Proceedings of the ASME 2017 International Mechanical Engineering Congress and Exposition
IMECE2017
November 3-9, 2017, Tampa, Florida, USA

IMECE2017-70899

Convexity and Optimality Conditions for Constrained Least-Squares Fitting of Planes and Parallel Planes to Establish Datums

Craig M. Shakarji
Physical Measurement Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899, U.S.A.
craig.shakarji@nist.gov

Vijay Srinivasan
Engineering Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899, U.S.A.
vijay.srinivasan@nist.gov

Abstract

This paper addresses some important theoretical issues for constrained least-squares fitting of planes and parallel planes to a set of input points. In particular, it addresses the convexity of the objective function and the combinatorial characterizations of the optimality conditions. These problems arise in establishing planar datums and systems of planar datums in digital manufacturing. It is shown that even when the input points are in general position: (1) a primary planar datum can contact 1, 2, or 3 input points, (2) a secondary planar datum can contact 1 or 2 input points, and (3) two parallel planes can each contact 1, 2, or 3 input points, but there are some constraints to these combinatorial counts. In addition, it is shown that the objective functions are convex over the domains of interest. The optimality conditions and convexity of objective functions proved in this paper will enable one to verify whether a given solution is a feasible solution, and to design efficient algorithms to find the global optimum solution.

1. Introduction

Planar datums and systems of planar datums arise frequently in the specification and verification of product geometries before, during, and after manufacturing [1-7]. Traditionally, such datum planes were established on a manufactured part using physical devices such as surface plates, angle blocks, and expanding and closing vises [8]. In the current era of digital manufacturing, one faces the task of establishing datums by *fitting* planes and lines to a cloud of points, which may number in the millions, sampled on a manufactured part using coordinate measuring systems (CMS). Standards development organizations such as ISO (International Organization for Standardization) and ASME are now responding to this trend by moving beyond supporting merely analog (i.e., physical) inspection devices to more general standards definitions that also support such digital (i.e., coordinate measurement) technologies. Experts in ISO/TC 213 and ASME Y14 standards committees are now engaged in defining the proper fitting criteria that simulate in the digital world what has been practiced in the physical world thus far.

Mathematically and computationally, fitting can be viewed as an optimization problem. Least-squares (including total least-squares) fitting of lines and planes has a long and colorful history over the past two centuries in many fields of science and engineering [9-11]. In computational coordinate metrology, least-squares fitting has enjoyed an enduring appeal [12, 13] that got a boost from recent interest in the form of constrained least-squares fitting to establish datums. Fueled by urgent requests from ISO and ASME standards committees, recent research has explored some of the theoretical and algorithmic issues of constrained least-squares fitting [14, 15]. These investigations were directed towards planar datums and systems of planar datums, which influenced the standards committees to consider constrained least-squares fitting as the default datum fitting criterion.

This paper consolidates earlier theoretical results for constrained least-squares fitting of planes, completes them with formal proofs, and extends the results to parallel planes and intersecting planes. It will address only the combinatorial characterizations of the optimality conditions and the convexity of objective functions, leaving the algorithmic details to other published sources and future research. However, the theoretical results reported this paper form the basis for much of the algorithm design and analysis.

The rest of the paper is organized as follows. Section 2 provides some motivating examples for planar datums. The optimization problem for plane fitting is then formulated in Section 3. A convex hull filter is introduced in Section 4, wherein a summary of combinatorial characterizations of

optimality conditions for various choices of objective functions is also provided. Section 5 gives brief descriptions of Gauss maps and linear maps that will be exploited later in this paper. Section 6 proves the optimality conditions for fitting lines and planes to establish datums. Convexity of the objective functions is the subject of Section 7. Thus the major contributions of this paper are contained in Sections 6 and 7. Finally, Section 8 summarizes the results of the paper and offers some directions for future research.

2. Motivating Examples

Consider the problem of specification and verification involving datum planes. Figure 1(a) shows how a designer may graphically present the specification of position tolerancing of a cylindrical hole in a part with respect to a system of primary and secondary planar datums. Figure 1(b) illustrates how such a system of primary and secondary datum planes may be established on a manufactured instance of the part.

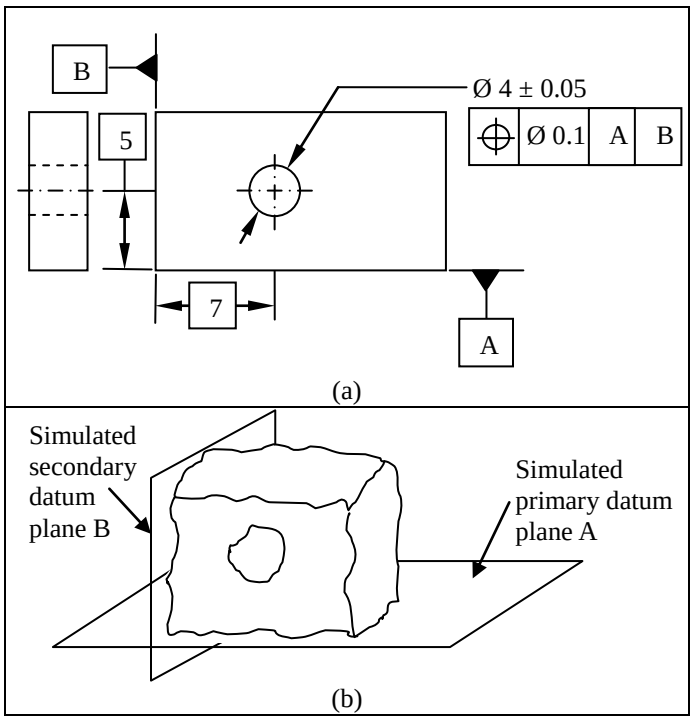

Figure 1. A simple example of (a) specification during design of a part, and (b) establishment of a system of primary and secondary datum planes on a manufactured instance of the part. The secondary datum plane B is required to be perpendicular to the primary datum plane A.

In the current era of digital manufacturing, a manufactured part, such as the one shown in Fig. 1(b), can be scanned by a CMS to generate a large number of discrete points, each with three-dimensional Cartesian coordinates. Those points corresponding to the datum feature A can then be processed to fit a primary datum plane A, and similarly for a secondary datum plane B. (For the sake of simplicity, a tertiary datum is not shown in Fig. 1.) The optimization problem for performing the plane fitting can be defined in several ways, depending on the choice of the objective function. In all cases, the datum planes are required to lie outside the material of the manufactured part while remaining as close to the part as possible (where 'as close as possible' is determined by the objective function).

It is also possible to specify datum planes for slabs and slots, as indicated in Fig. 2. Here two sets of points, each measured on each of the two features that correspond to the parallel plane features, are subjected to fitting by two parallel planes; the datum plane is the *median* plane of the two parallel planes that are thus fitted. In both cases involving slabs and slots, the fitted parallel planes are required to lie outside the material of the manufactured part while remaining as close to the part as possible. These median planes can also be used as secondary (or tertiary) datums; in that case additional constraints, such as the secondary datum being perpendicular to a primary datum, will be enforced.

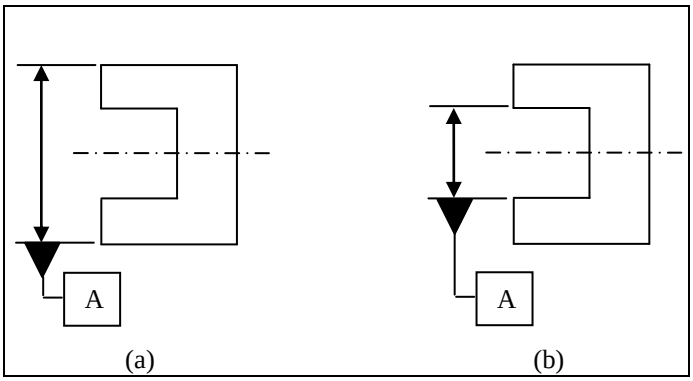

Figure 2. Specification of a datum plane for (a) a slab (external width) and (b) a slot (internal width). In each case, the datum is the median plane (indicated by dashes and dots) of two parallel planes that are fitted to two sets of points on two planar surface features (indicated by extension lines) on a manufactured instance of the part.

When a datum plane is a primary datum, the fitting is an optimization problem involving input points in space. This will also be referred to as a 3D (three-dimensional) problem in this paper. When a datum plane is a secondary datum, as datum B shown in Fig. 1, the fitting may be reduced (by projecting input points onto the primary datum plane) to an optimization problem involving points that lie on a (primary datum) plane; in this case, it is a line fitting problem for a set of input points in a plane. This will be referred to as a 2D (two-dimensional) problem in this paper. So in Fig. 1(b), the set of points measured on a manufactured part that correspond to the datum feature B are projected perpendicularly onto the datum plane A, to provide the input set of points for fitting a line.

A similar situation arises for establishing secondary datums involving slabs and slots. Here again, measured points on a manufactured part can be projected perpendicularly to a

primary datum plane, and parallel lines will be fitted to the two sets of input points in a plane. Therefore, the 2D problems of fitting lines and parallel lines to input points in a plane will be treated as equally important as 3D problems in the rest of the paper. In fact, several illustrations will deal with such line fittings in 2D while paying careful attention to the extendibility of the ideas to plane fittings in 3D.

3. Formulation of the Problem

Consider an arbitrary, continuous surface patch S in space and a plane P as shown in Fig. 3(a). Here $d(p,P)$ indicates the perpendicular distance of any point p in S to P. An infinitesimal area element around p is indicated as dA. Similar notations for an arbitrary, continuous curve C and a line L in a plane are shown in Fig. 3(b). Here $d(p,L)$ indicates the perpendicular distance of any point p in C to L. An infinitesimal line element around p is indicated as dl. It is assumed that the material of a manufactured part under consideration lies to one side of the surface patch S (and, similarly, to one side of the curve C). So when the plane P (or line L) is said to lie outside of the material, it should be apparent which side of S (or C) the plane P (or line L) should lie.

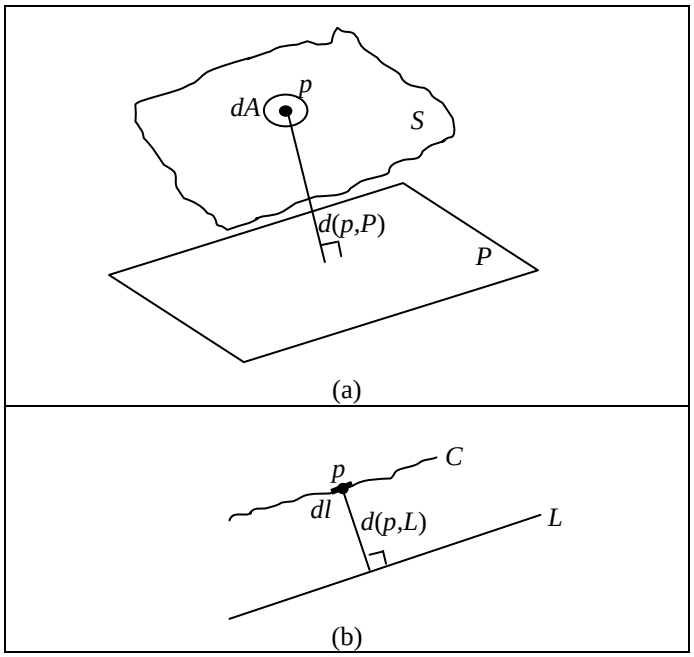

Figure 3. Notations for (a) fitting a plane, and (b) fitting a line.

A similar set of notations can be defined for parallel planes and parallel lines. For the sake of simplicity, only the case of parallel lines in a plane is illustrated in Fig. 4; the notations for the case of parallel planes in space can be inferred easily from this figure. Referring to Fig. 4, $d(p_1,L_1)$ denotes the perpendicular distance of any point p_1 in C_1 to line L_1, and $d(p_2,L_2)$ denotes the perpendicular distance of any point p_2 in C_2 to line L_2.

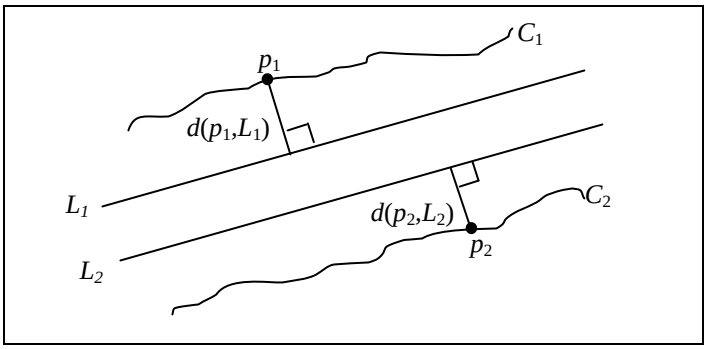

Figure 4. Notations for fitting parallel lines.

With these notations, and referring to Fig. 3(a), the constrained least-squares plane fitting problem can be posed as the following optimization problem:

$$\min_{P} \int_{S} d^2(p,P)\,dA \tag{1}$$

subject to P lying outside the material.

Similarly, referring to Fig. 3(b), the constrained least-squares line fitting problem can be posed as

$$\min_{L} \int_{C} d^2(p,L)\,dl \tag{2}$$

subject to L lying outside the material.

Posing the optimization problems for constrained least-squares fitting of parallel planes and parallel lines follows a similar pattern. For parallel planes P_1 and P_2, it is

$$\min_{P_1,P_2} \left[\int_{S_1} d^2(p_1,P_1)\,dA + \int_{S_2} d^2(p_2,P_2)\,dA \right] \tag{3}$$

subject to P_1 and P_2 being parallel, and both P_1 and P_2 lying outside the material.

For parallel lines L_1 and L_2, it is

$$\min_{L_1,L_2} \left[\int_{C_1} d^2(p_1,L_1)\,dl + \int_{C_2} d^2(p_2,L_2)\,dl \right] \tag{4}$$

subject to L_1 and L_2 being parallel, and both L_1 and L_2 lying outside the material.

The search space for planes and lines in the optimization problems posed in Eqs. (1) through (4) can be simplified considerably by the fact that the optimal planes and lines must contact at least one point of the input set of surface patches and

curves. For, if not, the plane (line) could be moved parallel to itself towards the input set while reducing the objective function of Eqs. (1) through (4), till it contacts at least one point of the input set. Such planes and lines have a special name. A plane (line) that contacts at least one point of an input set, which can be continuous or discrete, and keeps all the other points of the input set either on it or to one side of it, is called a *supporting* plane (line) of the input set [16]. Hence a simple result that gives a necessary condition for optimality follows:

Lemma 0: Solutions to the optimization problems posed in Eqs. (1) through (4) must be supporting planes or supporting lines to the input set of surface patches or curves, respectively.

Supporting lines and supporting planes generalize the concept of tangents to both discrete and continuous sets of points. The simple Lemma 0 also serves as the first combinatorial characterization of the optimality condition that reduces the search space. There is another subtle, but important, condition that can be associated with the supporting lines and planes. It is that the supporting line *L* (or supporting plane *P*) also keeps the immediate, infinitesimal material neighborhoods of the input curve (or surface patch) either on or to one side of *L* (or *P*). This removes any ambiguity as to which supporting lines and planes lie *outside* the material, and cuts down the search space of supporting lines and planes even further.

4. Convex Hull Filters and Comparison of Various Objective Functions

The optimization problems posed in Section 3 are quite appropriate for datum specification purposes during the design phase, when only perfect or imperfect continuous geometric entities are considered. When manufactured parts are measured using CMS for the purpose of verification, such lofty continuous objectives can seldom be realized. The output from CMS is usually a set of three-dimensional coordinates of discrete points, which may number in the millions. A practical and ingenious way to bridge the discrete world of CMS to the continuous world of surface patches and curves defined in Section 3 has been devised by the ISO and ASME standards committees using the convex hull filter. It is worth noting that the choice of convex hull for filtering is strongly motivated by Lemma 0.

Let $S = \{p_1, p_2, \dots, p_n\}$ be a set of input points from measurements made on a manufactured surface patch or a curve. For the sake of theoretical convenience, and without any loss of generality, let these input points lie 'more or less' on a horizontal plane or a horizontal line. Also let *CH*(*S*) be the convex hull of *S*. Depending on the dimension, *CH*(*S*) can be formed as either (in 3D) a convex polyhedron with triangular faces and hull edges that are line segments, or (in 2D) a convex polygon. If the input points are in general position, then *CH*(*S*) will consist of hull vertices, hull edges that are line segments, and hull faces that are triangles. (If the points are not in general position, then the hull edges and faces can be partitioned into line segments and triangles, respectively.)

Consider the 2D problem first. If *CH*(*S*) is a convex polygon in a plane, then every point of a hull edge (besides the end points) has one and only one supporting line – namely the one that contains that edge. There can be an infinite number of supporting lines through a hull vertex; the outward-pointing unit normals to these supporting lines form a wedge (a two-dimensional cone) with the hull vertex at the apex. However, not all the supporting lines will be candidates mentioned in Lemma 0. Only those supporting lines that lie outside the material will qualify for further consideration. This allows a partitioning of *CH*(*S*) into an *outer* part of the convex hull (which is simply referred to in this paper as the 'outer convex hull') and an *inner* part of the convex hull ('inner convex hull'), with only the outer convex hull contributing to the search space of supporting lines referred to in Lemma 0.

A similar observation can be made if *CH*(*S*) is a convex polyhedron in space. Every point in the interior of a triangular hull face has one and only one supporting plane – namely the one that contains that face. There can be an infinite number of supporting planes through a hull edge or hull vertex. In each case, the outward-pointing unit normals to these supporting planes form a wedge or a cone with the corresponding hull edge or hull vertex at the apex. Here again, using the material neighborhoods, a partitioning of *CH*(*S*) into an outer part of the convex hull and an inner part of the convex hull can be effected, with only the outer convex hull contributing the search space of supporting planes referred to in Lemma 0.

The vertices of an outer convex hull of a given set of input set *S* of points form a subset of *S*. The rest of the points in *S* need not be considered further for the purpose of datum establishment. With this fact in mind, it is possible to compare the optimality conditions for various optimization criteria that have been studied in literature and practiced in industry. Table 1 gives a summary of the combinatorial characterizations of the optimality conditions for fitting lines and parallel lines in terms of the minimum number of contact with the outer convex hull(s). Table 2 provides a similar summary for planes and parallel planes. In the case of parallel lines and parallel planes, they are 'inscribed' for slots and 'circumscribed' for slabs.

In Tables 1 and 2, the indicated minimum contacts with the outer convex hulls occur when the input points are in general position. In degenerate cases, where the points are not in general position, more faces, edges, and vertices can contact the fitted planes. The utility of the results summarized in Tables 1 and 2 lies in the fact that these results may significantly reduce the search space to find solutions to the optimization problems of Eqs. (1) through (4). So it is worth discussing these optimization problems in some detail.

Various optimization problems shown in Tables 1 and 2 have a long and colorful history in engineering practice. The optimality conditions displayed in these tables have been exploited in commercial software to deliver practical solutions to industry. The optimization problems designated as CL_1P and CL_1L provide solutions that are in harmony with the 3-2-1 contacts for primary, secondary, and tertiary datum plane system, popularized in the engineering folklore [1]. Such

popularity is due to the apparent mechanical stability offered by the contacting planes to the manufactured parts. These plane and line fittings are also known as fitting under the L_1-norm.

Table 1. Comparison of optimality conditions for constrained least-squares fitting of lines and parallel lines to the outer convex hull(s).

	Designation	Optimization problem	Minimum contact with outer convex hull(s)	Select Ref.
Lines	CL_1L	Minimize integral of the distances to the supporting line.	edge	[17]
	SL_2L	Minimize integral of the square of the distances to a line without any constraint, then shift the fitted line to the outermost point(s).	vertex	[13,18]
	MZL	Minimize the maximum distance to the supporting line.	vertex	[19]
	CL_2L	Minimize integral of the square of the distances to the supporting line, per Eq. (2).	vertex or edge	[14,15, *]
Parallel lines	MILL	Maximize the distance between two parallel, inscribing lines.	vertex on each line.	[20]
	MCLL	Minimize the distance between two parallel, circumscribing lines.	vertex on one line and edge on the other line.	[19]
	SL_2LL	Minimize the integral of squares of distances to two parallel lines, and shift them to the outermost point(s).	vertex on each line.	[13]
	CL_2ILL	Minimize the sum of integrals of squares of distances to two parallel, inscribing lines, per Eq. (4).	(1) vertex on each line, or (2) vertex on one line and edge on the other line.	[*]
	CL_2CLL	Minimize the sum of integrals of squares of distances to two parallel, circumscribing lines, per Eq. (4).	(1) vertex on each line, or (2) vertex on one line and edge on the other line.	[*]

* indicates this paper.

The optimization problems designated as MZP and MZL have been the default definitions for datum planes and lines in ISO standards. These plane and line fittings are also known as fitting under the L_∞-norm [19]. SL_2P and SL_2L employ shifting a (total) least-squares fitting, also known as fitting under the L_2-norm, to the outermost point(s); these are well-known in research literature and lend themselves to elegant software implementations [13, 18]. Such least-squares fitting are also known for their numerical stability; that is, small changes in the input set of points result in only small changes in the fitted lines and planes. Their constrained counterparts, namely CL_2P and CL_2L, are relatively new and are the focus of recent research [14, 15] and this paper. These constrained least-squares fittings seem to combine the benefits of mechanical stability and numerical stability.

Table 2. Comparison of optimality conditions for constrained least-squares fitting of planes and parallel planes to the outer convex hull(s).

	Designation	Optimization problem	Minimum contact with outer convex hull(s)	Select Ref.
Planes	CL_1P	Minimize integral of the distances to the supporting plane.	face	[17]
	SL_2P	Minimize integral of the square of the distances to a plane without any constraint, then shift the fitted plane to the outermost point(s).	vertex	[13,18]
	MZP	Minimize the maximum distance to the supporting plane.	vertex or edge	[19]
	CL_2P	Minimize integral of the square of the distances to the supporting plane, per Eq. (1).	vertex, edge, or face	[14,15, *]
Parallel planes	MIPP	Maximize the distance between two parallel, inscribing planes.	vertex on each plane	[20]
	MCPP	Minimize the distance between two parallel, circumscribing planes.	(1) vertex on one plane and face on the other plane, or (2) edge on each plane	[19]
	SL_2PP	Minimize the integral of squares of distances to two parallel planes, and shift them to the outermost point(s).	vertex on each plane	[13]
	CL_2IPP	Minimize the sum of integrals of squares of distances to two parallel, inscribing planes, per Eq. (3).	(1) vertex on each plane, or (2) edge on each plane, or (3) vertex on one plane and edge on the other plane, or (4) vertex on one plane and face on the other plane.	[*]
	CL_2CPP	Minimize the sum of integrals of squares of distances to two parallel, circumscribing planes, per Eq. (3).	(1) vertex on each plane, or (2) edge on each plane, or (3) vertex on one plane and edge on the other plane, or (4) vertex on one plane and face on the other plane.	[*]

* indicates this paper.

A similar comparison can be made for fitting parallel planes and lines, as shown in Tables 1 and 2. MIPP and MILL have been the default fitting for slots (as shown in Fig. 2(b) for internal widths) in ASME and ISO standards. These optimization problems can be posed as quadratic programming problems [20]. MCPP and MCLL have also enjoyed the default status for fitting in ASME and ISO standards for slabs (as shown in Fig. 2(a) for external widths); these optimization problems can be posed as computations of widths of sets and are well studied in research literature [19]. SL_2PP and SL_2LL are the shifted versions of least-squares (under L_2-norm) fitting of parallel planes and parallel lines, which are otherwise unconstrained [13]. The problems designated as CL_2IPP, CL_2CPP, CL_2ILL, and CL_2CLL are relatively new, and are discussed in the rest of the paper along with CL_2P and CL_2L.

5. Gauss Maps and Linear Maps

Unconstrained (total) least-squares fitting of planes and lines to a set of discrete points or to an outer convex hull has been well explored in literature [13]. Results from these explorations can be extended to the constrained least-squares fitting of planes and lines, but these extensions require considerations of Gauss maps and linear maps, as described in Sections 5.1 and 5.2, respectively. These maps are combined to construct composite ellipsoids in Section 6 to provide the optimality conditions for constrained least-squares fitting.

5.1 Gauss Maps

Recall that supporting lines and supporting planes to the outer convex hulls form the feasible solutions to the optimization problems posed in Eqs. (1) through (4). There is a useful mapping, called the Gauss map, between these supporting planes (or lines) to points on a unit sphere (or circle) as shown in Fig. 5. Here, the outward-pointing unit normal of each supporting plane (or line) is mapped to a unique point on the unit sphere (or circle).

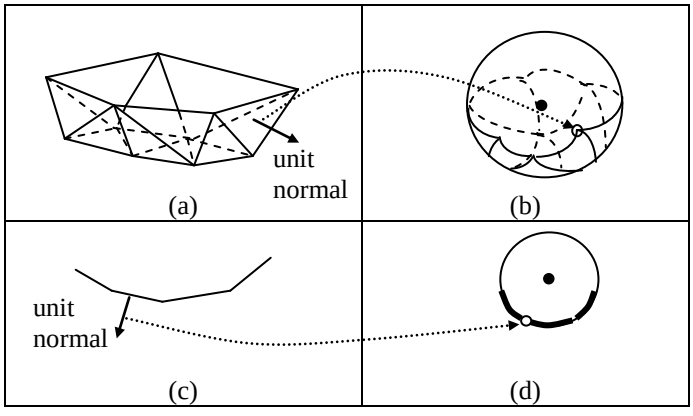

Figure 5. Gauss map obtained by mapping each unit normal of a supporting plane of an outer convex hull in space in (a) to a point on a unit sphere in (b). A similar Gauss map is obtained by mapping each unit normal of a supporting line of an outer convex hull in a plane in (c) to a point on a unit circle in (d).

For example, the outward-pointing unit normal *v* to a face of an outer convex hull in Fig. 5(a) is mapped to a point *p* on the unit sphere in Fig. 5(b), where the translated unit normal vector *v* starts at the center of the unit sphere and ends at the point *p* on the unit sphere. A similar mapping can be seen in Figs. 5(c) and 5(d), where an outward pointing unit normal *v* to an edge of an outer convex hull is mapped to a point *p* on a unit circle; the translated unit normal vector *v* starts at the center of the unit circle and ends at the point *p* on the unit circle.

The Gauss map partitions the unit sphere in Fig. 5(b) into faces, edges, and vertices that are (graph theoretic) dual to the vertices, edges, and faces, respectively, of the outer convex hull in Fig. 5(a). Similarly, the Gauss map partitions the unit circle in Fig. 5(d) into edges and vertices that are dual to the vertices and edges, respectively, of the outer convex hull in Fig. 5(c). Invoking the analogy of a global map on a sphere, the Gauss map in Figs. 5(b) and 5(d) can be viewed as covering the 'southern hemisphere' that includes the South Pole. If the outer convex hull in Figs. 5(a) and 5(c) were to be turned downside up, the Gauss map can be viewed as covering the 'northern hemisphere' that includes the North pole.

The incidence and adjacency relationships of the vertex-edge-face graph structure of the convex hull are preserved in the adjacency and incidence relationships of the face-edge-vertex graph structure in the Gauss map. The Gauss map is unique (injective, that is, one-to-one) for convex hulls, and is a powerful conceptual and computation tool to discuss the optimization problem of constrained least-squares fitting.

The topic of Gauss map is general, and is well developed in the mathematical literature. The Gauss map as visualized on a unit sphere has an appealing geographical metaphor in the form of a globe partitioned into regions. Additional notions such as the equator, and the North and South Poles on the unit sphere help the conceptual development of solutions to the optimization problem later in this paper.

5.2 Linear Maps

A linear map *M* is a linear transformation of vectors from $\mathbb{R}^n$ to $\mathbb{R}^m$. It can be represented by a matrix $M \in \mathbb{R}^{m\times n}$. Let $m > n$ for the purpose of this paper. The singular value decomposition (SVD) of *M* is given by

$$M = U\Sigma V^T, \tag{5}$$

where $U \in \mathbb{R}^{m\times m}$ is a left singular matrix, $\Sigma \in \mathbb{R}^{m\times n}$ is a diagonal matrix of singular values, and $V \in \mathbb{R}^{n\times n}$ is a right singular matrix [21-23]. Both U and V are ortho-normal matrices. By convention, the *n* singular values $\{\sigma_i, i = 1, \dots, n\}$ occupy the diagonal of the top $n \times n$ block of Σ and the remaining cells of Σ are filled with zeros. Also, the right *m*-*n* columns of *U* are filled with arbitrary ortho-normal columns to maintain the relation $U^TU = I$.

The particular relationship between a left singular vector u_j and a right singular vector v_j can be obtained from Eq. (5) as

$$Mv_j = \sigma_j u_j. \quad (6)$$

The SVD of M is closely related to the eigenvalue problem of $M^{\mathrm{T}}M$. The eigenvalues of $M^{\mathrm{T}}M$ are the squares of the singular values of M, and the eigenvectors of $M^{\mathrm{T}}M$ are the right singular vectors of M. So it follows that

$$(M^{\mathrm{T}}M)v_j = \sigma_j^2 v_j. \quad (7)$$

A geometrical interpretation of the SVD in Eq. (5) is given by the following result [23].

Theorem 1: A linear map $M \in \mathbb{R}^{m\times n}$ transforms a unit sphere in $\mathbb{R}^n$ to an ellipsoid in $\mathbb{R}^m$. The principal axes of the ellipsoid are aligned with the left singular vectors in U and the intercepts (semi-axes) of the ellipsoid with the principal axes are given by the singular values in Σ.

Figure 6 illustrates the geometric interpretation given by Theorem 1 when M is a 2×2 matrix. Under this linear map M a unit sphere (in this case, a unit circle) is mapped to an ellipsoid (in this case, an ellipse). The singular values and singular vectors of M are related by Eq. (6) as $Mv_1 = \sigma_1 u_1$ and $Mv_2 = \sigma_2 u_2$. In general, the right singular vectors are the preimages of the principal semi-axes of the ellipsoid realized in the left singular vector space.

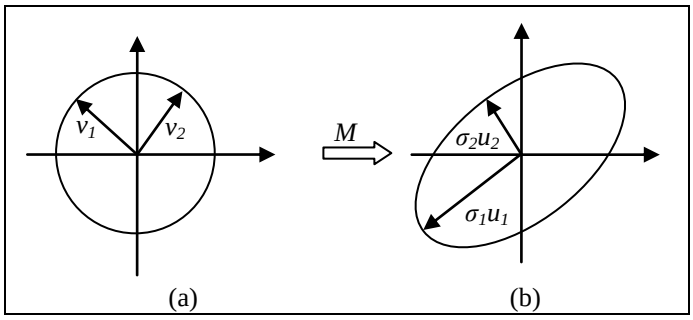

Figure 6. Geometrical illustration of the linear mapping M of (a) a unit circle to (b) an ellipse using the singular value decomposition of M.

Although the ellipsoidal nature of the linear map in the left singular vector space, as illustrated in Fig. 6(b), is very appealing, the right singular vector space is more useful for further discussion. As Eq. (7) suggests, the eigenvectors of $M^{\mathrm{T}}M$ live in the right singular vector space of M. A combination of Theorem 1 and Eq. (7) leads to the following important result for optimization.

Theorem 2: The function $v^{\mathrm{T}}M^{\mathrm{T}}Mv$ of unit vectors v on the unit sphere in $\mathbb{R}^n$ reaches a minimum value of σ_j^2 at $v = v_j$, where σ_j is the smallest singular value of M and v_j is the corresponding right singular vector.

An interesting and useful geometrical interpretation of Theorems 1 and 2 is provided by the following result in the right singular vector space.

Theorem 3: The spherical plot of the function $\sqrt{v^{\mathrm{T}}M^{\mathrm{T}}Mv}$ of unit vectors v defined on the unit sphere in $\mathbb{R}^n$ is an ellipsoid in $\mathbb{R}^n$, with the principal axes of the ellipsoid aligned with the right singular vectors of M and the semi-principal axes assuming the singular values of M.

The ellipsoid (and its specialization to an ellipse in 2D) of Theorem 3 will play a very important role in the theoretical developments of this paper.

All the results presented on linear maps in this section are general, and are well established in literature [21-23]. The main advantage of Theorem 3 is that all geometrical reasoning can be based on ellipses in 2D and ellipsoids in 3D for the constrained least-squares fitting problem addressed in the following section.

6. Optimality Conditions

The minimization achieved in Theorem 2 is made possible by analyzing a single ellipsoid described in Theorem 3. For the constrained minimization problems addressed in this paper, no single ellipsoid is sufficient. It is shown in this section that a composite ellipsoid, which is a continuous surface consisting of several ellipsoidal patches, is needed. The Gauss map and the linear map described in Section 5.1 and 5.2 will be used in constructing such a composite ellipsoid, based on some real matrices.

The case of fitting single lines and planes will be discussed in Section 6.1. The results will be extended in Section 6.2 to the case of fitting of pairs of lines and pairs of planes, where the pairs may intersect or be parallel. From now on, it should be noted that the notion of 'objective function' will implicitly include the constraints as well. This is because Lemma 0 enables automatic incorporation of the constraints by restricting the planes and lines in Eqs. (1) through (4) to be supporting planes and supporting lines, respectively.

6.1 Single Lines and Planes

The type of real matrix M dealt with in the rest of this paper will have two important features. First, it has either three columns (n = 3) when the fitting is done in three-dimensional space, or two columns (n = 2) when the fitting takes place in a plane. The second feature is that the origin of the coordinate system involved in defining the matrix M will be shifted, by pure translation, to a convenient point p for mathematical and computational purposes. To make this clear, the notation M_p will be used to indicate the point p that is used as the origin to construct that particular matrix M_p.

With these preliminaries, the objective functions defined in Eqs. (1) and (2) can be obtained by integration over the outer convex hulls of Figs. 5(a) and 5(c) treated as a surface and as a curve, respectively. For this, consider the origin of the coordinate system of the CMS to be an arbitrary point O in 3D and a plane P (or line L in 2D case) through O that has a unit

normal v. Then the integral of the square of the distance of a point on the outer convex hull from the plane P (or line L) can be expressed as the square of the second-norm of a linear map expanded as

$$\|M_O v\|_2^2 = (M_O v)^{\mathrm{T}}(M_O v) = v^{\mathrm{T}} M_O^{\mathrm{T}} M_O v, \tag{8}$$

where M_O is a matrix that will be described in some detail in the following Eqs. (9) and (10). Note the striking similarity between the objective function expressed in Eq. (8) and the function whose spherical plot is described in Theorem 3. Mathematically speaking, minimizing an objective function (as in Eq. (8) and Theorem 2) is the same as minimizing its square-root (as in Theorem 3).

In three-dimensional cases, it has been shown [14] that an exact integration of the expression in Eq. (1) over an outer convex hull, such as the one in Fig. 5(a), can be carried out using the Simpson's rule (over triangles in 3D) so that

$$M_O = \sqrt{\frac{1}{12}} \begin{bmatrix} \sqrt{A_1}x_{1A} & \sqrt{A_1}y_{1A} & \sqrt{A_1}z_{1A} \\ \sqrt{A_1}x_{1B} & \sqrt{A_1}y_{1B} & \sqrt{A_1}z_{1B} \\ \sqrt{A_1}x_{1C} & \sqrt{A_1}y_{1C} & \sqrt{A_1}z_{1C} \\ 3\sqrt{A_1}\bar{x}_1 & 3\sqrt{A_1}\bar{y}_1 & 3\sqrt{A_1}\bar{z}_1 \\ \vdots & \vdots & \vdots \\ \sqrt{A_N}x_{NA} & \sqrt{A_N}y_{NA} & \sqrt{A_N}z_{NA} \\ \sqrt{A_N}x_{NB} & \sqrt{A_N}y_{NB} & \sqrt{A_N}z_{NB} \\ \sqrt{A_N}x_{NC} & \sqrt{A_N}y_{NC} & \sqrt{A_N}z_{NC} \\ 3\sqrt{A_N}\bar{x}_N & 3\sqrt{A_N}\bar{y}_N & 3\sqrt{A_N}\bar{z}_N \end{bmatrix}, \tag{9}$$

where the outer convex hull is comprised of N triangles T_i, each having area A_i and vertices (x_{iA}, y_{iA}, z_{iA}), (x_{iB}, y_{iB}, z_{iB}), and (x_{iC}, y_{iC}, z_{iC}), their average being ($\bar{x}_i$, $\bar{y}_i$, $\bar{z}_i$).

Similarly, in two-dimensional cases it has been shown [14] that the application of Simpson's rule results in an exact integration of the expression in Eq. (2) over an outer convex hull, such as the one in Fig. 5(c), yielding the $3N \times 3$ matrix

$$M_O = \sqrt{\frac{1}{6}} \begin{bmatrix} \sqrt{L_1}(x_1) & \sqrt{L_1}(y_1) \\ 2\sqrt{L_1}\left(\frac{x_1+x_2}{2}\right) & 2\sqrt{L_1}\left(\frac{y_1+y_2}{2}\right) \\ \sqrt{L_1}(x_2) & \sqrt{L_1}(y_2) \\ \vdots & \vdots \\ \sqrt{L_N}(x_N) & \sqrt{L_N}(y_N) \\ 2\sqrt{L_N}\left(\frac{x_N+x_{N+1}}{2}\right) & 2\sqrt{L_N}\left(\frac{y_N+y_{N+1}}{2}\right) \\ \sqrt{L_N}(x_{N+1}) & \sqrt{L_N}(y_{N+1}) \end{bmatrix}, \tag{10}$$

where the outer convex hull is comprised of N line segments each having length L_i, and $N+1$ vertices each with coordinates (x_i, y_i).

The matrices in Eqs. (9) and (10) can be used to construct composite ellipsoids in 3D and composite ellipses in 2D, respectively. For simplicity of exposition, the construction of a composite ellipse is described first in some detail. Figure 7 illustrates an outer convex hull, its Gauss map, and a composite ellipse that consists of several elliptic arcs. Reference to Figs. 5(c) and 5(d) can be made for comparison.

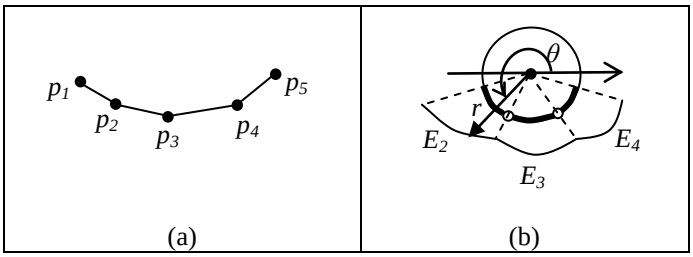

Figure 7. Illustration of (a) an outer convex hull, and (b) a composite ellipse using polar plots over a Gauss map.

The vertices of the outer convex hull in Fig. 7(a) are indicated as $\{p_1, p_2, \ldots, p_5\}$. The Gauss map of the outer convex hull is shown in Fig. 7(b) on a unit sphere (in this case, a unit circle), where the three thick circular arcs correspond to the cones of normals at vertices $p_2, p_3,$ and p_4. Now consider the vertex p_2 and the matrix M_{p_2} defined in Eq. (10). Here, note that the origin for the coordinate system is translated to the vertex p_2 to compute the coordinates involved in M_{p_2}. Then Theorem 3 defines an ellipse, and the elliptic arc E_2 in Fig. 7(b) is the restriction of this ellipse to the Gauss map corresponding to the cone of outer normals at the vertex p_2. Note that the elliptic arc E_2 is drawn as a polar plot of $r = \sqrt{v^{\mathrm{T}} M_{p_2}^{\mathrm{T}} M_{p_2} v}$ as a function of the angle θ between the unit vector v and the horizontal axis, as shown in Fig. 7(b). The other elliptic arcs E_3 and E_4 in Fig. 7(b) are constructed similarly using polar plots to obtain the composite ellipse.

Once the composite ellipse is constructed, the optimization problem reduces to finding the unit vector v (or, equivalently, a point on the unit circle) in Fig. 7(b) that yields the smallest radius in the polar plot of the composite ellipse, and a corresponding point p on the outer convex hull in Fig. 7(a) for which v is the outer normal. Then, the constrained least-squares fitting line is uniquely found as the one passing through p and having v as its normal.

This leads to only two possibilities for the optimality condition in 2D, as described in the following case analyses.

<u>*Case 1*</u>: The minimum radius of the polar plot is realized at an interior point of an elliptic arc. In this case, the constrained least-squares line contacts only the vertex of the outer convex hull responsible for that elliptic arc.

<u>*Case 2*</u>: The minimum radius of the polar plot occurs at the intersection of two adjacent elliptic arcs. In this case, the constrained least-squares line contacts the edge of the outer

convex hull that joints the two vertices that correspond to the two adjacent elliptic arcs.

Thus, the optimality condition for the 2D problem is given by the following theorem.

Theorem 4a: The constrained least-squares line for a set of input points in a plane can contact one or two points of the input set. These contact points correspond to a vertex or an edge of the outer convex hull.

A similar characterization for fitting a plane to points in space can be obtained by constructing a composite ellipsoid as an extension of Fig. 7. Starting with an outer convex hull, such as the one in Fig. 5(a), the composite ellipsoid is constructed as a spherical plot using the unit sphere, the Gauss map, and Theorem 3. The composite ellipsoid consists of a continuous collection of ellipsoidal patches $\{E_i\}$, each E_i corresponding to a vertex p_i on the outer convex hull in 3D. Here E_i arises from the ellipsoid of Theorem 3 constructed with the matrix M_{p_i} from Eq. (9). The ellipsoid itself can be constructed as a spherical plot of the radius $r = \sqrt{v^{\mathrm{T}} M_{p_i}^{\mathrm{T}} M_{p_i} v}$ as a function of two angles (say, the latitude θ and longitude φ) associated with the unit vector v in the unit sphere. E_i is then the restriction of this ellipsoid to the Gauss map corresponding to the cone of outer normals at the vertex p_i.

In seeking the minimum radius of the spherical plot that represents the composite ellipsoid, there are only three possibilities, as described in the following case analyses.

Case 1: The minimum radius of the spherical plot is realized at an interior point of an ellipsoidal patch. In this case, the constrained least-squares plane contacts only the vertex of the outer convex hull responsible for that ellipsoidal patch.

Case 2: The minimum radius of the spherical plot occurs at the intersection of two adjacent ellipsoidal patches. In this case, the constrained least-squares plane contacts the edge of the outer convex that joints the two vertices that correspond to the two adjacent ellipsoidal patches.

Case 3: The minimum radius of the spherical plot occurs at the intersection of three adjacent ellipsoidal patches. In this case, the constrained least-squares line contacts the triangular face of the outer convex that is defined by the three vertices that correspond to the three adjacent ellipsoidal patches.

Thus, the optimality condition for the 3D problem is given by the following theorem.

Theorem 4b: The constrained least-squares plane for a set of input points in space can contact one, two, or three points of the input set. These contact points correspond to a vertex, an edge, or a face of the outer convex hull.

Theorems 4(a) and 4(b) form the foundations for some of the optimality conditions enumerated in Tables 1 and 2. In particular, comparisons can be made between CL_2L and all the other line fittings (CL_1L, SL_2L, and MZL) in Table 1. Similarly, comparisons can be made between CL_2P and all the other plane fittings (CL_1P, SL_2P, and MZP) in Table 2. The advantage of the constrained least-squares fitting over the others is that it combines the benefits of mechanical stability (responsible for the popularity of CL_1L and CL_1P) and numerical stability (responsible for the popularity of SL_2L and SL_2P). This is perhaps the strongest reason for its attraction to the ASME and ISO standards community.

The process of establishing the optimality conditions of Theorems 4(a) and 4(b) has also revealed some elegant algorithmic ideas that can be used to find the optimum solution. Construction of outer convex hulls, their Gauss maps, the linear maps of Eqs. (9) and (10), and the composite ellipsoids/ellipses can all be part of an algorithm to find the constrained least-squares fitting solution.

6.2 Pairs of Lines and Pairs of Planes

The problem of fitting more than one single plane or one single line occurs while establishing datums for slabs and slots. Also, especially in ISO standards [3, 4], datums can be established for wedges and angular slots without regard to the included angle. To address these needs, Section 6.2.1 tackles the problem of fitting intersecting lines and intersecting planes. Section 6.2.2 takes up the problem of fitting parallel lines and parallel planes.

6.2.1 Intersecting Lines and Intersecting Planes

The constrained least-squares fitting of lines and planes can be used directly for establishing datum systems for wedges and angular slots. As shown in Fig. 8(a) for 2D cases, each curve of a solid wedge can be independently subjected to a constrained least-squares fitting of a line. The datum is then a system consisting of the median line (i.e., angle bisector) of the intersecting lines and a point (e.g., at the intersection of fitted lines) on the median line. A similar approach is illustrated in Fig. 8(b) for an angular slot. Note that such a datum system enables the establishment of a complete two-dimensional reference frame.

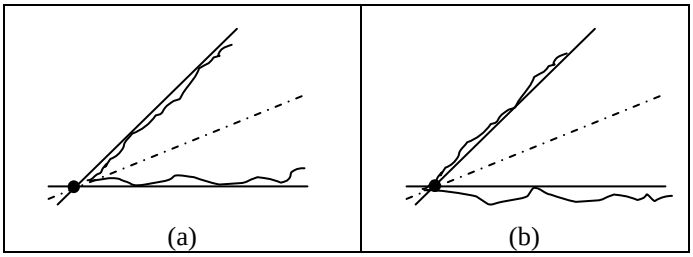

Figure 8. Illustration of datum for (a) a wedge, and (b) an angular slot. The datum is a system consisting of the median line indicated with dashes and dots, and a point indicated with a filled circle on the median line.

Extending these ideas to 3D cases of wedges and angular slots, a constrained least-squares plane can be fitted on each surface independently, resulting in intersecting planes. The datum is then a system consisting of the median plane (i.e., angle bisector) of the intersecting planes and a line (e.g., at the intersection of the fitted planes) on the median plane. Such a datum system will correspond to a prismatic class of datums [1, 4].

6.2.2 Parallel Lines and Parallel Planes

The constrained least-squares fitting of parallel lines and parallel planes can be treated in a manner similar to that of constrained least-squares fitting of single line and single plane. Starting with the optimization problems defined in Eqs. (3) and (4), the notions of outer convex hulls, Gauss maps, and linear maps can still be applied with the only additional constraint that the lines and planes be parallel. In addition, the Gauss map will have two connected components, one in the northern hemisphere and the other in the southern hemisphere of the unit sphere; also, the composite ellipsoid will have two connected components, one in each hemisphere. These two components correspond to the surfaces S_1 and S_2 in Eqn. (3), or to the curves C_1 and C_2 in Eqn. (4).

For simplicity of exposition, the 2D problem is considered first in some detail. Figure 9 illustrates a simple 2D example involving an upper convex hull and a lower convex hull as in Fig. 9(a). Both are outer convex hulls for the respective set of input points. The unit circle and the composite ellipse are shown in Fig. 9(b). The composite ellipse consists of two connected components of elliptic arcs.

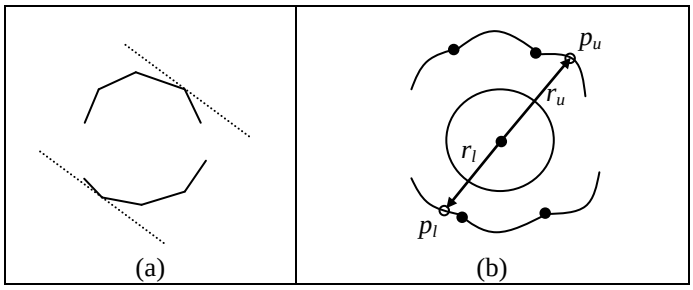

Figure 9. (a) An example of upper and lower convex hulls, and (b) the unit circle and the corresponding composite ellipse consisting of two connected components.

The two parallel supporting lines shown dotted in Fig. 9(a) define a normal direction along which two antipodal (that is, diametrically opposite) points p_u and p_l can be located on the composite ellipse as in Fig. 9(b). The corresponding radii of the composite ellipse are designated as r_u and r_l. The objective function of Eq. (4) is then given by $r_u^2 + r_l^2$. This leads to only two possibilities for the optimality condition in 2D, as described in the following case analyses.

Case 1: The minimum of $r_u^2 + r_l^2$ is realized when both the antipodal points p_u and p_l are at interiors of elliptic arcs. In this case, the constrained least-squares parallel lines contact a vertex from each of the outer convex hulls.

Case 2: The minimum of $r_u^2 + r_l^2$ is realized when one antipodal point is at an interior of an elliptic arc and the other antipodal point is at the intersection of two adjacent elliptic arcs. In this case, the constrained least-squares parallel lines contact a vertex on one outer convex hull and an edge on the other outer convex hull.

The possibility of realizing a minimum when both the antipodal points p_u and p_l are at the intersections of elliptic arcs is considered a special case, where the input points are in a degenerate position. Therefore, the edge-edge contact is not considered for the optimality conditions when the input points are in general position. Thus, the optimality condition for the 2D problem is given by the following theorem.

Theorem 5a: The constrained least-squares fitting of parallel lines to two sets of input points in a plane can have one contact point on each parallel line, or one contact point on one parallel line and two contact points on the other parallel line.

These contact points on the two outer convex hulls correspond to a vertex on each parallel line, or a vertex on one parallel line and an edge on the other parallel line. Table 1 summarizes these optimality conditions for CL_2ILL and CL_2CLL.

A similar set of case analyses can be conducted for the 3D problem of fitting parallel planes. Generalizing Fig. 9 to 3D results in a composite ellipsoid in 3D, with two connected components of ellipsoidal patches. Two parallel supporting planes, generalizing Fig. 9(a), will yield a common normal direction that defines two antipodal points p_u and p_l on the composite ellipsoid with two radii r_u and r_l. The objective function of Eq. (3) is then given by $r_u^2 + r_l^2$. This leads to only four possibilities for the optimality condition in 3D, as described in the following case analyses.

Case 1: The minimum of $r_u^2 + r_l^2$ is realized when both the antipodal points p_u and p_l are at interiors of ellipsoidal patches. In this case, the constrained least-squares parallel planes contact a vertex from each of the outer convex hulls.

Case 2: The minimum of $r_u^2 + r_l^2$ is realized when both the antipodal points are at interiors of elliptic arcs (each being the intersection of two adjacent ellipsoidal patches). In this case, the constrained least-squares parallel planes contact an edge from each of the outer convex hulls. Note that these two edges will form two skew lines in space.

Case 3: The minimum of $r_u^2 + r_l^2$ is realized when one antipodal point is at the interior of an ellipsoidal patch and the other antipodal point is at the interior of an elliptic arc (being

the intersection of two adjacent ellipsoidal patches). In this case, the constrained least-squares parallel planes contact a vertex on one outer convex hull and an edge on the other outer convex hull.

Case 4: The minimum of $r_u^2 + r_l^2$ is realized when one antipodal point is at the interior of an ellipsoidal patch and the other antipodal point at an ellipsoidal vertex (being the intersection of three adjacent ellipsoidal patches). In this case, the constrained least-squares parallel planes contact a vertex on one outer convex hull and a face on the other outer convex hull.

The possibility of realizing a minimum when both the antipodal points p_u and p_l are at the (ellipsoidal) vertices of the composite ellipsoid is considered a special case, where the input points are in a degenerate position. Therefore, the face-face contact is not considered for the optimality conditions when the input points are in general position. A similar consideration also rules out the edge-face contact. Thus, the optimality condition for the 3D problem is given by the following theorem.

Theorem 5b: The constrained least-squares fitting of parallel planes to two sets of input points in space can have one contact point on each parallel plane, or two contact points on each parallel plane, or one contact points on one parallel plane and two contact points on the other parallel plane, or one contact point on one parallel plane and three contact points on the other parallel plane.

These contact points on the two outer convex hulls correspond to a vertex on each parallel plane, or an edge on each parallel plane, or a vertex on one parallel plane and an edge on the other parallel plane, or a vertex on one parallel plane and a face on the other parallel plane. Table 2 summarizes these optimality conditions for CL_2IPP and CL_2CPP.

7. Convexity of Objective Functions

A blind application of the ideas outlined in Section 6 to find a global solution to the constrained least-squares fitting may lead to an exhaustive search over all the ellipsoidal patches of the composite ellipsoid. This can be avoided if the objective function is convex. Then only a local minimum is needed because it will also serve as the global minimum under convexity. This section shows that the objective functions are indeed convex even under the constraints of Eqs. (1) through (4). Convexity of the objective functions is indeed a critically important property for practical and efficient computation, especially in light of the numerous combinatorial conditions given by Theorems 4a, 4b, 5a and 5b, all of which would have be checked in an exhaustive search for a global minimum in the absence of convexity.

The notions of composite ellipse and composite ellipsoid are central to further discussion. As noted earlier, there is an important one-to-one correspondence among the following three entities: (1) an outward pointing normal *v* of a supporting plane (or line), (2) a point on the Gauss map etched on a unit sphere (or unit circle), and (3) a point on a composite ellipsoid (or ellipse). Recall that the faces, edges, and vertices of a Gauss map are dual to the vertices, edges, and faces of an outer convex hull. So the ellipsoidal patches (or elliptic arcs), ellipsoidal edges, and ellipsoidal vertices (or vertices between elliptic arcs) are dual to the vertices, edges, and faces of an outer convex hull.

While discussing the behavior of the objective function, frequent references will be made to points in the interior of an ellipsoidal patch and to points across (that is, at the intersection of) ellipsoidal patches to evoke a geometrical perception of the optimization problem. At those moments, imagining the correspondence of these points to the normals of supporting planes, or to points on Gauss map etched on a unit sphere, will greatly aid the comprehension of the theoretical arguments. In addition, references to the South Pole, southern hemisphere, and equatorial region will be made on the Gauss map on a unit sphere. Also, in much of the discussion, the relevant domain of interest for optimization will be a (fairly large) portion of the 'southern hemisphere' near the South Pole because that is where the solution of interest lies.

Armed with these observations, the objective functions will be shown to be convex by proving that they are convex in every neighborhood within the domain of interest. Section 7.1 considers the neighborhoods in the interior of elliptic arcs and ellipsoidal patches. Section 7.2 examines the convexity along the edges of intersection of the ellipsoidal patches. Section 7.3 then examines the neighborhoods across the elliptic arcs and ellipsoidal patches. Section 7.4 provides the convexity argument for fitting pairs of lines and pairs of planes.

7.1 Convexity in the Interior of Patches

Consider first the 2D problem of constrained least-squares fitting of lines in a plane. The composite ellipse shown in Fig. 7(b) can then be used as an example. Figure 10(a) reproduces Fig. 7(b) for the polar plot of the composite ellipse. In Fig. 10(b) a Cartesian plot of the integral of the squares of the distances (which is r^2) is presented as a function of the single degree of freedom θ. The goal is to establish that this objective function in Fig. 10(b) is convex over the entire domain of interest.

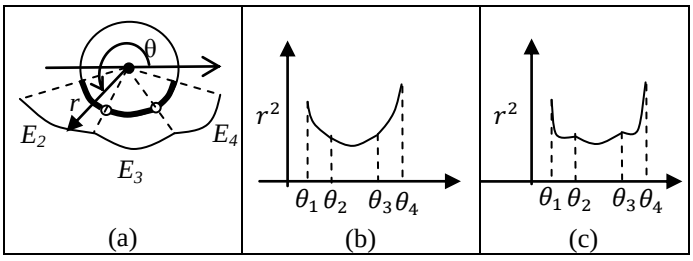

Figure 10. Illustrations of (a) a polar plot, (b) a Cartesian plot of the objective function as being convex, and (c) a non-convex objective function.

The elliptic arc E_2 of Fig. 10(a) is mapped in the interval (θ_1, θ_2) in Fig. 10(b) after squaring the radius function. Similar mappings of E_3 and E_4 can be seen in Fig. 10(b) in the subsequent intervals. In the interior of each of these intervals, such as (θ_1, θ_2), the radius r as a function of the angle θ can be seen to be convex because of the property of ellipse in the neighborhood of its minor axis (that is, near the South Pole of the unit circle in Fig. 10(a)). Since the square of a positive convex function is convex, the objective function r^2 is also convex in the interior of the interval (θ_1, θ_2). This result holds good in every one of the subsequent intervals shown in Fig. 10(b), thus establishing the following fact.

Lemma 1a: The objective function is convex at every neighborhood within each of the elliptic arcs in the composite ellipse over the domain of interest for the constrained least-squares fitting of lines in a plane.

The argument for the convexity of the objective function in the interior of the elliptic arcs can be extended to the convexity of the objective function in the interior of ellipsoidal patches in the 3D cases. Here again the radius r of an ellipsoid plotted as a Cartesian function of two degrees of freedom (say, the angles of longitude θ and latitude φ) is convex near the minor axis of the ellipsoid (that is, near the South Pole of the unit sphere). Therefore, the objective function r^2 must be convex in the interior of the patches in the θ-φ domain that correspond to the interior of the ellipsoidal patches, leading to the following fact.

Lemma 1b: The objective function is convex at every neighborhood within each of the ellipsoidal patches in the composite ellipsoid over the domain of interest for the constrained least-squares fitting of planes in space.

7.2 Convexity Along Edges of Patches

A composite ellipsoid will have elliptic arcs, with each arc as the intersection of two adjacent ellipsoidal patches. These intersection curves are pieces of planar curves (ellipses) and the radius function $r(\theta, \varphi)$ is convex in the southern hemisphere near the South Pole. So r^2 must also be convex in the domain of interest, leading to the following result.

Lemma 1c: The objective function is convex along the interior of every edge (that is, excluding its end points) of intersection of the ellipsoidal patches in the composite ellipsoid over the domain of interest for the constrained least-squares fitting of planes in space.

The objective function is convex even at the end points (that is, the vertices) of the edges of the ellipsoidal patches, as shown in the next section.

7.3 Convexity Across Patches

Again, consider the 2D problem first. Having established that the objective function is convex in the interiors of the elliptic arcs, it is now necessary to show that situations such as Fig. 10(c) do not occur *across* the elliptic arcs. For this, an interesting alternative expression of the objective function is helpful. This involves the consideration of the centroid g of the outer convex hull. Following the arguments presented in [14, 15] it can be shown that in 2D cases

$$v^{\mathrm{T}} M_p^{\mathrm{T}} M_p v = v^{\mathrm{T}} M_g^{\mathrm{T}} M_g v + L(|v^{\mathrm{T}}(p - g)|)^2, \qquad (11)$$

where L is the total length of the outer convex hull. Figure 11 provides an illustration to give a geometrical interpretation of Eq. (11) in 2D.

Consider the outer convex hull in Fig. 11 that is a reproduction of the one in Fig. 5(c). In Fig. 11 the centroid of the outer convex hull is indicated as g. Now consider a supporting line l_p for the outer convex hull through a hull vertex p and a unit normal vector v, and a line l_g that is parallel to l_p and passing through g. Let h be the distance between these two parallel lines l_p and l_g. Then, according to Eq. (11), the integral of the square of the distances of the convex hull from l_p is equal to the integral of the square of the distances of the convex hull from l_g plus the length of the convex hull times the square of the distance h between these two parallel lines. This property is sometimes referred to as the 'parallel axis theorem.' The point p can be any point in the plane and does not have to be a convex hull vertex for Eq. (11) to be valid.

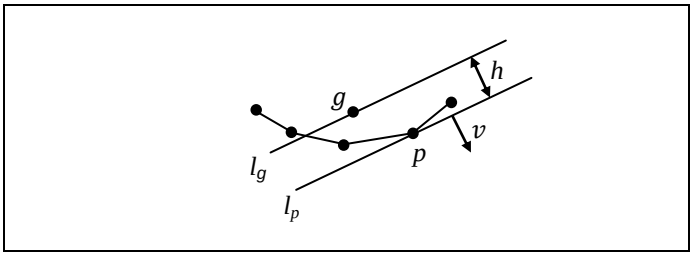

Figure 11. Illustration for the interpretation of Eq. (11).

Since the sum of convex functions is convex, the convexity of the objective function in Eq. (11) across the elliptic arcs is proved if each of the two terms on the right side of Eq. (11) is proved to be a convex function. Consider the first term $v^{\mathrm{T}} M_g^{\mathrm{T}} M_g v$ in the right side of Eq. (11). According to Theorem 3, the polar plot of its square-root is an ellipse, and its Cartesian plot with respect angle θ is convex and positive in the neighborhood of the minor axis (that is, near the South Pole). So its square is also convex even across the elliptic arcs of Fig. 10(a).

It then only remains to prove that the second term $L(|v^T(p - g)|)^2$ in the right side Eq. (11), which is the same as Lh^2, is convex across the elliptic arcs. It is possible to show that h^2 is a convex *across* the elliptic arcs (and thus Lh^2 is a convex function of the angle θ) using case analyses as illustrated in Fig. 12. In each of the three figures in the top row of Fig. 12, a continuous family of supporting lines passes through the vertex p_i and similarly another continuous family of supporting lines

passes through the vertex p_j. The transition from one elliptic arc to another occurs at the instant when the pivot for the family of supporting lines switches from p_i to p_j. That moment of transition is captured in Fig. 12 using three case analyses.

In particular, the case analyses of Fig. 12 involve the behavior of the perpendicular distance h from the centroid g to the supporting lines as a function of the rotation angle θ. Since the search for the minimum takes place in the (fairly large) vicinity of the South Pole, it can be first observed from the top row of Fig. 12 that h varies as $\cos\theta$ in the domain of interest. Therefore, the h^2-θ plots in the bottom row of Fig. 12 consist of pieces of cosine-square functions.

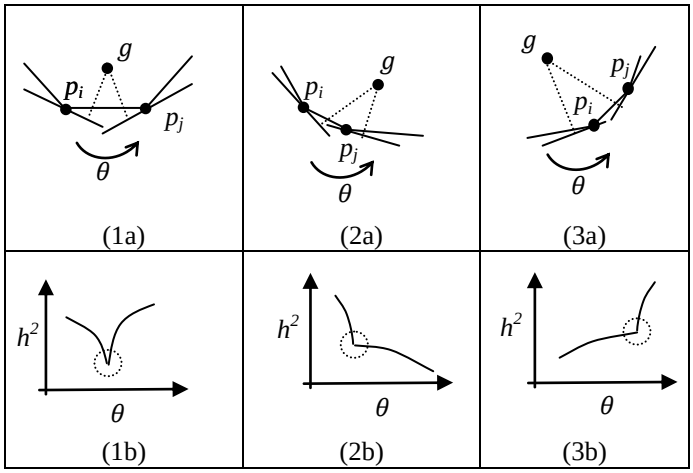

Figure 12. Illustration of three cases of transition and the related objective functions. The perpendiculars h from the centroid g to the supporting lines are shown dotted in the top row. The transition neighborhoods are shown within dotted circles in the h^2-θ plots in the bottom row.

Even though the pieces of cosine-square functions in the bottom row of Fig. 12 are themselves not convex, they behave like convex functions in the transition neighborhoods indicated within dotted circles again in the bottom row of Fig. 12. The concept of local convexity of a function is illustrated in Fig. 13. For a function *f(x)* to be convex over an arbitrary (including vanishingly small) interval (x_1, x_2) in Fig. 13, it is sufficient to show that the line-segment connecting any two points, such as *a* and *b*, on the graph of the function within that interval lies entirely above the graph of that function.

Figure 12 illustrates all the three cases of possible transitions. The first case illustrated in Figs. 12(1a) and 12(1b) shows how the perpendicular distance h from the centroid g to the supporting lines changes with respect to the rotation angle θ. As shown in the h^2-θ plot in Fig. 12(1b), the transition reaches a local minimum and is locally convex. The second case illustrated in Figs. 12(2a) and 12(2b) also shows that the transition in the h^2-θ plot is kept locally convex. The third case shown in Figs. 12(3a) and 12(3b) is similar to the second case, except for the fact the h^2-θ plot shows an increase in h^2 with respect to θ while still maintaining local convexity at the transition.

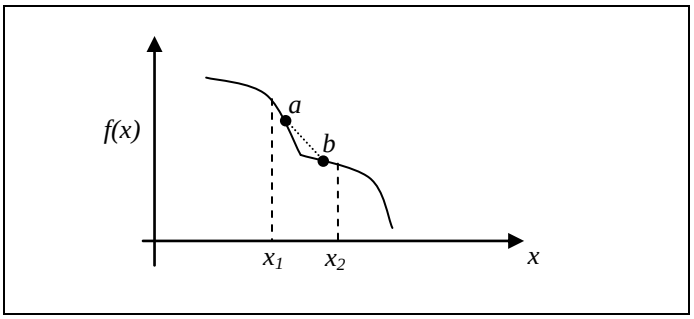

Figure 13. Illustration of local convexity of a function within an interval (x_1, x_2). The line-segment joining *a* and *b* (shown dotted) lies entirely above *f(x)*.

The behavior of the transition in the neighborhoods indicated within dotted circles in Fig. 12 can be verified by simple trigonometric calculations based on the figures shown in the top row of Fig. 12. It is worth noting that as θ increases in each of these transitions, there is a jump from a lower value for the left derivative of the function to a higher value for the right derivative of the function. Thus situations such as the ones depicted in Fig. 10(c) are avoided, and the following fact is established.

Lemma 2a: The objective function is convex at all transitions of the elliptic arcs in the composite ellipse over the domain of interest for the constrained least-squares fitting of lines in a plane.

The arguments for the convexity of the objective function across the elliptic arcs can be extended to the convexity of the objective function across the ellipsoidal patches in the 3D cases. Following the arguments presented in [14, 15] it can be shown that the equivalent of Eq. (11) in 3D cases is the following 'parallel plane theorem'

$$v^{\mathrm{T}} M_p^{\mathrm{T}} M_p v = v^{\mathrm{T}} M_g^{\mathrm{T}} M_g v + A(|v^{\mathrm{T}}(p-g)|)^2, \qquad (12)$$

where *A* is the total area of the outer convex hull. To establish local convexity of the objective function in Eq. (12) at a transition from one elliptic patch to an adjoining elliptic patch, consider any two supporting planes P_1 and P_2 each through one of any two adjacent vertices of the outer convex hull. These two planes will intersect at a line *l*. Now, transform the coordinate system so that the *z*-axis is aligned with *l*. Then the projections onto the new *xy*-plane of outer convex hull, the centroid, and the supporting planes P_1 and P_2 will yield case analyses that are identical to those shown in Fig. 12 thus establishing the local convexity of the objective function in that projected view. Since this true for *any* two supporting planes each through one of *any*

two adjacent vertices of the outer convex hull, the following fact is established.

Lemma 2b: The objective function is convex at all transitions of the elliptic patches in the composite ellipsoid over the domain of interest for the constrained least-squares fitting of planes in space.

The following theorems for convexity are then obtained by combining Lemmas 1a, 1b, 1c, 2a, and 2b.

Theorem 6a: The objective function is convex in the domain of interest for the constrained least-squares fitting of lines in a plane.

Theorem 6b: The objective function is convex in the domain of interest for the constrained least-squares fitting of planes in space.

7.4 Convexity for Pairs of Lines and Pairs of Planes

Consider the cases of wedges and angular slots shown in Fig. 8. Since the each of the intersecting lines or intersecting planes is fitted independently, the convexity proofs for objective functions provided thus far apply directly to these intersecting cases.

The cases of slabs and slots need some additional argument. Fitting parallel planes and parallel lines requires the consideration of the optimization problems posed in Eqs. (3) and (4). Note that the objective function in each of these two equations is the sum of two objective functions, each of which has been shown to be convex by Theorems 6a and 6b. Also, note that the two parallel lines or parallel planes will have equal and opposing normals due to parallelism. So, with simple linear transformation, the sum of the two functions in Eq. (3) or Eq. (4) can be posed as the sum of two convex functions over the same domain (say, involving longitude θ and latitude φ in the 3D case, and just one angle θ in the 2D case). Figure 14 illustrates this using a 2D example previously seen in Fig. 9. Since the sum of convex functions that share the same domain is convex, the following facts are established.

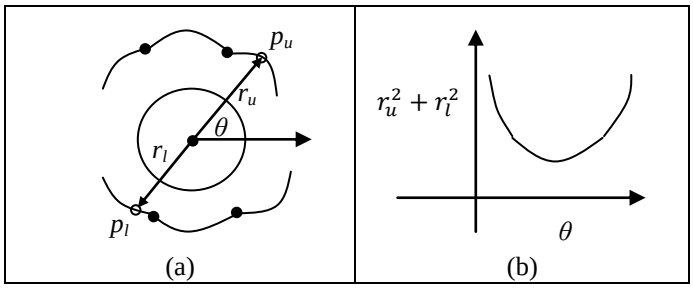

Figure 14. (a) An example of a composite ellipse in 2D, and (b) plot of the objective function $r_u^2 + r_l^2$ as a function of just one angle θ.

Theorem 7a: The objective function is convex in the domain of interest for the constrained least-squares fitting of two parallel lines in a plane.

Theorem 7b: The objective function is convex in the domain of interest for the constrained least-squares fitting of two parallel planes in space.

8. Summary and Concluding Remarks

This paper addressed the problem of establishing planar datums and systems of planar datums using constrained least-squares fitting to input points sampled on single planar features, wedges, angular slots, slabs and parallel slots. Combinatorial characterizations of the optimality conditions for these constrained optimization problems were provided with rigorous proofs. In addition, convexity of the objective functions over the domain of interest was proved.

These optimality conditions, in the form of the minimum number of points of contact, enable one to verify if a given solution is a feasible solution. In addition, the theoretical arguments that proved these optimality conditions and the proofs of convexity of the objective functions will inspire the design of new and efficient algorithms to find global optimum solutions, as demonstrated earlier using research software [14, 15]. Algorithms that exploit the results of this paper may also employ GPUs (Graphics Processing Units) to speed up the search. Further research is necessary to investigate such hardware-assisted algorithmic issues.

The constrained least-squares fitting has gained considerable support in the standards bodies owing to its theoretical advantage (as a combination of mechanical stability and numerical stability) and practical effectiveness (in software implementation). Robust implementations of constrained least-squares fitting algorithms in commercial software will be important for their acceptance by industry. Preliminary response from leading CMS vendors indicates that they are indeed beginning to implement and test their software using the results of this paper.

Acknowledgment and Disclaimer

The authors thank ISO and ASME standards experts whose advice and suggestions were invaluable in initiating and sustaining this research investigation. Any mention of commercial products or systems in this article is for information only; it does not imply recommendation or endorsement by NIST.

References

[1] ASME Y14.5-2009, Dimensioning and Tolerancing, The American Society of Mechanical Engineers, New York, 2009.

[2] ASME Y14.5.1-1994, Mathematical Definition of Dimensioning and Tolerancing Principles, The American Society of Mechanical Engineers, New York, 1994.

[3] ISO 1101:2017, Geometrical product specifications (GPS) – Geometrical tolerancing – Tolerances of form,

orientation, location and run-out, International Organization for Standardization, Geneva, 2017.
[4] ISO 5459:2011, Geometrical product specifications (GPS) – Geometrical tolerancing – Datums and datum systems, International Organization for Standardization, Geneva, 2011.
[5] Nielsen, H.S., The ISO Geometrical Product Specifications Handbook – Find your way in GPS, International Organization for Standardization, Geneva, 2012.
[6] Krulikowski, A., Fundamentals of Geometric Dimensioning and Tolerancing, 3rd Edition, Effective Training Inc., Westland, MI, 2012.
[7] Srinivasan, V., Geometrical product specification, in CIRP Encyclopedia of Production Engineering, Laperrière, L. and Reinhart, G. (Eds), Springer, 2014.
[8] Wick, C. and Veilleux, R.F., Tool and Manufacturing Engineers Handbook, Vol. IV Quality control and assembly, 4th Edition, Society of Manufacturing Engineers, Dearborn, MI, 1987.
[9] Golub, G. H. and Van Loan, C. F., An analysis of the total least squares problem, SIAM Journal of Numerical Analysis, Vol. 17, No. 6, pp. 883-893, December 1980.
[10] Van Huffel, S. and Vandewalle, J., The Total Least Squares Problem, SIAM Frontiers in Applied Mathematics, Philadelphia, 1991.
[11] Nievergelt, Y., Total least squares: State-of-the-art regression in numerical analysis, SIAM Reivew, Vol.36, No. 2, pp. 258-264, June 1994.
[12] Srinivasan, V., Shakarji, C.M. and Morse, E.P., On the enduring appeal of least-squares fitting in computational coordinate metrology, ASME Journal of Computing and Information Science in Engineering, Vol. 12, March 2012.
[13] Shakarji, C.M. and Srinivasan, V., Theory and algorithms for weighted total least-squares fitting of lines, planes, and parallel planes to support tolerancing standards, ASME Journal of Computing and Information Science in Engineering, 13(3), 2013.
[14] Shakarji, C.M. and Srinivasan, V., A constrained L_2 based algorithm for standardized planar datum establishment, ASME IMECE2015-50654, Proceedings of the ASME 2015 International Mechanical Engineering Congress and Exposition, Houston, TX, Nov. 13-19, 2015.
[15] Shakarji, C.M. and Srinivasan, V., Theory and algorithm for planar datum establishment using constrained total least-squares, 14th CIRP Conference on Computer Aided Tolerancing, Gothenburg, Sweden, 2016.
[16] O'Rourke, J., Computational Geometry in C, 2nd Edition, Cambridge University Press, 1998.
[17] Shakarji, C.M. and Srinivasan, V., Datum planes based on a constrained L_1 norm, ASME Journal of Computing and Information Science in Engineering, Vol. 15, December 2015.
[18] Butler, B.P., Forbes, A.B., and Harris, P.M., Algorithms for geometric tolerance assessment, NPL Report DITC 228/94, National Physical Laboratory, U.K., 1994.
[19] Anthony, G.T. et al., Chebyshev Best-Fit Geometric Elements, NPL Report DITC 221/93, National Physical Laboratory, U.K., 1993.
[20] Schoenherr, S., Quadratic Programming in Geometric Optimization: Theory, Implementation, and Applications, Doctoral Thesis, ETH Zurich, 2002.
[21] Golub, G. H., and Van Loan, C. F., Matrix Computations, 3rd Edition, The Johns Hopkins University Press, Baltimore, 1996.
[22] Gill, P.E., Murray, W., and Wright, M.H., Practical Optimization, Emerald Group Publishing Ltd., 1982.
[23] Trefethen, L. N., and Bau III, D., Numerical Linear Algebra, SIAM, Philadelphia, 1997.