along the line through the point and the center of the sphere; this distance is also denoted “radial Euclidean distance” in the computer vision literature.

Other generic criteria are those based on the algebraic error (AE) (also called equation error), and are of the form  $\sum_i F^2(x_i)$ , where  $F(x) = 0$  is an implicit parametric representation of the surface. Notice that different equations may represent the same surface. For instance, we may represent *same* sphere by the equation  $F_1(x) = \|x - c\| - r = 0$ , or by the equation  $F_2(x) = \|x - c\|^2 - r^2 = 0$ , or also by  $F_3(x) = ax^T x + b^T x + c = 0$ , therefore an algebraic criterion is related to one particular representation of the geometric surface. The representation  $F_3$  gives rise to the criterion

$$f_{\text{AE}}(a, b, c) = \sum_{i=1}^N (ax_i^T x_i + b^T x_i + c)^2. \quad (3)$$

We notice however, that these criteria are not equivalent in practice: optimizing with respect to  $F_1$  or  $F_2$  yields a surface that minimizes the EG criterion or the DOS criterion, respectively, while it is geometrically unclear what we are minimizing when using  $F_3$ . Moreover, in this latter case, one needs to add a further constraint on the parameters in order to avoid trivial solutions. A common choice is for instance  $\|[a \ b^T \ c]\| = 1$ , see [9].

The numerical cost for solving the minimization is also *not* the same for the three problems above: the most used criterion is AE, since its solution reduces to a simple linear Least-Squares problem. As we remarked, the EG criterion would be geometrically preferable, but its solution requires an iterative nonlinear optimization algorithm. The DOS criterion retains the geometric interpretation, and provides approximation performances which are very close to those obtained with the EG criterion, as shown by the numerical tests in Section IV.A. Moreover, its solution may be determined in closed form, as described in the following.  $\triangle$

Define  $\bar{x}$  as the barycenter of the data points

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4)$$

and  $\bar{X}$  as the matrix of centered data

$$\bar{X} = X - \bar{x}u^T \quad (5)$$

where  $u = [1 \ 1 \ \dots \ 1]^T \in \mathbb{R}^N$ . The following theorem holds.

*Theorem 1:* Let  $\bar{X} \in \mathbb{R}^{n, N}$  be full-row-rank, and let  $H \in \mathbb{R}^{N, N-n}$  be the matrix containing by columns a basis for the null space of  $\bar{X}$ . Define further

$$\begin{aligned} \Omega &= [2X^T \ H] \in \mathbb{R}^{N, N}, \\ \nu &= [x_1^T \ x_2^T \ \dots \ x_N^T]^T. \end{aligned} \quad (6)$$

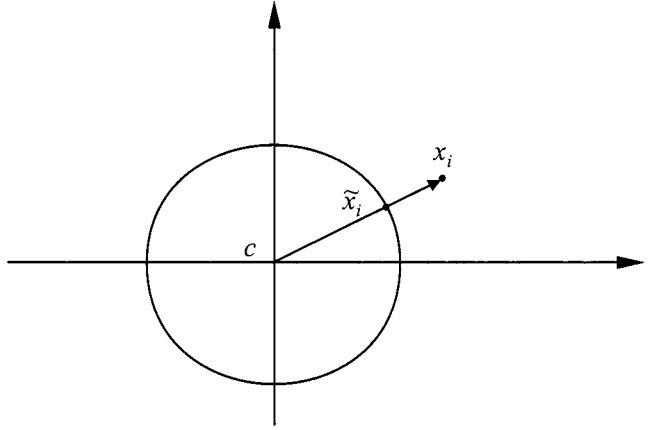


Fig. 1. Distance from a data point  $x_i$  to the sphere.

Then, the sphere that minimizes the criterion function  $f(c, r)$  in (1) is unique, and has center in

$$c = [I_n \ 0]\Omega^{-1}\nu \quad (7)$$

where  $I_n$  denotes the  $n \times n$  identity matrix. The optimal radius is given by

$$r = \left[ \frac{1}{N} \sum_{i=1}^N \|x_i - c\|^2 \right]^{1/2}. \quad (8)$$

*Proof:* Let  $\rho \doteq r^2$ , and  $g_i(c) \doteq \|x_i - c\|^2$ . With this definition, we have

$$f(c, \rho) = \sum_{i=1}^N (g_i(c) - \rho)^2.$$

The first-order optimality conditions for  $f(c, \rho)$  are

$$\frac{\partial f(c, \rho)}{\partial \rho} = -2 \sum_{i=1}^N (g_i(c) - \rho) = 0 \quad (9)$$

$$\frac{\partial f(c, \rho)}{\partial c} = \sum_{i=1}^N 2(g_i(c) - \rho) \frac{\partial g_i(c)}{\partial c} = 0 \quad (10)$$

where  $(\partial g_i(c)/\partial c) = -2(x_i - c)$ . From the first condition, we immediately obtain

$$\rho(c) = \frac{1}{N} \sum_{i=1}^N g_i(c). \quad (11)$$

Substituting now (11) in (10), after some manipulation we obtain the condition for the optimal center

$$\sum_{i=1}^N g_i(c)(x_i - \bar{x}) = 0 \quad (12)$$

where  $\bar{x}$  is the barycenter of the data, defined in (4). Defining  $\bar{X}$  as in (5), we notice that the condition (12) is satisfied if and only if the vector  $g(c) \doteq [g_1(c) \ g_2(c) \ \dots \ g_N(c)]^T$  lies in the null space of  $\bar{X}$ . Since  $\text{Rank}(\bar{X}) = n$ , a basis for  $\text{Null}(\bar{X})$  is composed of  $N - n$  independent vectors, that we arrange in the columns of the matrix  $H \in \mathbb{R}^{N, N-n}$ . Notice that the vector  $g(c)$

may be expressed in the form  $g(c) = \nu - 2X^T c + u c^T c$ , where  $\nu$  is defined in (6), therefore the condition  $g(c) \in \text{Null}(\bar{X})$  is expressed as

$$2X^T c + H\xi - u c^T c = \nu$$

for some  $\xi \in \mathbb{R}^{N-n}$ . It may be easily verified that  $u \in \text{Null}(\bar{X})$ , therefore there exist a vector  $\xi_u$  such that  $u = H\xi_u$ , and we may write  $H\xi - u c^T c = H\eta$ , for  $\eta \doteq \xi - \xi_u c^T c$ . The previous equation becomes now linear in  $c$  and  $\eta$

$$\begin{bmatrix} 2X^T & H \end{bmatrix} \begin{bmatrix} c \\ \eta \end{bmatrix} = \nu.$$

Notice that the matrix  $\Omega = \begin{bmatrix} 2X^T & H \end{bmatrix}$  is square and full-rank ( $\text{Rank}[\bar{X}^T \ H] = N$  from the definition of orthogonal complement; substituting  $\bar{X}^T = X^T - u\bar{x}^T = X^T - H\xi_u\bar{x}^T$ , it follows that  $\text{Rank}[X^T \ H] = N$ ), therefore the optimal solution for the center is unique, and the statement of the theorem immediately follows.  $\square$

As a simple extension of the above result, we may consider *weights* on the data points. This may be useful for instance when it is known that different data points are subject to different levels of noise, or in outliers rejection schemes. If  $w_i \geq 0, i = 1, \dots, N$  are the weights, the (weighted) DOS criterion writes

$$f_{W \text{ DOS}}(c, r) = \sum_{i=1}^N w_i (\|x_i - c\|^2 - r^2)^2. \quad (13)$$

The results of Theorem 1 can then be extended as follows. Let  $W = \text{diag}(w_1, w_2, \dots, w_N)$ , define the weighted barycenter and the weighted centered data matrix as

$$\bar{x}_w \doteq \frac{1}{\sum_{i=1}^N w_i} \sum_{i=1}^N w_i x_i; \quad \bar{X}_w \doteq (X - \bar{x}_w u^T) W$$

and let  $H_w$  be a matrix containing by columns a basis for the null space of  $\bar{X}_w$ . Then, the center of the sphere which is optimal for the weighted DOS criterion is

$$c = [I_n \ 0] [2X^T \ H_w]^{-1} \nu$$

and the (square of the) corresponding optimal radius is

$$r^2 = \frac{1}{\sum_{i=1}^N w_i} \sum_{i=1}^N w_i \|x_i - c\|^2.$$

Numerical tests for spherical fitting are reported in the examples Section IV.A.

### III. ELLIPSOIDAL FITTING

In this section, we study the problem of fitting an ellipsoidal surface to a set of  $n$ -dimensional data points, according to a DOS error criterion. As previously discussed, current results available in the literature do not guarantee convergence to the global optimum for the generic  $n$ -dimensional ellipsoid-fitting

problem, and are generally based on equation error (AE) criteria, which do not always have a direct geometric interpretation. The existing AE-based fitting algorithms in two-dimensional (2-D) and 3-D usually require the addition of further constraints on the design variables, whose geometric significance is often unclear, see for instance [2], [14], [24], [27]. Moreover, the resulting conic is not in general guaranteed to be an ellipse, so that further iterative refinements or rejection schemes need to be introduced, see [7], [18], [22]. Recently, a direct “ellipse-specific” fitting method has been proposed in [8]. This method is based on the solution of a generalized eigenvalue problem, and can therefore be considered a “closed-form” method. The main drawback of this method is that it cannot be extended to deal with multidimensional data (i.e., it works only for points on the plane), since in the multidimensional case the “ellipticity” constraint cannot be expressed as a simple quadratic constraint on the equation parameters. A key observation for our development is that the ellipticity constraint is however a *convex* constraint on suitable equation parameters (see Theorems 2 and 3).

More sophisticated fitting algorithms use the solution of an AE-based fitting method to provide a starting point for iterative nonlinear optimization algorithms (such as curvature weighted algorithms, [9]) that aim to the minimization of the EG criterion. These algorithm suffer from usual problems related to nonlinear optimization schemes: convergence is not guaranteed, nor it is the determination of the *global* optimum, see e.g., [8], [9], [22], [24], [28].

Another issue that we tackle is *robustness* with respect to data points that violate the statistical assumptions of the estimator, often referred to as *outliers*. We will not enter into the details of robust estimation here, but only recall that criteria based on the sum of squared residuals are particularly sensitive to the presence of outliers in the data. It is generally recognized that it is better in this respect to consider criteria based on the sum of the absolute values of the residuals, [11], [15], [23], [28]. Other policies for outlier rejection are those based on consensus methods, like the RANSAC paradigm, see for instance [4] and references therein. In the sequel, we provide the results for both the sum-of-squares and the absolute-value-sum criteria.

The main point in this section is that ellipsoidal fitting in  $n$ -dimensions is a *convex* problem. This fundamental fact seems to have gone unrecognized in the literature, but has several important implications. Notably, the solution is guaranteed to converge to the *global* optimum of the problem, thus eliminating the issue of local stationarity points. Moreover, the problem may be solved with great numerical efficiency by means of recently developed interior-point methods, as discussed in the sequel. In particular, the ellipsoid-fit problem will be cast as an SDP (see for instance [30], [33]), which is a class of convex optimization problems, for which extremely efficient numerical algorithms are available.

Let the ellipsoid  $\mathcal{E}$  be described as

$$\mathcal{E} = \{x: (x - c)^T A (x - c) = r^2 \quad \text{Tr } A = 1\} \quad (14)$$

where  $c \in \mathbb{R}^n$  is the center, and  $A \in \mathbb{R}^{n,n}$  is a symmetric positive definite (this will be expressed as  $A \succ 0$ ) matrix that defines the *shape* of the ellipsoid. We consider the following

DOS-based optimality criteria. First, we consider the sum-of-squares criterion

$$f(c, A, r) = \sum_{i=1}^N (d^2(x_i, c) - d^2(\tilde{x}_i, c))^2 \quad (15)$$

where  $d(a, b)$  denotes a metric distance between the points  $a$  and  $b$ , and  $\tilde{x}_i$  denotes the point on the ellipsoid that lies on the ray  $\tilde{x}_i c$ , see Fig. 2. It may be easily verified that

$$\tilde{x}_i = c + r \frac{x_i - c}{\sqrt{(x_i - c)^T A (x_i - c)}}.$$

If we choose to measure distances according to the metric induced by  $A$ , then  $d^2(a, b) = (a - b)^T A (a - b)$ , and the criterion (15) assumes the simple form

$$f_2(c, A, r) = \sum_{i=1}^N ((x_i - c)^T A (x_i - c) - r^2)^2. \quad (16)$$

To take into account robustness issues regarding the sensitivity of the estimate with respect to outliers, we also consider the following criterion based on the absolute-value-sum

$$f_1(c, A, r) = \sum_{i=1}^N |(x_i - c)^T A (x_i - c) - r^2|. \quad (17)$$

*Remark:* As we remarked for spherical fitting (which is of course a particular case of ellipsoidal fitting), other optimality criteria are currently used in the literature. The most geometrically significant criterion is again the Euclidean error criterion

$$f_{EG}(c, A, r) = \sum_{i=1}^N (\sqrt{(x_i - c)^T A (x_i - c)} - r)^2. \quad (18)$$

Minimization of this criterion requires iterative nonlinear optimization algorithms and, reportedly, presents severe convergence problems (see [9]). The most frequently used criterion is instead one based on the algebraic (equation) error, using the equation  $x^T A x + b^T x + c = 0$  to represent the ellipsoid. Contrary to the spherical case, further difficulties arise here, due to the fact that the “ellipticity” constraint  $A \succ 0$  must be enforced by the algorithm. Several “tricks” and weighting schemes (see again [9], and [2], [24], [18], [7], [22]) have been proposed in the literature in order to force the the solution of the AE-based problem to approximate the solution of the EG problem. The limitations of these methods have been discussed in [8], where a direct “ellipse-specific” algorithm is presented. The main drawback of the algorithm proposed in [8] is that it is limited to the 2-D case. Below, we propose a globally convergent, ellipse specific algorithm that is based on the DOS criterion, and works without restrictions on the space dimension  $n$ .  $\triangle$

The computation of the optimal ellipsoid, for the criteria (16) and (17) is summarized in the following two theorems.

*Theorem 2:* Let

$$g_i(\Omega) \doteq [x_i^T \quad 1] \Omega \begin{bmatrix} x_i \\ 1 \end{bmatrix}, \quad \Omega \doteq \begin{bmatrix} A & b \\ b^T & d \end{bmatrix} \quad (19)$$

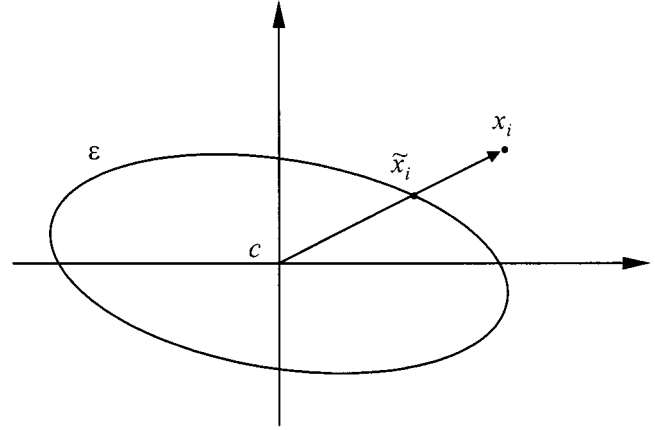


Fig. 2. Ellipsoid  $\mathcal{E}$  and points on the ray  $\tilde{x}_i c$ .

where  $A = A^T \in \mathbb{R}^{n,n}$ ,  $b \in \mathbb{R}^n$ , and  $d \in \mathbb{R}$ ; let also  $\gamma \in \mathbb{R}^N$  be a vector whose components  $\gamma_i, i = 1, 2, \dots, N$  are positive scalars. Consider the following SDP problem in the variables  $A, b, d, \gamma_i$

$$\text{minimize } u^T \gamma \text{ subject to} \quad (20)$$

$$\begin{bmatrix} 1 & g_i(\Omega) \\ g_i(\Omega) & \gamma_i \end{bmatrix} \succ 0, \quad i = 1, 2, \dots, N \quad (21)$$

$$A \succ 0; \quad \text{Tr } A = 1 \quad (22)$$

and let  $A_{\text{opt}}, b_{\text{opt}}, d_{\text{opt}}, \gamma_{\text{opt}}$  represent the values at the optimum of the variables of the previous SDP. Then, the ellipsoid (14), with

$$\begin{aligned} A &= A_{\text{opt}}, \\ c &= -A_{\text{opt}}^{-1} b_{\text{opt}}, \\ r^2 &= |c^T A c - d_{\text{opt}}| \end{aligned}$$

is the *global* minimizer for the criterion (16).

*Proof:* Expanding the product in (14), we obtain the ellipsoid equation in the form

$$x^T A x - 2c^T A x + (c^T A c - r^2) = 0; \quad \text{Tr } A = 1.$$

Considering the change of variables  $b = -Ac, d = c^T A c - r^2$ , we rewrite the ellipsoid equation in the form

$$x^T A x + 2b^T x + d = 0, \quad \text{Tr } A = 1.$$

Substituting  $x_i$  in the previous equation, we obtain  $g_i(\Omega)$ , as defined in (19). Let  $g(\Omega)$  be the vector of components  $g_i(\Omega)$ . The approximation problem with respect to the criterion (16) amounts therefore to finding  $A \succ 0, b, d$  that minimize  $\|g(\Omega)\|^2$ . If  $u$  is a vector of ones of appropriate dimension, this is in turn equivalent to minimizing  $u^T \gamma$  subject to  $g_i(\Omega)^2 < \gamma_i$ , for  $i = 1, 2, \dots, N$ . Using the Schur complements rule (see for instance [30]), the latter condition is equivalent to the matrix inequality

$$\begin{bmatrix} 1 & g_i(\Omega) \\ g_i(\Omega) & \gamma_i \end{bmatrix} \succ 0.$$

Since  $g_i$  is affine in the variables  $A, b, d$ , the previous condition is an LMI (linear matrix inequality) constraint on the variables.

The resulting optimization problem belongs therefore to a particular class of convex problems known in the literature as SDP problems, and takes the form

$$\text{minimize } u^T \gamma \text{ subject to} \quad (23)$$

$$\begin{bmatrix} 1 & g_i(\Omega) \\ g_i(\Omega) & \gamma_i \end{bmatrix} \succ 0, \quad i = 1, 2, \dots, N \quad (24)$$

$$A \succ 0; \quad \text{Tr } A = 1. \quad (25)$$

The statement of the theorem then follows easily from the previous considerations.  $\square$

The next theorem treats the case of robust ellipsoidal fitting, i.e., data approximation performed according to the criterion (17).

**Theorem 3:** Let  $g_i(\Omega)$  and  $\gamma$  be defined as in Theorem 2. Consider the following SDP problem in the variables  $A, b, d, \gamma_i$

$$\text{minimize } u^T \gamma \text{ subject to} \quad (26)$$

$$\gamma_i - g_i(\Omega) > 0, \quad i = 1, 2, \dots, N \quad (27)$$

$$\gamma_i + g_i(\Omega) > 0, \quad i = 1, 2, \dots, N \quad (28)$$

$$\gamma_i > 0, \quad i = 1, 2, \dots, N \quad (29)$$

$$A \succ 0; \quad \text{Tr } A = 1 \quad (30)$$

and let  $A_{\text{opt}}, b_{\text{opt}}, d_{\text{opt}}, \gamma_{\text{opt}}$  represent the values at the optimum of the variables of the previous SDP. Then, the ellipsoid (14), with

$$\begin{aligned} A &= A_{\text{opt}} \\ c &= -A_{\text{opt}}^{-1} b_{\text{opt}} \\ r^2 &= |c^T A c - d_{\text{opt}}| \end{aligned}$$

is the *global* minimizer for the criterion (17).

*Proof:* The first part of the proof follows the same line as that of the previous theorem. The approximation problem with respect to the criterion (17) amounts to finding  $A \succ 0, b, d$  that minimize  $\sum_i |g_i(\Omega)|$ . This is in turn equivalent to minimizing  $u^T \gamma$  subject to  $|g_i(\Omega)| < \gamma_i, \gamma_i > 0$ , for  $i = 1, 2, \dots, N$ , which is equivalent to

$$\begin{aligned} \gamma_i - g_i(\Omega) &> 0 \\ \gamma_i + g_i(\Omega) &> 0, \end{aligned} \quad \text{for } i = 1, 2, \dots, N.$$

Since  $g_i$  is affine in the variables  $A, b, d$ , the previous conditions, together with the positive definite condition on  $A$ , and the positivity condition on  $\gamma_i$ , are a set of LMI constraints on the variables. The statement of the theorem then follows easily from the previous considerations.  $\square$

*Remark:* We stress the fact that, since for both the considered criteria ellipsoidal fitting may be cast as an SDP (which is a particular class of convex programs), an appropriate numerical algorithm will *always* converge to the *global* optimum in a pre-defined maximum number of iterations. In particular, the complexity result for the SDPs in Theorems 2 and 3 is as follows: using a *general-purpose* primal-dual interior point algorithm as the one described in [30], the number of operations to attain the global optimum with given accuracy grows with problem size as  $O(N^2 n^4)$ , see [16], [30]. Notice also that the previous is a

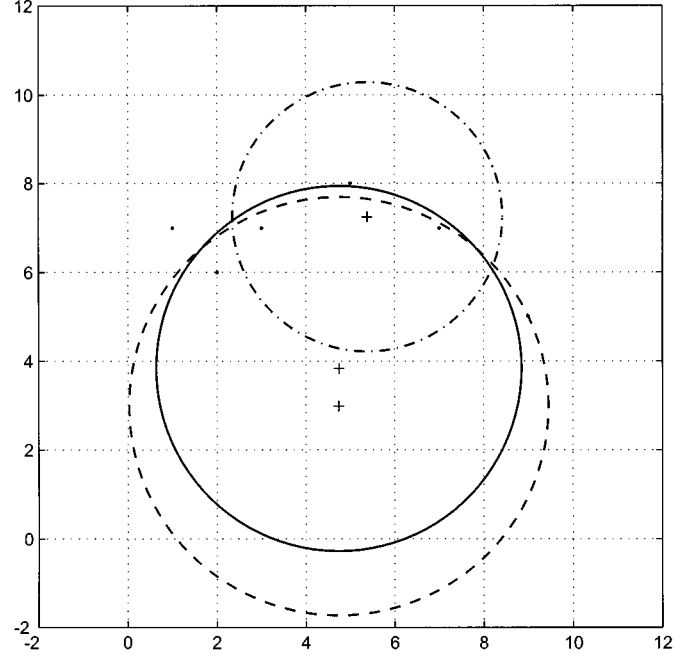


Fig. 3. Optimal closed-form solution for the DOS criterion: (1) (solid line), comparison with the algebraic criterion (3) (dash-dotted line), and the EG criterion (2) (dashed line).

*worst-case* complexity estimate, and that in practice these algorithms have been observed to behave much better than predicted by the theoretical estimate (see comments in [30]).

We also remark that, similar to the spherical case, weights on the data can be introduced in the optimality criterion, and the results of Theorem 2 and Theorem 3 can be modified accordingly. This straightforward extension is left to the reader.  $\triangle$

#### IV. NUMERICAL EXAMPLES

In this section, we report numerical experiments for the spherical and ellipsoidal fitting methods previously described.

##### A. Spherical Fitting

We start with a simple 2-D example proposed in [9], with data points in

$$X = \begin{bmatrix} 1 & 2 & 5 & 7 & 9 & 3 \\ 7 & 6 & 8 & 7 & 5 & 7 \end{bmatrix}.$$

The optimal circle computed according to Theorem 1 (DOS criterion) has center in  $c = [4.7423 \ 3.8351]^T$  and radius  $r = 4.1088$ , and it is depicted in Fig. 3, together with the solution obtained by minimization of the equation error criterion (3) (which has center in  $[5.3794 \ 7.2532]^T$  and radius 3.037, as reported in [9]), and the solution that minimizes the sum of actual squared EG criterion (2). The EG solution required an iterative Gauss-Newton method, and yielded the center in  $[4.7398 \ 2.9835]^T$  and radius 4.7142. To compare these three solutions, we computed *a posteriori* the final Euclidean residual  $\bar{d} = (\sum_{i=1}^N (\|x_i - c\| - r)^2)^{1/2}$ , obtaining  $\bar{d}_{ae} = 3.2944$  for the circle computed according to the AE criterion (3),  $\bar{d}_{\text{dos}} = 1.1825$  for the one computed according to DOS criterion (1), and  $\bar{d}_{eg} = 1.1080$  for the one resulting from the EG criterion

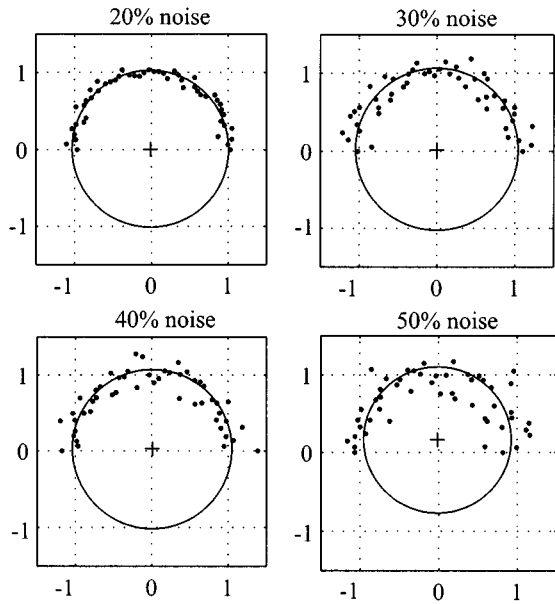


Fig. 4. Noise sensitivity of optimal fit circles for different noise ratios in the data.  $N = 50$  data points.

(2), this latter being of course the minimal achievable value of  $\bar{d}$ . From this example, we observe that the solution obtained from the DOS criterion is close to that resulting from minimization of the EG criterion, which comes at much higher computational cost, since it requires an iterative nonlinear optimization algorithm. This fact is supported by further experimental validation: we run a large number of tests using  $N = 50$  randomly generated points in  $\mathbb{R}^3$ , uniformly distributed in a 3-D box of unit side length, and computed the  $\bar{d}$  residuals for the solutions obtained from the DOS and EG criteria. The results are as follows: on average, the residual from DOS was only 2.1% larger than the minimum achievable residual (the one computed minimizing the EG criterion), while in the less favorable experiment the DOS residual was 7.0% larger than the minimum achievable residual.

We next tested the stability of the proposed method with respect to noise in the data. We considered occluded data (a semi-circle) affected by an increasing amount of Gaussian noise. The data have been generated as  $x = \rho[\cos \theta \ \sin \theta]^T$ , where  $\theta$  is uniformly distributed in the interval  $[0, \pi]$ , and  $\rho$  is Gaussian with mean one and variance  $\sigma^2$ . The noise intensity was set as  $3\sigma = 0.2; 0.3; 0.4; 0.5$  in the respective experiments. The method showed remarkable stability for noise levels up to 50% of the nominal radius (see Fig. 4).

We point out that the proposed fitting method, as it is common to all methods based on quadratic cost criteria, may be sensitive to outliers. One possibility for reducing outlier sensitivity is to solve spherical fitting as a special case of ellipsoidal fitting, using the solution based on the absolute-value-sum criterion proposed in Theorem 3. Otherwise, when the level of noise associated to each data point is known *a priori*, one can use a weighted criterion, with weights proportional to the inverse of the noise intensity. In the example shown in Fig. 5, the data have been generated as  $x = \rho[\cos \theta \ \sin \theta]^T$ , where  $\theta$  is uniformly distributed in  $[0, \pi]$ . To simulate the presence of outliers,

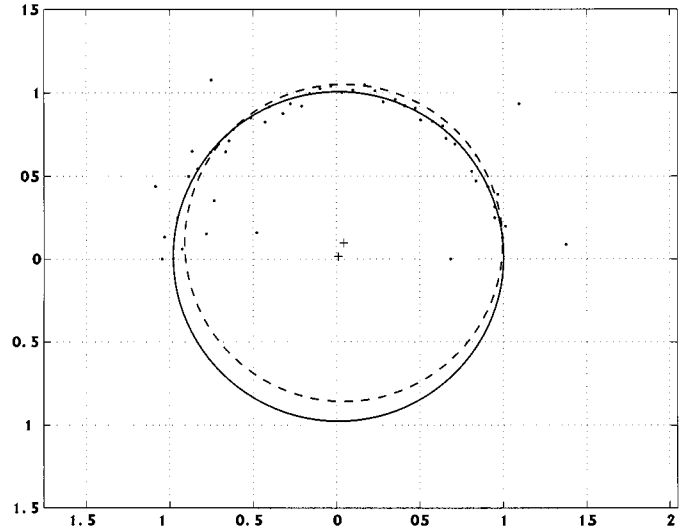


Fig. 5. Data having low noise intensity  $\sigma_{\text{low}} = 0.03$  with probability 0.9 and high noise intensity  $\sigma_{\text{hi}} = 0.3$  with probability 0.1. Solid line shows the weighted optimal circle. Dashed line shows the standard fit.

the radius  $\rho$  was taken as a Gaussian variable with mean one, having standard deviation  $\sigma_{\text{low}} = 0.03$  with probability 0.9 and  $\sigma_{\text{hi}} = 0.3$  with probability 0.1. We run a large number of tests using data instances with this kind of noise, and then computed the statistics on the sensitivity of the optimal solutions which are reported in Table I.

### B. Ellipsoidal Fitting

We first considered two data sets proposed in [9]. The results of ellipsoidal approximation computed according to Theorem 2 on these data sets is shown in Fig. 6. The optimal ellipsoid for the data in Fig. 6(a) has center  $c = [5.1978 \ 5.1093]^T$ ,  $A = \begin{bmatrix} 0.3784 & 0.1770 \\ 0.1770 & 0.6216 \end{bmatrix}$ , and  $r^2 = 4.3763$ . The value of the criterion (16) at the optimum is  $f_2 = 12.5831$ . The actual average Euclidean distance from the points to the optimal ellipse,  $\bar{d} = (\sum_i \|x_i - \tilde{x}_i\|^2)^{1/2}$  was computed a-posteriori, obtaining  $\bar{d} = 1.5824$ . Notice that this value of the residual Euclidean error falls in between of the values reported in [9], Section 3, for algebraic fit ( $\bar{d} = 1.80$  reported) and direct Euclidean fit (computed using an iterative nonlinear optimization scheme,  $\bar{d} = 1.17$  reported). For the data in Fig. 6(b), we determined an optimal ellipsoid with center  $c = [-0.3973 \ 0.5149]^T$ ,  $A = \begin{bmatrix} 0.1944 & 0.0276 \\ 0.0276 & 0.8056 \end{bmatrix}$ , and  $r^2 = 73.3422$ . The average Euclidean distance  $\bar{d}$  from the points to the optimal ellipse, computed a-posteriori, is  $\bar{d} = 3.5193$ .

Fig. 7 shows the sensitivity of the optimal DOS ellipsoid to different levels of Gaussian noise affecting the data.

Fig. 8 shows a comparison between the criteria (16) and (17) on a data set contaminated by two outliers. Fig. 8(a) shows that the solution computed by means of Theorem 2 [i.e., minimization of sum-of-squares criterion (16)] is particularly sensitive to the two outliers, while the solution computed according to Theorem 3 [i.e., minimization of absolute-value-sum criterion (17)] shows a remarkable stability, see Fig. 8(b).

TABLE I  
STATISTICS FOR OPTIMAL PARAMETERS OF APPROXIMATING CIRCLE. SEMI-CIRCLE DATA AFFECTED BY GAUSSIAN NOISE AND OUTLIERS.  
 $X_c, Y_c$  INDICATE THE COORDINATES OF THE CENTER

	avg. $X_c$	avg. $Y_c$	avg. radius	stdev. $X_c$	stdev. $Y_c$	stdev. radius
DOS fit	-0.0022	0.0499	0.9759	0.0233	0.0657	0.0373
Weighted DOS fit	-0.0004	0.0097	0.9949	0.0069	0.0154	0.0100

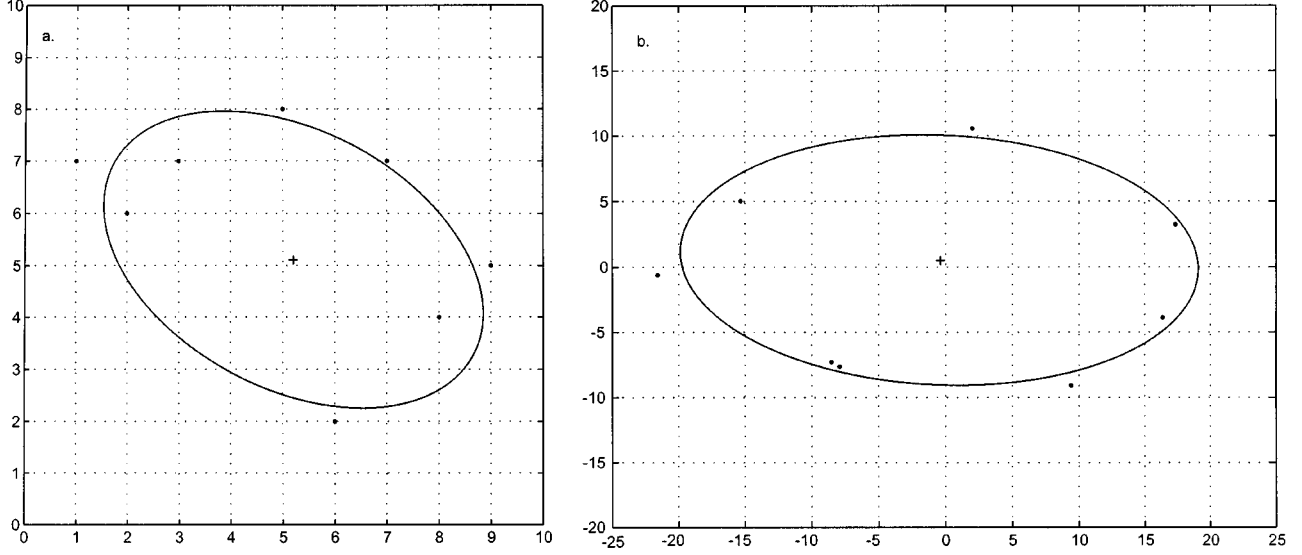


Fig. 6. Optimal ellipsoids computed according to criterion (16) on two data sets presented in [9].

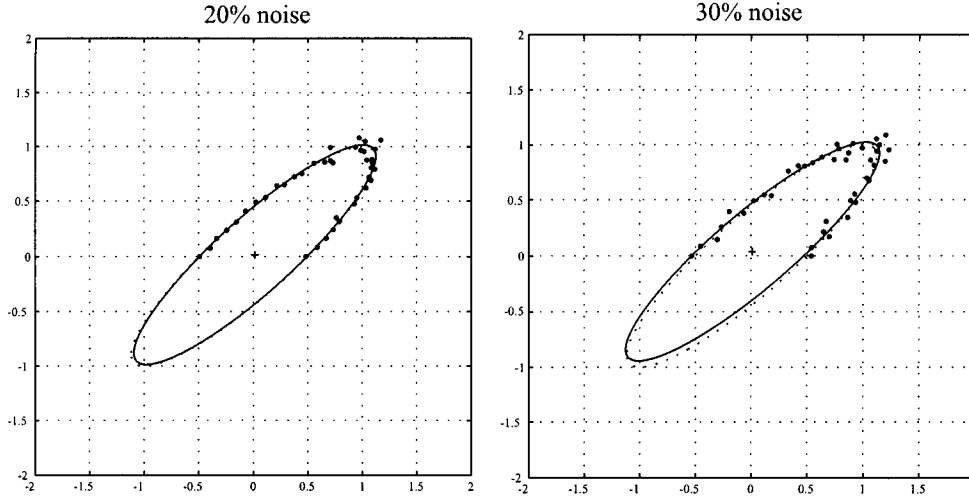


Fig. 7. Sensitivity of optimal ellipsoids computed according to criterion (16), for different levels of Gaussian noise. The dotted curves indicate the nominal ellipsoid used to generate the data.

To conclude this examples section, we considered three-dimensional data generated as

$$x = \rho \begin{bmatrix} .5 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} z$$

where  $z \in \mathbb{R}^3$  is uniformly distributed on the surface of a unit sphere. For  $\rho = 1$ , the points  $x$  lie on the surface of a nominal ellipsoid centered in the origin. We artificially created noisy

data affected by outliers generating values of  $\rho$  which with probability 0.9 are Gaussian with unit mean and standard deviation  $\sigma_{\text{low}} = 0.03$ , and with probability 0.1 are Gaussian with unit mean and standard deviation  $\sigma_{\text{hi}} = 0.3$ . We run several problem instances with  $N = 50$  noisy points, and for each instance we computed the optimal ellipsoids, using both criteria (16) and (17). The average performances of the fits are reported in Table II. All examples in this section were computed under Matlab, using the SP SDP solver [32], and the LMItol interface, [6].

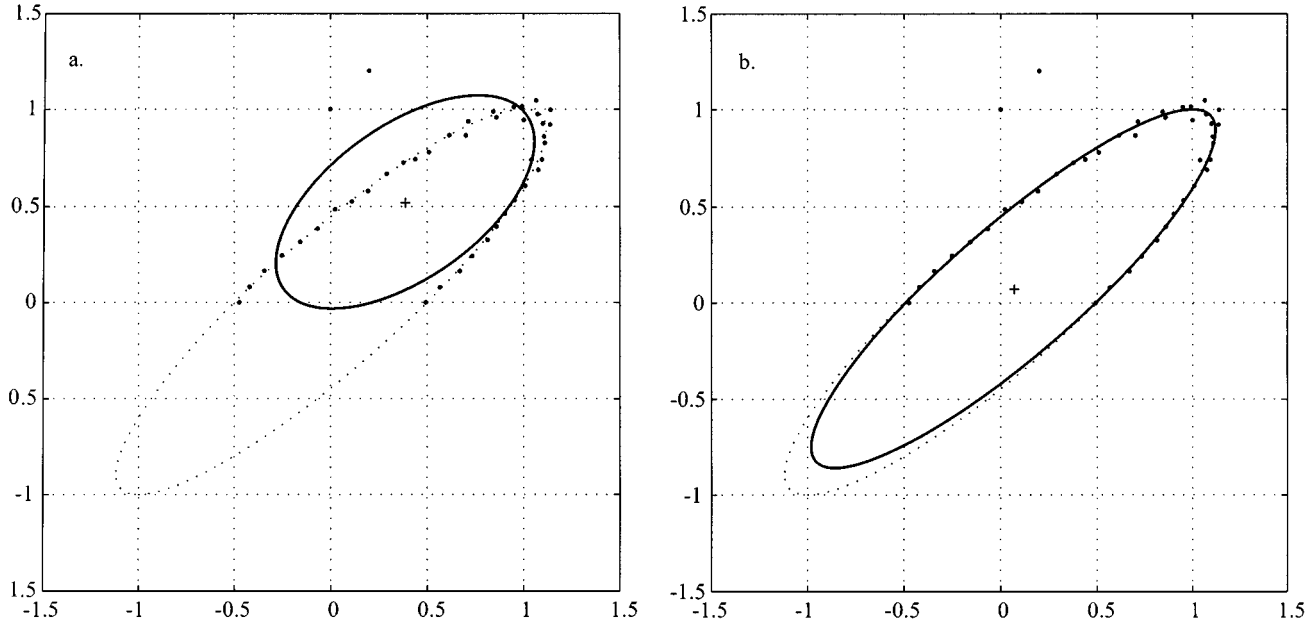


Fig. 8. Comparison of sum-of-squares criterion (left picture) and absolute-value-sum criterion (right picture), for a data set contaminated by two outliers.

TABLE II  
STATISTICS FOR OPTIMAL PARAMETERS OF APPROXIMATING ELLIPSOIDS, FOR DATA AFFECTED BY GAUSSIAN NOISE AND OUTLIERS. THE ERROR  $e$  REPRESENTS THE AVERAGE RADIAL DISTANCE FROM THE POINTS TO THE APPROXIMATING SURFACE

	avg. $A$			avg. $c$			avg. $r$	avg. $e$
sum-squares DOS fit	0.435	-0.441	-0.001	0.000	-0.001	-0.004	0.313	0.006
	-0.441	0.541	0.001					
	-0.001	0.001	0.023					
abs-sum DOS fit	0.432	-0.435	-0.000	0.000	-0.001	0.001	0.324	0.003
	-0.435	0.542	-0.000					
	-0.000	-0.000	0.025					

## V. OTHER ELLIPSOIDAL APPROXIMATION PROBLEMS

In the previous sections, we discussed the problem of approximating data in the least-squares sense via ellipsoidal surfaces. In some engineering applications it is instead of interest to determine other types of ellipsoidal approximations of the data. For instance, it is desired to compute an ellipsoid that contains the data, or to separate two different classes of points by means of an ellipsoidal surface. These geometric problems may again be cast as semidefinite optimization problems, and are briefly reviewed below.

### A. Minimal Ellipsoid Containing Given Points

A problem that arises in several applications of computer vision, cluster analysis [20], and robust statistics (e.g., in ellipsoidal peeling methods for outliers detection, [23]) is the computation of the minimal ellipsoid containing given  $n$ -dimensional points  $x_1, \dots, x_N$ , or equivalently, their convex hull, [30], [31]. We adopt here the following representation of the ellipsoid  $\mathcal{E}$

$$\{x: (x - c)^T P^{-1} (x - c) - 1 \leq 0\} \quad (31)$$

where  $P \in \mathbb{R}^{n,n} \succ 0$  denotes the “shape matrix” of the ellipsoid, and  $c \in \mathbb{R}^n$  is the center. The “size” of the ellipsoid is

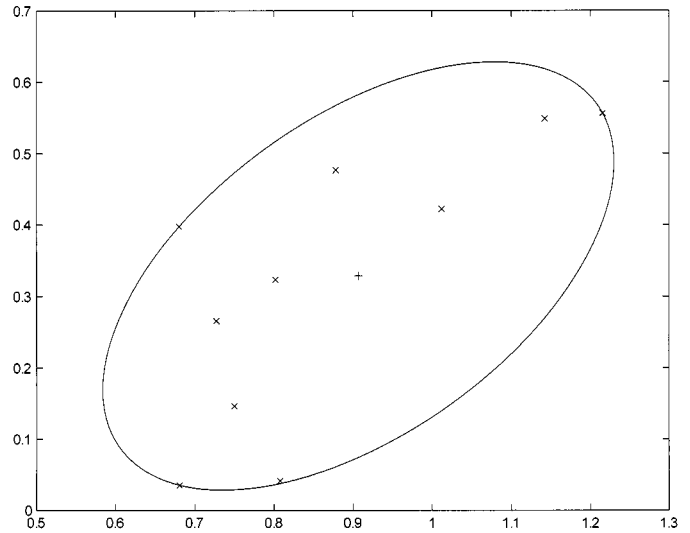


Fig. 9. Minimal ellipsoid covering a set of points.

measured by the sum of the squared semi-axes lengths, which is equal to  $\text{Tr } P$ . An alternative measure of size is related to the volume of the ellipsoid (which is proportional to  $\det P$ ), see for instance [31]. Clearly, a point  $x_i$  belongs to the ellipsoid (31) if and only if  $(x_i - c)^T P^{-1} (x_i - c) - 1 \leq 0$ . Using the Schur



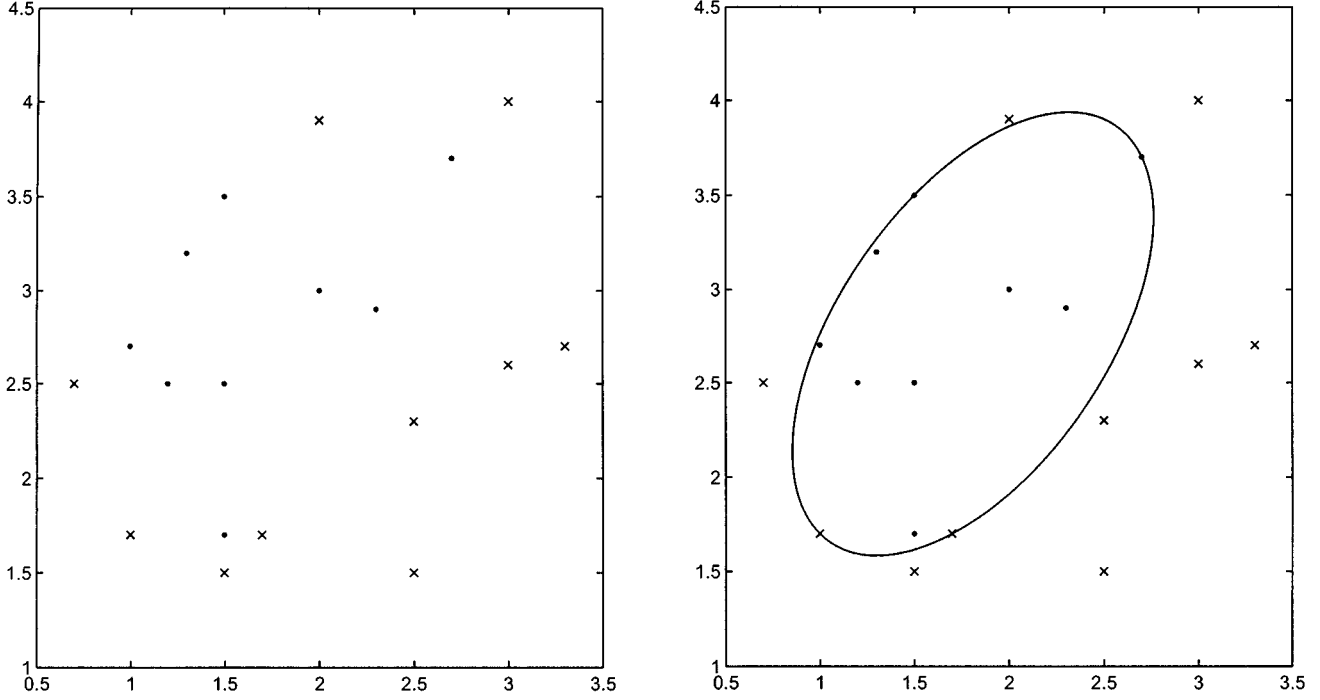


Fig. 10. Ellipsoid separating two data patterns. Points represent  $x$ -type data, and crosses represent  $y$ -type data.

complements rule, this condition (together with the condition  $P \succ 0$ ) is equivalent to

$$\begin{bmatrix} P & (x_i - c) \\ (x_i - c)^T & 1 \end{bmatrix} \succ 0$$

which is an LMI constraint on the problem variables  $P, c$ . Minimal ellipsoidal covering can therefore be solved by means of semidefinite programming, as detailed in the following: let  $P_{\text{opt}}, c_{\text{opt}}$  be the values of the variables of the following SDP at the optimum

$$\text{minimize } \text{Tr } P \text{ subject to} \quad (32)$$

$$\begin{bmatrix} P & (x_i - c) \\ (x_i - c)^T & 1 \end{bmatrix} \succ 0, \quad i = 1, 2, \dots, N. \quad (33)$$

Then, the ellipsoid  $\{x: (x - c_{\text{opt}})^T P_{\text{opt}}^{-1} (x - c_{\text{opt}}) - 1 \leq 0\}$  is the minimal ellipsoid containing the points  $x_i, i = 1, 2, \dots, N$ . This result has been also presented in [31], where the “size” of the ellipsoid was measured in terms of volume.

Fig. 9 shows the minimal covering ellipse for the data points, shown in the equation at the bottom of the page, which results in

$$P_{\text{opt}} = \begin{bmatrix} 0.1044 & 0.0515 \\ 0.0515 & 0.0899 \end{bmatrix}, \quad c_{\text{opt}} = \begin{bmatrix} 0.9068 \\ 0.3281 \end{bmatrix}.$$

As a simple special case of the previous problem, we may consider the minimal sphere containing given points. If  $c \in \mathbb{R}^n$  denotes the center of the sphere, and  $r \geq 0$  denotes the radius,

then the optimal sphere may be computed solving the following (convex) optimization problem in the variables  $c, r$

$$\text{minimize } r \text{ subject to} \quad (34)$$

$$\|x_i - c\| \leq r, \quad i = 1, 2, \dots, N; \quad (35)$$

$$r \geq 0. \quad (36)$$

This problem belongs to the class of second-order cone programs (SOCP), which are a particular class of SDP, and for which there exist dedicated algorithms that exploit the structure of the problem for extra efficiency, see [25], [26] and references therein.

### B. Pattern Separation by Ellipsoids

A problem of cluster classification arising in pattern recognition is that of separating two set of points  $\{x_1, \dots, x_{N_x}\}$  and  $\{y_1, \dots, y_{N_y}\}$  by means of an ellipsoidal surface, see [20]. An SDP solution to this problem has been proposed in [30]. The ellipsoid  $\{x: x^T A x + 2b^T x + d \leq 0\}$  contains all the points  $x_i$  and none of the  $y_i$  if and only if  $A \succ 0$  and

$$\begin{aligned} x_i^T A x_i + 2b^T x_i + d &\leq 0, \quad i = 1, \dots, N_x, \\ y_i^T A y_i + 2b^T y_i + d &\geq 0, \quad i = 1, \dots, N_y. \end{aligned} \quad (37)$$

These form a set of LMI conditions on the problem variables  $A, b, d$ . We can therefore find the optimal ellipsoid separating the two clusters by minimizing the size of  $A$  under the LMI constraints (37). Notice that if the two clusters are not separable

$$X = \begin{bmatrix} 0.7500 & 0.7274 & 1.1422 & 0.8786 & 0.8018 & 0.6802 & 1.2152 & 0.8077 & 1.0121 & 0.6807 \\ 0.1463 & 0.2659 & 0.5483 & 0.4766 & 0.3228 & 0.3975 & 0.5559 & 0.0414 & 0.4220 & 0.0352 \end{bmatrix}.$$

by an ellipsoidal surface, the solution algorithm will return an unfeasibility message. Fig. 10 shows an example with the two data clusters and the computed ellipsoid that separates them.

There are several variations, with interesting engineering applications, of the geometric problems mentioned in the previous sections. To mention a few, the problem of finding a minimal ellipsoid containing  $N$  given ellipsoids is treated in [3]. The minimal volume ellipsoid containing a generic set is treated for instance in [10]. The problem of computing the maximum volume ellipsoid contained inside a polyhedron described by means of a set of linear inequalities, as well as other problems related to ellipsoidal approximation are treated in [3], [12], [16], [31].

## VI. CONCLUSION

From the numerical tests, we observed that the proposed solutions based on the DOS criterion give approximation performances that are close to those provided by algorithms based on the minimization of EG criterion. These latter algorithms do not have a convergence guarantee and are computationally expensive. On the other hand, the proposed ellipsoidal fitting algorithms always guarantee convergence to the global optimum, for data in spaces of generic dimension  $n$ , and their computational complexity gracefully scales as  $O(N^2n^4)$  with data cardinality  $N$  and space dimension  $n$ . To the best of the author's knowledge, no algorithm with such characteristics (either based on algebraic or Euclidean geometric criteria) can be found in the literature, with the exception of [8], which is however restricted to data in two-dimensions.

In the test cases, the DOS-optimal solutions showed remarkable stability with respect to Gaussian noise, and also to the presence of outliers, in the case of the robust algorithm proposed in Theorem 3.

As a final word of caution, we remark that the used implementations of the SDP optimization algorithms permit to treat with ease problems of medium size, i.e., up to several hundreds of points. For data sets of larger size, one should expect numerical problems from these general-purpose SDP solvers, and should move to problem-specific SDP codes that exploit the data structure.

## REFERENCES

- [1] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *J. SIAM*, vol. 5, no. 1, pp. 13–51, 1995.
- [2] F. L. Bookstein, "Fitting conic sections to scattered data," *Comput. Graph. Image Process.*, vol. 9, pp. 56–71, 1979.
- [3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA: SIAM, 1994.
- [4] G. C. Calafiore, "Outliers robustness in multivariate orthogonal regression," *IEEE Trans. Syst., Man., Cybern. A*, vol. 30, pp. 674–679, Dec. 2000.
- [5] F. L. Chernousko, "Guaranteed estimates of undetermined quantities by means of ellipsoids," *Sov. Math. Dokl.*, vol. 21, pp. 396–399, 1980.
- [6] L. El Ghaoui and J.-L. Commeau, (1999, Jan.) LMITool Version 2.0. [Online]. Available: <http://www.ensta.fr/~gropco>.
- [7] T. Ellis, A. Abbod, and B. Brillault, "Ellipse detection and matching with uncertainty," *Image Vis. Comput.*, vol. 10, no. 2, pp. 271–276, 1992.
- [8] A. Fitzgibbon, M. Pilu, and R. B. Fiske, "Direct least square fitting of ellipses," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 476–480, May 1999.
- [9] W. Gander, G. H. Golub, and R. Strebler, "Least-squares fitting of circles and ellipses," *BIT*, vol. 34, pp. 558–578, 1994.
- [10] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*. Berlin, Germany: Springer-Verlag, 1988.
- [11] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [12] L. G. Khachiyan and M. J. Todd, "On the complexity of approximating the maximal inscribed ellipsoid for a polytope," *Math. Progr.*, vol. 61, pp. 137–159, 1993.
- [13] Y. Leedan and P. Meer, "Heteroscedastic regression in computer vision: Problems with bilinear constraint," *Int. J. Comput. Vis.*, vol. 37, pp. 127–150, 2000.
- [14] G. Lukacs, R. R. Martin, and A. D. Marshall, "Faithful least-squares fitting of spheres, cylinders, cones and tori for reliable segmentation," in *Proc. ECCV*, vol. 1, 1998, pp. 671–686.
- [15] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim, "Robust regression methods for computer vision: A review," *Int. J. Comput. Vis.*, vol. 6, pp. 59–70, 1991.
- [16] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1993.
- [17] J. O'Rourke and N. I. Badler, "Decomposition of three-dimensional objects into spheres," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, pp. 295–305, 1979.
- [18] J. Porril, "Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter," *Image Vis. Comput.*, vol. 8, no. 1, pp. 37–41, 1990.
- [19] V. Pratt, "Direct least squares fitting of algebraic surfaces," *ACM J. Comput. Graph.*, vol. 21, no. 4, 1987.
- [20] J. B. Rosen, "Pattern separation by convex programming," *J. Math. Anal. Applicat.*, vol. 10, pp. 123–134, 1965.
- [21] P. L. Rosin, "A note on the least-squares fitting of ellipses," *Pattern Recognit. Lett.*, vol. 14, pp. 799–808, 1993.
- [22] P. L. Rosin and G. A. West, "Nonparametric segmentation of curves into various representations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 1140–1153, Dec. 1995.
- [23] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. Wiley, 1987.
- [24] P. D. Sampson, "Fitting conic sections to very scattered data: An iterative refinement of the Bookstein algorithm," *Comput. Graph. Image Process.*, vol. 18, pp. 97–108, 1982.
- [25] M. S. Lobo, L. Vandenberghe, and S. Boyd, "Software for second-order cone programming," *Linear Algebra and Its Applications*, vol. 284, pp. 193–228, 1998.
- [26] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Applicat.*, vol. 284, pp. 193–228, 1998.
- [27] G. Taubin, "Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations, with applications to edge and range image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 1115–1138, Nov. 1991.
- [28] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Upper Saddle River, NJ: Prentice-Hall, 1998.
- [29] S. Ullman and R. Basri, "Recognition by linear combinations of models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 992–1006, Oct. 1991.
- [30] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, 1996.
- [31] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *J. SIAM*, vol. 19, no. 2, pp. 499–533, 1998.
- [32] L. Vandenberghe and S. Boyd, *SP: Software for Semidefinite Programming. User's Guide*. Stanford, CA: ISL, Stanford Univ., 1999.
- [33] *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds., Kluwer, New York, 2000.



**Giuseppe Calafiore** was born in Torino, Italy, in 1969. He received the "Laurea" degree in electrical engineering and the Doc. degree in information and system theory from the Politecnico di Torino in 1993 and 1997, respectively.

Since 1998, he has been an Assistant Professor with the Dipartimento di Automatica e Informatica, Politecnico di Torino. He held visiting positions at Information Systems Laboratory, Stanford University, Stanford, CA, in 1995, at Ecole Nationale Supérieure de Techniques Avancées (ENSTA), Paris, France, in 1998, and at the University of California, Berkeley, in 1999. His research interests are in the field of analysis, identification and control of uncertain systems, pattern analysis and robotics, convex optimization, and randomized algorithms.