Minimum Volume Ellipsoid Scaled to Contain a Tangent Sphere, with Application to Integrity Monitoring

Jason H. Rife
Tufts University

>> Accepted Article <<

CITATION:
J. H. Rife and D. W. Schuldt, "Minimum volume ellipsoid scaled to contain a tangent sphere, with application to integrity monitoring," *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, Monterey, CA, 2014, pp. 324-333.
doi: 10.1109/PLANS.2014.6851390

CORRESPONDING AUTHOR:
Jason Rife
jason.rife@tufts.edu

FINANCIAL SUPPORT:
Federal Aviation Administration GBAS Program (Grant FAA-10-G-006)

Minimum Volume Ellipsoid Scaled to Contain a Tangent Sphere, with Application to Integrity Monitoring

Dr. Jason H. Rife and Dieter W. Schuldt
Dept. of Mechanical Engineering
Tufts University
Medford, MA
jason.rife@tufts.edu

***Abstract*—An ellipsoid bound is introduced and proven to be a tight and conservative approximation of a sphere to which it is tangent. The bounding ellipsoid is constructed with a fixed shape that can be scaled arbitrarily (so, for example, the ellipsoid shape matches the contours of a given probability density function). The ellipsoid bound is proven conservative in that the bound always contains the sphere to which it is tangent. The ellipsoid bound is proven tight in that its volume is the minimum guaranteeing conservative bounding. Applications for the ellipsoid bound include analysis of vector integrity monitors with nominally chi-square distributions.**

Keywords- Chi-Square, Integrity, Minimum Volume Ellipsoid, RAIM, Signal Deformation Monitoring

I. Introduction

The main focus of this paper is the construction of an ellipsoid model that is a tight and conservative approximation of a sphere, when both the ellipsoid and sphere are tangent at a specified point and when the volume of the ellipsoid is scalable (but when its shape is otherwise constrained). The bound considered in this paper applies to vector spaces of arbitrary dimension and might thus also be described as a *hyperellipsoid* bound to a tangent *hypersphere*. The proposed bound has application to analysis of vector signal monitors, including many integrity monitors used in safety-of-life navigation.

The proposed bound is an example of a broader class of ellipsoid bounding problems that arises commonly in engineering. One related geometric bounding example is the problem of constructing a minimum-volume ellipsoid that contains a cluster of points, spheres, or other ellipsoids, a problem which can be solved numerically using convex optimization methods [1]. A modern library of numerical algorithms dedicated specifically to solving convex optimization problems and other geometric problems involving ellipsoids is described in [2]. Unfortunately, available numerical solvers do not yet provide standard tools to solve the problem considered here: that of determining a minimum-volume, fixed-shape ellipsoid to contain a tangent sphere.

Whereas recent research on geometric bounding has emphasized *numerical* problem solving, this paper identifies an *analytical* ellipsoid bound. This analytical bound is advantageous, as compared to a numerical bound, because it can be used to reduce computation and to decompose large optimization problems into simpler ones.

The proposed ellipsoid bound has significance in risk analysis of integrity monitors that support safety-of-life GPS navigation. Such monitors include the Signal Deformation Monitor (SDM) used in Ground-Based and Satellite-Based Augmentation Systems (GBAS and SBAS) [3]-[5], ionosphere storm monitors [6]-[7], least-squares Receiver Autonomous Integrity Monitoring (RAIM) [8]-[11], and Collaboration-Enhanced Receiver Integrity Monitoring (CERIM) [12]-[14]. A common trait of all these monitors is that they detect signal anomalies whenever an input measurement vector falls outside a spherical threshold.

Typically, integrity monitor risk is characterized using a chi-square distribution; however, care must be taken in assessing risk when the parameters describing the input noise are uncertain [15],[16]. Monitor noise is nominally chi-square because, ideally, the measurement error distribution is pre-conditioned such that the elements of the input vector are independent and identically distributed (IID), with spherical contours of constant probability. In practice, however, pre-conditioning is imperfect because of the difficulties in precisely characterizing the parameters describing the input noise distribution. (For GPS, the position-solution error varies continuously due to satellite motion, multipath, radio-frequency interference, and atmospheric weather.) To account for uncertainty in the parameters of the input noise distribution, new methods are desired. The proposed bound serves this purpose, providing a bound on integrity risk probability that is tighter than previously proposed conservative bounds [15],[16] because the shape of the bound is sculpted to match that of the monitor threshold (which is hyperspherical).

The remainder of the paper is organized as follows. The next section introduces a tight but conservative ellipsoid bound for a tangent sphere. A subsequent section applies the ellipsoid bound to integrity-monitor performance analysis. The paper concludes with a brief summary and an appendix that provides detailed proofs establishing conservatism and tightness.

II. Bounding Ellipsoid

This section introduces a locally tight yet conservative ellipsoidal bound for a sphere to which it is tangent. As described in the next section, the sphere can be regarded as the threshold in vector integrity monitoring applications.

A. Defining the Sphere and Tangent Ellipsoid

As a first step, construct an ellipsoid bound for a sphere when the two surfaces are tangent at a chosen contact point. Define the sphere to be ∂F:

$$\partial F = \{\mathbf{z} \mid f(\mathbf{z}) = 1\}. \tag{1}$$

Here f is a level-set function constructed for vectors $\mathbf{z} \in \mathbb{R}^N$ that creates a sphere of radius $\sqrt{T}$ where f is unity.

$$f(\mathbf{z}) = \frac{1}{T}\mathbf{z}^T\mathbf{z}. \tag{2}$$

Next construct ∂G to be an ellipsoid that is tangent to the sphere ∂F at a particular contact point $\mathbf{z}_c$ (see Fig. 1). This ellipsoid ∂G can be defined in terms of a level-set function g.

$$\partial G = \{\mathbf{z} \mid g(\mathbf{z}) = 1\} \tag{3}$$

The level-set function g is characterized by three parameters: a covariance matrix $\mathbf{\Lambda} \in \mathbb{R}^{N\times N}$ that determines ellipsoid shape, a scaling parameter T_g, that determines ellipsoid size, and an offset parameter $\boldsymbol{\mu}_g \in \mathbb{R}^N$ that determines ellipsoid location.

$$g(\mathbf{z}) = \frac{1}{T_g}\left(\mathbf{z}-\boldsymbol{\mu}_g\right)^T \mathbf{\Lambda}^{-1}\left(\mathbf{z}-\boldsymbol{\mu}_g\right) \tag{4}$$

It is assumed that the shape of the ellipsoid is fixed (through the definition of a constant covariance matrix $\mathbf{\Lambda} \in \mathbb{R}^{N\times N}$). Thus, the enclosed volume can be varied only by adjusting the scaling parameter T_g. Also, the covariance matrix $\mathbf{\Lambda}$ is assumed diagonal, without loss of generality, since the coordinates of $\mathbf{z}$ are arbitrary and can be chosen to be aligned with the ellipsoid principal axes. The positive diagonal elements of $\mathbf{\Lambda}$, labeled λ_i. correspond to the principal axis lengths (squared) of the ellipsoid.

In order to ensure that the two surfaces ∂G and ∂F remain in contact at the point $\mathbf{z}_c$, the ellipsoid center $\boldsymbol{\mu}_g$ must be constructed to change with the scaling parameter T_g. Because the surfaces are assumed tangent at the contact point, the gradient vector fields for the two surfaces must align at $\mathbf{z}_c$. Thus,

$$\nabla f(\mathbf{z}_c) = \alpha \nabla g(\mathbf{z}_c), \tag{5}$$

where α is a scalar. Applying this tangency constraint, by substituting (2) and (4) into (5), allows the ellipsoid center location $\boldsymbol{\mu}_g$ to be obtained as a function of the scaling parameter T_g:

$$\boldsymbol{\mu}_g = \left(\mathbf{I} - \frac{\sqrt{T_g}}{\sqrt{\mathbf{z}_c^T\mathbf{\Lambda}\mathbf{z}_c}}\mathbf{\Lambda}\right)\mathbf{z}_c. \tag{6}$$

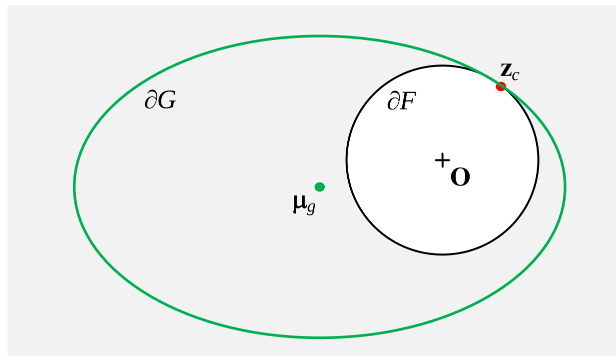

Fig. 1. Bounding Concept – An ellipsoid ∂G with fixed shape must contain a spherical volume ∂F when both surfaces are tangent at $\mathbf{z}_c$.

B. Ellipsoid Bounding Tangent Sphere

This section identifies a tight but conservative ellipsoidal bound ∂G for the sphere ∂F. To define conservative, it is useful first to identify the volumes associated with the ellipsoid and sphere surfaces. These volumes are labeled F and G and defined as follows.

$$F = \{\mathbf{z} \mid f(\mathbf{z}) \le 1\} \tag{7}$$

$$G = \{\mathbf{z} \mid g(\mathbf{z}) \le 1\} \tag{8}$$

A conservative bound is defined as one for which the ellipsoid fully contains the sphere, that is $G \supset F$.

The tight, conservative bound is found by varying the ellipse scaling parameter T_g until it is no longer true that $G \supset F$. As shown in Fig. 2, the tight bounding condition occurs either when a new tangent point appears on the surfaces ∂F and ∂G, or when the curvatures of the two surfaces become equal at the original tangent point $\mathbf{z}_c$. This process can be repeated for two cases, a first case in which the constructed contact point $\mathbf{z}_c$ has a known value and a second in which $\mathbf{z}_c$ is unknown.

When the tangent point $\mathbf{z}_c$ is constructed at a known location the tight, conservative bound occurs for $T_g = T_{g*}$.

$$T_{g*} = \frac{\mathbf{z}_c^T\mathbf{\Lambda}\mathbf{z}_c}{\lambda_{\min}^2} \tag{9}$$

A proof that the bound is tight and conservative when $T_g = T_{g*}$ is given in the appendix by Theorem 8 (Tight Bounding Theorem).

When the tangent point $\mathbf{z}_c$ may occur at any unknown location, then additional conservatism is required. To allow for any possible contact point location $\mathbf{z}_c \in \partial F$, the scaling parameter must be $T_g = T_*$.

$$T_* = \frac{\lambda_{\max}}{\lambda_{\min}^2} T \tag{10}$$

The proof that the bound is conservative for any contact point $\mathbf{z}_c$ is given in the appendix by Theorem 9 (Bounding for Unknown Contact Point).

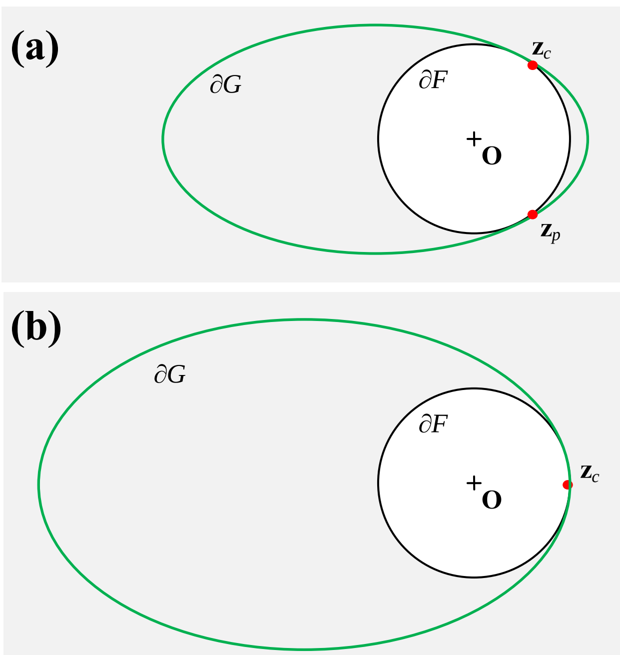

Fig. 2. Tight Bounding – As the scaling parameter shrinks, the ellipsoid ∂G becomes a tight bound under one of two conditions. Either a new tangent point appears, as shown at point $\mathbf{z}_p$ in case (a), or the curvatures of the two surfaces ∂G and ∂F become equal at the original tangent point $\mathbf{z}_c$, as shown in case (b).

III. Application to Integrity Monitoring

The problem of bounding a sphere with a tangent ellipsoid has significance to safety-of-life applications in which integrity monitors are used to verify the quality of vector signals. Typically, such signals are conditioned to be zero mean with additive random noise. As such, anomalies can be classified by detecting large deviations from zero of an input signal vector $\mathbf{z}$. To characterize these deviations, integrity monitors typically use a quadratic monitor statistic m.

$$m = \mathbf{z}^T \mathbf{z} \tag{11}$$

The monitor determines that a fault is present if the static exceeds a threshold T. In this case, an alert is issued.

$$m > T \rightarrow \text{alert} \tag{12}$$

With a few notable exceptions, such as in [14], the input vector $\mathbf{z}$ is typically decorrelated with preprocessing such that the random noise on each element of the input vector is nearly independent and identically distributed (IID). Ideally, this decorrelation step warps the noise distribution such that its isoprobability contours are spherical. In reality, the decorrelation process is imperfect, in large part because the input noise is nonstationary, due to time-varying conditions that affect the noise environment such as changing satellite location (for GPS monitoring) or radio-frequency interference (for monitoring of GPS or other radio navigation systems).

Accounting for imperfect decorrelation, and hence for input noise that is not IID, requires the analysis of input noise distributions with non-spherical isoprobability contours [15],[16]. In the case of a multivariate Gaussian density function p_N, the shape of the isoprobability contours is ellipsoidal, as characterized by a covariance matrix $\mathbf{Q}$.

$$p_N(\mathbf{z};\mathbf{Q},\boldsymbol{\mu}_z) = \left|2\pi\mathbf{Q}\right|^{-\frac{1}{2}} e^{-\frac{1}{2}\left((\mathbf{z}-\boldsymbol{\mu}_z)^T \mathbf{Q}^{-1}(\mathbf{z}-\boldsymbol{\mu}_z)\right)} \tag{13}$$

Here the distribution mean $\boldsymbol{\mu}$ is nonzero only in the presence of a fault which biases the distribution away from the center of the spherical threshold.

For integrity monitoring, it is critical to establish the probability that the monitor fails to detect a bias of a particular magnitude $b = \left\|\boldsymbol{\mu}_z\right\|$. A failed detection occurs when the bias magnitude b is larger than the threshold T at the same time that random noise pushes the monitor static m below T. The probability of missed detection can be assessed by integrating the probability distribution of the input vector $\mathbf{z}$ within a sphere of radius $\sqrt{T}$.

$$P_{md} = \int_{\mathbf{z}^T\mathbf{z} \le T} \cdots \int p_N(\mathbf{z};\mathbf{Q},\boldsymbol{\mu}_z)\, dV \tag{14}$$

Because the spherical shell is symmetric about the origin, this analysis is unchanged if a coordinate system is selected with orthogonal axes aligned with the principal axes of the covariance matrix $\mathbf{Q}$. Thus, without loss of generality, $\mathbf{Q}$ can be replaced by the positive-definite, diagonal matrix $\boldsymbol{\Lambda}$. This transformation of $\mathbf{Q}$ into $\boldsymbol{\Lambda}$ involves a simple rotation, and since rotation matrices are unitary, the length b of the mean vector $\boldsymbol{\mu}$ is unchanged. Hence,

$$P_{md} = \int_{F} \cdots \int p_N(\mathbf{z};\boldsymbol{\Lambda},\boldsymbol{\mu}_z)\, dV \tag{15}$$

Note here that (7) has been invoked to change the volume of integration to F.

System safety verification requires that the missed detection probability be computed by evaluating the integral (15) for a range of unknown biases $\boldsymbol{\mu}$ and unknown covariance matrices $\boldsymbol{\Lambda}$. Were the isoprobability contours aligned with the boundary of the spherical volume of integration F, a relative compact solution to (15) is possible. Specifically, when $\boldsymbol{\Lambda}=\mathbf{I}$, the integral results in P_{ncx}, the noncentral chi-square cumulative distribution function for N degrees of freedom.

$$\boldsymbol{\Lambda} = \mathbf{I} \rightarrow P_{md} = P_{ncx}\left(m \le T; N, b^2\right) \tag{16}$$

The noncentral chi-square distribution is tabulated by most statistical software packages.

When the isoprobability contours are elliptical and the volume of integration is spherical, the integral (15) results in a distribution is said to be generalized chi-square [17]. This distribution is significantly more involved to evaluate than the noncentral chi-square distribution, and typically requires a custom numerical integration solution (as provided in [16]).

This paper introduces a useful tool for analyzing integrity monitor performance that simplifies (15) through a conservative approximation. The approximation, denoted $\overline{P}_{md}$, is a conservative representation of the actual missed detection probability in that it is never smaller.

$$\overline{P}_{md} \ge P_{md} \tag{17}$$

To ensure conservatism as defined by (17), it is sufficient to increase the volume of integration, replaces F with a larger approximation G. The substitution is guaranteed conservative if $F \subset G$ as described by the equation below.

$$\begin{aligned} \int_G \cdots \int p_N(\mathbf{z})dV &= \int_F \cdots \int p_N(\mathbf{z})dV + \int_{G\backslash F} \cdots \int p_N(\mathbf{z})dV \\ &\ge \int_F \cdots \int p_N(\mathbf{z})dV \end{aligned} \tag{18}$$

In defining the ellipsoid G, the contact point $\mathbf{z}_c$ is assumed to be aligned with the bias $\boldsymbol{\mu}_z$.

$$\mathbf{z}_c = \frac{\sqrt{T}}{b}\boldsymbol{\mu}_z\,. \tag{19}$$

This choice makes G a relatively tight approximation of F where the probability density is highest (that is near $\boldsymbol{\mu}_z$), and allows for the bound to be looser where the probability is lowest (away from $\boldsymbol{\mu}_z$).

In a typical P_{md} analysis, the bias direction $\boldsymbol{\mu}_z$ is assumed unknown, and P_{md} is computed (for the worst-case direction) as a function of bias magnitude b. Since the bias direction is unknown, the contact point location $\mathbf{z}_c$ is also unknown. To guarantee that $F \subset G$, therefore, the level-set function $g(\mathbf{z})$ from equation (4) must be defined with $T_g = T_*$, where the scaling parameter T_* is given by (10). For $T_g = T_*$, the bias is $\boldsymbol{\mu}_z = \boldsymbol{\mu}_*$, where

$$\boldsymbol{\mu}_* = \left(\mathbf{I} - \boldsymbol{\Lambda}\frac{1}{\lambda_{\min}}\frac{\sqrt{T\lambda_{\max}}}{\sqrt{\mathbf{z}_c^T\boldsymbol{\Lambda}\mathbf{z}_c}}\right)\mathbf{z}_c\,. \tag{20}$$

To ensure a conservative approximation of P_{md}, the covariance matrix $\boldsymbol{\Lambda}$ should be set to its worst-case value, as discussed in [16].

To simplify analysis, it is now convenient to introduce a change of variables. Since the ellipsoid volume of integration G now has the same shape as the ellipsoid isoprobability contours of the distribution p_N, both can be converted into spheres through the same transformation. The appropriate change of variables is

$$\mathbf{w} = \boldsymbol{\Lambda}^{-1/2}\left(\mathbf{z} - \boldsymbol{\mu}_*\right). \tag{21}$$

For this change of variables,

$$\overline{P}_{md} = \int_{\mathbf{w}^T\mathbf{w} \le T_*} \cdots \int p_G(\mathbf{w}; \mathbf{I}, \boldsymbol{\mu}_{tot})dV\,, \tag{22}$$

where

$$\boldsymbol{\mu}_{tot} = \boldsymbol{\Lambda}^{-1/2}\left(\boldsymbol{\mu}_z - \boldsymbol{\mu}_*\right). \tag{23}$$

By analogy to (16), the integral (22) can be expressed as a noncentral chi-square distribution, where

$$\overline{P}_{md} = P_{ncx}\left(m \le T_*; N, b_*^2\right) \tag{24}$$

with

$$b_* = \frac{1}{\sqrt{\lambda_{\max}}}\left(b + \left(\frac{\lambda_{\max}}{\lambda_{\min}} - 1\right)\sqrt{T}\right). \tag{25}$$

This result is useful because (24) is simply a noncentral chi-square distribution, easily computed by standard statistical software, and because the set of distribution parameters is small and well characterized. Specifically, the parameters needed to evaluate (24) and (25) and (10) include only the monitor threshold T, the monitor degrees of freedom N, the parameterized bias magnitude b, and the extreme principal axes of the covariance matrix, $\lambda_{\max}$ and $\lambda_{\min}$.

IV. CONCLUSION

The paper introduces an ellipsoid model that is a tight, conservative approximation of a sphere to which it is tangent. The ellipsoid bound is constructed assuming that its shape is fixed but that its volume can be scaled to a minimum, beyond which further scaling would result in the ellipsoid failing to bound the sphere at one or more points.

The ellipsoid model can be applied for performance analysis of integrity monitors that alert when anomalous signals are detected. Specifically, the ellipsoid can be constructed as a model of a spherical monitor threshold, a model that is tangent (and therefore tight) at a point aligned with any fault-induced measurement bias. This construction ensures that the ellipsoid is a close approximation to the original threshold in the region where the probability density for monitor noise is highest, allowing for relatively accurate though approximate predictions of monitor missed-detection performance. Importantly, the ellipsoid bound allows for relatively simple (chi-square) analysis of monitor missed-detection performance, even when the parameters governing the monitor input noise are poorly characterized.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Federal Aviation Administration GBAS Program (Grant FAA-10-G-006) for supporting this research. The opinions discussed here are those of the authors and do not necessarily represent those of the FAA or other affiliated agencies.

APPENDIX

Theorem 1 (Large T_g Limit): *If T_g is sufficiently large and T is finite, then the volume F excluding the contact point $\mathbf{z}_c$ is contained in the interior of G.*

Background: Subsequent theorems will shrink T_g until G no longer contains F. This theorem provides a starting point by guaranteeing that G contains F if T_g is sufficiently large, with no contact points other than $\mathbf{z}_c$.

Proof: Define the hyperplane tangent to the hypersphere F through the contact point $\mathbf{z}_c$ to be H.

$$H = \left\{ \mathbf{z} \mid (\mathbf{z} - \mathbf{z}_c)^T \mathbf{z}_c = 0 \right\} \tag{26}$$

The normal vector field of the plane H is parallel to both $\nabla f(\mathbf{z}_c)$ and $\nabla g(\mathbf{z}_c)$. To simplify subsequent analysis, define $\hat{\mathbf{e}}_1$ to be a unit vector in this direction.

$$\hat{\mathbf{e}}_1 = \nabla H / \|\nabla H\| \tag{27}$$

Now introduce a mapping that transforms $\mathbf{z}$ to $\mathbf{w}$, where the first vector element in new coordinate system is the component in the $\hat{\mathbf{e}}_1$ direction.

$$\mathbf{w} = \mathbf{R}^T \mathbf{\Lambda}^{-1/2} (\mathbf{z} - \mathbf{z}_c) \tag{28}$$

This mapping performs three functions. First, shifting by $\mathbf{z}_c$ maps the contact point to the origin (such that the hyperplane H passes through the origin of the new coordinate system). Second, scaling by $\mathbf{\Lambda}^{-1/2}$ converts the volume G from a hyperellipsoid into a hypersphere. Third, multiplying by the rotation matrix $\mathbf{R}$ maps the $\hat{\mathbf{e}}_1$ direction into the first coordinate of $\mathbf{w}$ (if $\mathbf{R}$ is appropriately defined).

With this transformation, the tangent hyperplane H can be expressed

$$H = \left\{ \mathbf{w} \mid \mathbf{w}^T \hat{\mathbf{e}}_1 = 0 \right\} . \tag{29}$$

The mapping (28) is an affine transformation that preserves topology. As such, the volumes F and G remain in the negative halfspace (below H), touching H only at the contact point (i.e. at the origin of the $\mathbf{w}$ coordinate system).

Apply the mapping (28) to the level-set function $g(\mathbf{z})$. The result is

$$g(\mathbf{w}) = \left\| \mathbf{w} + \sqrt{T_g}\,\hat{\mathbf{e}}_1 \right\|^2 / T_g . \tag{30}$$

With this definition, ∂G can be expressed as

$$\partial G = \left\{ \mathbf{w} \;\middle|\; \|\mathbf{w}\|^2 + 2[\mathbf{w}]_1 \sqrt{T_g} = 0 \right\} . \tag{31}$$

In this paper, the notation $[\bullet]_i$ refers to the i^{th} component of the vector enclosed in brackets. As such, $[\mathbf{w}]_1$ is the first coordinate of the $\mathbf{w}$ vector, the distance to the plane H. By construction ∂G contains the origin; in addition, an appropriate choice of T_g allows ∂G to pass through any point in the halfspace with $[\mathbf{w}]_1 < 0$. In the limit when T_g approaches infinity, the volume G is the union of the origin and the entire halfspace with $[\mathbf{w}]_1 < 0$.

Each point in F, excluding the contact point ($F \setminus \mathbf{z}_c$), lies in the halfspace with $[\mathbf{w}]_1 < 0$. If T is finite, then T_g can be set sufficiently large such that $\{G \setminus \partial G\} \supset \{\partial F \setminus \mathbf{z}_c\}$ (You could use G^o and F^o here since the set of points G and F have been defined). In other words, the interior of G contains all points in F except for the contact point.

✈

Theorem 2 (Bounding Near Contact Point): *If $T_g \geq T_{g*}$, then, in the neighborhood of the contact point $\mathbf{z}_c$, ∂F is contained within G.*

Background: As T_g shrinks, G can fail to bound F at one of two locations: either (i) locally, near the specified contact point $\mathbf{z}_c$, or, (ii) if a new tangent point appears and then T_g shrinks further, at that point. This theorem considers the first case and establishes conditions under which bounding is guaranteed near the specified contact point $\mathbf{z}_c$.

Proof: Consider a trajectory that moves away from the contact point $\mathbf{z}_c$ in an arbitrary circular arc along the surface ∂F. Parameterizing by an angle θ, the circular arc trajectory can be represented by the equation

$$\mathbf{z} = \cos(\theta)\mathbf{z}_c + \sin(\theta)\mathbf{z}_t , \tag{32}$$

where $\mathbf{z}_t$ is defined to be orthogonal to $\mathbf{z}_c$ with the same length (such that $\mathbf{z}_c^T\mathbf{z}_t = 0$ and $\|\mathbf{z}_c\| = \|\mathbf{z}_t\| = \sqrt{T}$). Along this trajectory, the level-set function g changes by an amount $\Delta g(\mathbf{z}) = g(\mathbf{z}) - g(\mathbf{z}_c)$. Noting that the contact point is on the surface ∂G, where the level-set function is unity, the change can be simplified to $\Delta g(\mathbf{z}) = g(\mathbf{z}) - 1$. For the point $\mathbf{z}\in\partial F$ to lie inside the volume G, it must be true that $g(\mathbf{z})\leq 1$, or equivalently

$$\Delta g(\mathbf{z}) \leq 0 \rightarrow \mathbf{z} \in G , \tag{33}$$

The change in the level-set function g can be computed by integrating the gradient ∇g along the circular arc.

$$\Delta g(\mathbf{z}) = \int_0^\theta \nabla g(\mathbf{z}) \cdot \frac{\partial \mathbf{z}}{\partial \theta} d\theta , \tag{34}$$

Evaluating this integral using (4) and (32), and taking the limit for small angles θ gives

$$\Delta g(\mathbf{z}) = \frac{\theta^2}{T_g 4}\left(\mathbf{z}_t^T \Lambda^{-1} \mathbf{z}_t - \frac{T\sqrt{T_g}}{\sqrt{\mathbf{z}_c^T \Lambda^{-1} \mathbf{z}_c}} \right). \tag{35}$$

From this formula, it is evident that at small angles (i.e., in the neighborhood of $\mathbf{z}_c$), that condition (33) is satisfied if the term above in parentheses is nonpositive. Rearranging terms, (33) is satisfied if

$$\frac{\sqrt{T_g}}{\sqrt{\mathbf{z}_c^T \Lambda^{-1} \mathbf{z}_c}} \geq \frac{\mathbf{z}_t^T \Lambda^{-1} \mathbf{z}_t}{T} . \tag{36}$$

The maximum value of the right-hand side of the (36) occurs when the direction of $\mathbf{z}_t$, which is arbitrary, aligns with the shortest principal axis $\lambda_{\min}$. Hence, allowing for all possible $\mathbf{z}_t$, (33) is satisfied if

$$\frac{\sqrt{T_g}}{\sqrt{\mathbf{z}_c^T \Lambda^{-1} \mathbf{z}_c}} \geq \frac{1}{\lambda_{\min}} . \tag{37}$$

Using (9), this inequality can be rewritten $T_g \geq T_{g*}$. Since (33) is satisfied when $T_g \geq T_{g*}$, we can say that all points $\mathbf{z}$ on ∂F in the neighborhood of $\mathbf{z}_c$ are contained in G.

✈

Theorem 3 (Matching Orthant): *Consider a point $\mathbf{z}^* \in \{\partial F \cap \partial G\}$. If the sign of any component of $\mathbf{z}^*$ is opposite the sign of the corresponding component of $\mathbf{z}_c$, and if $T_g > T_{g*}$, then there exists a point $\bar{\mathbf{z}}^*$ such that $\bar{\mathbf{z}}^* \in \partial F$ and $\bar{\mathbf{z}}^* \notin G$.*

Background: Whereas the previous theorem concerned the first case of bounding failure (failure near $\mathbf{z}_c$), this theorem considers the second case (failure at a new tangent point other than $\mathbf{z}_c$). In fact, no new tangent point can appear for $T_g > T_{g*}$. This result is developed in two steps. First, this theorem (Theorem 3) establishes that new tangent points cannot appear in any orthant except the orthant containing $\mathbf{z}_c$. Next Theorem 4 will complete the contradiction, proving that a new tangent point cannot appear in the same orthant as $\mathbf{z}_c$, either.

Proof: To provide a means of comparing orthants for two vectors, define a vector function $\mathbf{s}(\mathbf{a},\mathbf{b}) \in \mathbb{R}^N$ with elements computed as follows.

$$[\mathbf{s}(\mathbf{a},\mathbf{b})]_i = \begin{cases} [\mathbf{b}]_i & 0 \leq [\mathbf{b}]_i [\mathbf{a}]_i \\ -[\mathbf{b}]_i & 0 > [\mathbf{b}]_i [\mathbf{a}]_i \end{cases} \tag{38}$$

Using this function, define the subspace J to be

$$J = \left\{ \mathbf{z} \,\middle|\, \mathbf{z} = \mathbf{s}(\mathbf{z}_c, \mathbf{z}) \right\} . \tag{39}$$

In other words, J is the closed orthant containing $\mathbf{z}_c$.

Now consider a point $\mathbf{z}^*$, where the surfaces ∂F and ∂G are tangent. Assume one element of $\mathbf{z}^*$ is opposite to the sign of the corresponding element of $\mathbf{z}_c$. Then, $\mathbf{z}^*$ is not in the same orthant as $\mathbf{z}_c$ and hence $\mathbf{z}^* \notin J$.

Also consider a point $\bar{\mathbf{z}}^* \in J$, where

$$\bar{\mathbf{z}}^* = s(\mathbf{z}_c, \mathbf{z}^*). \tag{40}$$

Combining (38) and (40), it is straightforward to show that

$$\mathbf{z}_J^T (\mathbf{z}^* - \bar{\mathbf{z}}^*) < 0 \tag{41}$$

for any point $\mathbf{z}_J \in J$.

Now evaluate the level-set function $g(\mathbf{z})$ at both $\mathbf{z}^*$ and $\bar{\mathbf{z}}^*$. Forming a difference and canceling equal quadratic terms gives

$$g(\mathbf{z}^*) - g(\bar{\mathbf{z}}^*) = \frac{2}{T_g} \boldsymbol{\mu}_g^T \boldsymbol{\Lambda}^{-1} (\mathbf{z}^* - \bar{\mathbf{z}}^*). \tag{42}$$

This difference can be rewritten

$$g(\mathbf{z}^*) - g(\bar{\mathbf{z}}^*) = \mathbf{z}_j^T (\mathbf{z}^* - \bar{\mathbf{z}}^*), \tag{43}$$

where substituting (6) for $\boldsymbol{\mu}_g$ gives

$$\mathbf{z}_j = \frac{2}{T_g}\mathbf{\Lambda}^{-1}\left(\frac{\sqrt{T_g}}{\lambda_{\min}\sqrt{T_{g*}}}\mathbf{\Lambda} - \mathbf{I}\right)\mathbf{z}_c . \quad (44)$$

In (43), $\mathbf{z}_j$ can be shown to belong to the same orthant as $\mathbf{z}_c$ ($\mathbf{z}_J \in J$), because the leading terms of (43) are positive and because the term in parentheses is also positive, since $T_g > T_{g*}$ and $\lambda_i \geq \lambda_{\min}$. Combining (41) and (43) implies that $g\left(\mathbf{z}^*\right) - g\left(\overline{\mathbf{z}}^*\right) < 0$. Furthermore, because $\mathbf{z}^* \in \partial G$, $g\left(\mathbf{z}^*\right) = 1$. Thus, $g\left(\overline{\mathbf{z}}^*\right) > 1$ and so $\overline{\mathbf{z}}^* \notin G$.

Also, because $\mathbf{z}^* \in \partial F$ by construction, $f\left(\mathbf{z}^*\right) = 1$. With

$$f\left(\mathbf{z}^*\right) = f\left(\overline{\mathbf{z}}^*\right) = \sum_{i=1}^{N}\frac{1}{T}\left[\mathbf{z}^*\right]_i^2 , \quad (45)$$

it follows that $f\left(\overline{\mathbf{z}}^*\right) = 1$ and hence that $\overline{\mathbf{z}}^* \in F$.

✈

Theorem 4 (Appearance of New Tangent Points): *If T is finite, $T_g > T_{g*}$ and ∂F is tangent to ∂G at* $\mathbf{z}_c$, then $\mathbf{z}_c$ *is the only point where the surfaces ∂F and ∂G are tangent.*

Background: As a companion to the prior theorem, this theorem completes the proof that no new tangent point appears so long as $T_g > T_{g*}$.

Proof: Consider all values of T_g, incrementally decreasing from infinity to T_{g*}. At the large T_g limit, Theorem 1 states that, given finite T, all points in F lie on the interior of G with the exception of the contact point $\mathbf{z}_c$. Hence, by construction, the only contact point present in the large T_g limit is the tangent point $\mathbf{z}_c$.

Because the surfaces ∂F and ∂G are smooth and continuously differentiable, any new contact point that appears on both surfaces as T_g shrinks must be a tangent point. Hypothesize the existence of such a tangent point $\mathbf{z}_p$. Because $\mathbf{z}_p \in \partial F \cap \partial G$,

$$f\left(\mathbf{z}_p\right) = g\left(\mathbf{z}_p\right) = 1 . \quad (46)$$

Furthermore, because the surfaces are locally tangent,

$$\nabla f\left(\mathbf{z}_p\right) = \beta \nabla g\left(\mathbf{z}_p\right), \quad (47)$$

where β is a positive scalar.

A test for the appearance of a new contact point at any value of T_g can be constructed from the two constraints (46) and (47). First re-arrange (47) to solve for $\mathbf{z}_p$. The result is

$$\left[\mathbf{z}_p\right]_i = \frac{n_i}{d_i}\left[\mathbf{z}_c\right]_i , \quad (48)$$

where the denominator d_i is

$$d_i = \frac{T_g\lambda_i}{T\beta} - 1 . \quad (49)$$

And the numerator n_i is

$$n_i = \sqrt{\frac{T_g}{T_{g*}}}\frac{\lambda_i}{\lambda_{\min}} - 1 . \quad (50)$$

If $T_g > T_{g*}$, the numerator n_i is strictly positive ($n_i > 0$), because $\lambda_i \geq \lambda_{\min}$.

Equation (48) indicates that, when the value of T_g is fixed, the set of allowable tangent points $\mathbf{z}_p$ is characterized by a single tunable parameter β. This scalar parameter β is the ratio of the level-set function gradients, as described by (47). Because the level-set gradients, which are parallel to the surface normal vectors, must both be outward facing when tangency first occurs, this parameter is positive $\left(\beta > 0\right)$. The implication is that, if a tangent point exists, it will be found by considering all positive, real values of β.

Now start at the large T_g limit and progressively increment through infinitesimally smaller values of T_g. For each T_g, it is possible to show no tangent point appears for any β. Hence no new tangent point appears so long as $T_g > T_{g*}$ and thus $\mathbf{z}_c$ is the only point where the two surfaces ∂F and ∂G are tangent.

At each incrementally decreasing value of T_g, it can be shown that no tangent point appears by considering four subsets that span the range of allowable β values. These subsets are: $\beta \in \left(0, \alpha\right)$, $\beta = \alpha$, $\beta \in \left(\alpha, \gamma\right]$, $\beta \in \left(\gamma, \infty\right)$, where the parameter

$$\alpha = \frac{\sqrt{T_g T_g^*}\lambda_{\min}}{T} , \quad (51)$$

and where the parameter

$$\gamma = \frac{T_g\lambda_{mnz}}{T} . \quad (52)$$

Note that λ_{mnz} is the smallest of the λ_i (diagonal elements of Λ) that corresponds to a non-zero element of $\mathbf{z}_c$. Symbolically,

$$\lambda_{mnz} = \min_{i \in \left\{ i \mid \left[\mathbf{z}_c\right]_i \neq 0 \right\}}\left(\lambda_i\right) . \quad (53)$$

For each of the four subsets of β values, suppose that $\mathbf{z}_p$ is a tangent point and establish a contradiction by showing that at least one of the requirements (46) and (47) is violated. To simplify evaluation of the gradient requirement, (47), consider its reformulation as (48) and invoke the Matching Orthant Theorem (Theorem 3). Theorem 3 states that a new contact point only appears in a different orthant from $\mathbf{z}_c$ if a corresponding contact point has already appeared (at a larger T_g value) in the same orthant as $\mathbf{z}_c$. Hence, if no new tangent point has appeared at a larger T_g value, then the sign of all elements of $\mathbf{z}_p$ must match the sign of the elements of $\mathbf{z}_c$, and hence, the ratio of the n_i and d_i terms in (48) must be non-negative. Noting also that $n_i > 0$ for $T_g > T_{g*}$,

$$d_i \geq 0 \;\; \forall \; \left\{ i \;\middle|\; \left[\mathbf{z}_c\right]_i \neq 0 \right\} \tag{54}$$

on this range of T_g values. In short, on the range $T_g > T_{g*}$, tangency is only possible if the two constraints (46) and (54) are met.

Consider the first range $\beta \in (0, \alpha)$, and evaluate constraint (46). Note that $d_i > n_i$ on this range. This result can be verified by substituting (51) into (49) to find that $n_i = d_i$ when $\beta = \alpha$ and by further noting that smaller values of β increase d_i. Because $d_i > n_i$ for all i, the magnitude $\left\|\mathbf{z}_p\right\|$ is less than the magnitude $\left\|\mathbf{z}_c\right\|$ according to (48). Because $f(\mathbf{z}) = \|\mathbf{z}\|^2 / T$ by (2), $f(\mathbf{z}_p) < f(\mathbf{z}_c)$; furthermore, because $\mathbf{z}_c \in \partial F$ and because points $\mathbf{z}$ can only belong to ∂F if $f(\mathbf{z})=1$, $f(\mathbf{z}_p) \notin \partial F$, and hence $\mathbf{z}_p$ fails to satisfy requirement (46).

Next consider when $\beta = \alpha$, which presents a special case. Substituting (51) into (49), gives the result that $n_i = d_i$ when $\beta = \alpha$, and hence $\mathbf{z}_p = \mathbf{z}_c$. The constraints (46) and 54) are satisfied; however, $\mathbf{z}_p$ is clearly identical to the existing tangent point $\mathbf{z}_c$, and hence no new contact point appears.

Now consider the range $\beta \in (\alpha, \gamma]$. For all nonzero elements of $\mathbf{z}_c$, $n_i > d_i \geq 0$ on this range. The upper bound on d_i can be established by noting that increasing values of β monotonically decrease d_i according to (49) , and that all values of $d_i < n_i$ because $n_i = d_i$ when $\beta = \alpha$. The lower bound on d_i can be established by substituting (52) into (49) and noting that $d_i = 0$ when $\lambda_i = \lambda_{mnz}$, which occurs where $\beta = \gamma$. Because $n_i > d_i \geq 0$ for all nonzero elements of $\mathbf{z}_c$, the magnitude $\left\|\mathbf{z}_p\right\|$ is greater than the magnitude $\left\|\mathbf{z}_c\right\|$ according to (48), with magnitude tending toward infinity in the extreme case ($\beta = \gamma$). Appealing to (2), $f(\mathbf{z}_p) > f(\mathbf{z}_c)$ over the entire range $\beta \in (\alpha, \gamma]$. This result implies that $\mathbf{z}_p$ lies outside F and that $f(\mathbf{z}_p) > 1$, which violates constraint (46).

Finally, consider the range $\beta \in (\gamma, \infty)$. Since increasing β monotonically decreases d_i according to (49), and since $\beta = \gamma$ results in $d_i = 0$ for at least one nonzero element of $\mathbf{z}_p$, then $d_i < 0$ for at least one nonzero element of $\mathbf{z}_p$ on the range $\beta \in (\gamma, \infty)$, violating constraint (54) as long as $T_g > T_{g*}$.

Since at least one constraint is violated for every β, except the case $\beta = \alpha$ that corresponds to the existing tangent point $\mathbf{z}_c$, no new tangent point appears at any value of T_g incrementing downward from large T_g through all values with $T_g > T_{g*}$.

Theorem 5 (Conservative Bounding): *If T is finite and $T_g \geq T_{g*}$, then $G \supset F$.*

Background: Here all the above theorems are combined to prove G is a conservative bound for F.

✈

For any point on the sphere ∂F to leave the volume G as T_g shrinks, that point would need to exit G in the neighborhood of a tangent point. Theorem 2 states that G continues to bound F in the neighborhood of the specified tangent point $\mathbf{z}_c$, so long as $T_g \geq T_{g*}$. Theorem 4 states that no new tangent point appears on the range $T_g > T_{g*}$. By extension, were T_g to shrink infinitesimally such that $T_g = T_{g*}$, any new tangent point appearing would still belong to both the surfaces ∂F and ∂G, and so no point on ∂F would leave G. With no points on ∂F leaving G for $T_g \geq T_{g*}$, it can be said that $G \supset F$ over this range.

✈

Theorem 6 (Bounding Failure Near Contact Point) *If the component of $\mathbf{z}_c$ aligned with the smallest principal axis of Λ is zero, then G fails to contain ∂F in the neighborhood of $\mathbf{z}_c$ when $T_g < T_{g*}$.*

Background: To establish that the bound from the previous theorem is tight, it is sufficient to show that a failure happens at any smaller value of T_g, either because of a bounding failure at $\mathbf{z}_c$ (as shown in this proof) or at a new tangent point (as shown in the next proof).

Proof: In developing Theorem 2, condition (36) was necessary for G to contain points on ∂F in the neighborhood of $\mathbf{z}_c$. In the proof, a vector $\mathbf{z}_t$ was identified as orthogonal to $\mathbf{z}_c$. If the component of $\mathbf{z}_c$ in the direction of the smallest principal axis of Λ is zero, then one possible vector $\mathbf{z}_t$ is aligned with the smallest principal axis, and for this case, $\mathbf{z}_t^T \Lambda^{-1} \mathbf{z}_t / T = 1/\lambda_{\min}$. To ensure bounding for points in the neighborhood of $\mathbf{z}_c$, the right side of (36), equivalent to $\sqrt{T_g / T_{g*}} / \lambda_{\min}$, must therefore exceed $\lambda_{\min}$. This requirement is a contradiction for $T_g < T_{g*}$ and hence G fails to contain ∂F in the neighborhood of $\mathbf{z}_c$ when $T_g < T_{g*}$.

✈

Theorem 7 (Bounding Failure at New Tangent) *If the component of* $\mathbf{z}_c$ *aligned with the smallest principal axis of* Λ *is nonzero, then a new tangent point* $\overline{\mathbf{z}}_c \in (\partial F \cap \partial G) \setminus \mathbf{z}_c$ *appears when* T_g *shrinks to* $T_g = T_{g*}$. This point $\overline{\mathbf{z}}_c \notin G$ if $T_g < T_{g*}$.

Background: This theorem establishes that the second failure condition, where a new tangent point appears and then leaves *G* as T_g shrinks, occurs when $T_g < T_{g*}$.

Proof: First, define a vector function $\mathbf{r}(\mathbf{a}) \in \mathbb{R}^N$ that reflects the first element of a vector

$$\left[\mathbf{r}(\mathbf{a})\right]_i = \begin{cases} -[\mathbf{a}]_i & i = 1 \\ [\mathbf{a}]_i & i \neq 1 \end{cases} \tag{55}$$

Now consider the two boundaries ∂F and ∂G defined by (1) and (3), respectively. Assume without loss of generality, that the diagonal elements of Λ are sorted in ascending order by magnitude, such that the smallest diagonal element $\lambda_{\min}$ is also the first diagonal element ($\lambda_1 = \lambda_{\min}$). Modifying the general equation (6) for the ellipse center $\boldsymbol{\mu}_g$ using the relation $\lambda_1 = \lambda_{\min}$ and the assertion that $T_g = T_{g*}$, the following results.

$$\boldsymbol{\mu}_g = \left(\mathbf{I} - \boldsymbol{\Lambda}\frac{1}{\lambda_1}\right)\mathbf{z}_c . \tag{56}$$

By construction, ∂F and ∂G are tangent at $\mathbf{z}_c$. For $T_g = T_{g*}$, the two boundaries can also be shown to be tangent at a complementary point $\overline{\mathbf{z}}_c$:

$$\overline{\mathbf{z}}_c = \mathbf{r}(\mathbf{z}_c), \tag{57}$$

where the point $\overline{\mathbf{z}}_c$ is distinct from $\mathbf{z}_c$ because the component of $\mathbf{z}_c$ aligned with the smallest principal axis of Λ is nonzero (i.e., $[\mathbf{z}_c]_1 \neq 0$). To obtain this result, it is first necessary to show that the point $\overline{\mathbf{z}}_c$ lies on both boundaries. The point $\overline{\mathbf{z}}_c$ lies on ∂F because $\|\overline{\mathbf{z}}_c\| = \|\mathbf{z}_c\|$ and so, by (2), $f(\overline{\mathbf{z}}_c) = f(\mathbf{z}_c) = 1$. The point $\overline{\mathbf{z}}_c$ lies on ∂G since (57) guarantees that $\left|\left[\overline{\mathbf{z}}_c - \boldsymbol{\mu}_g\right]_i\right| = \left|\left[\mathbf{z}_c - \boldsymbol{\mu}_g\right]_i\right|$ for all components *i* including the first component, since $\left[\boldsymbol{\mu}_g\right]_1 = 0$ by (56). It follows from (4) that $g(\overline{\mathbf{z}}_c) = g(\mathbf{z}_c) = 1$ and hence that $\overline{\mathbf{z}}_c \in \partial G$.

To establish that ∂F and ∂G are tangent at $\overline{\mathbf{z}}_c$, it is also necessary to show that the normal vector fields for both boundaries, ∂F and ∂G, are parallel at $\overline{\mathbf{z}}_c$. The normal vector fields at $\overline{\mathbf{z}}_c$ are given by the following equations.

$$\nabla f(\overline{\mathbf{z}}_c) = \frac{2}{T}\overline{\mathbf{z}}_c \tag{58}$$

$$\nabla g(\overline{\mathbf{z}}_c) = \frac{2}{T_g}\boldsymbol{\Lambda}^{-1}\left(\overline{\mathbf{z}}_c - \boldsymbol{\mu}_g\right) \tag{59}$$

Combining (57) with (58) gives

$$\nabla f(\overline{\mathbf{z}}_c) = \mathbf{r}\left(\nabla f(\mathbf{z}_c)\right). \tag{60}$$

Noting that $\boldsymbol{\mu}_{g,1} = 0$ and that $\left[\overline{\mathbf{z}}_c\right]_i = \left[\mathbf{z}_c\right]_i$ for all i≠1, the formula (59) gives

$$\nabla g(\overline{\mathbf{z}}_c) = \mathbf{r}\left(\nabla g(\mathbf{z}_c)\right). \tag{61}$$

Now substitute (5) into (61).

$$\nabla g(\overline{\mathbf{z}}_c) = \frac{1}{\alpha}\mathbf{r}\left(\nabla f(\mathbf{z}_c)\right) \tag{62}$$

Comparing (60) and (62), it is evident that $\nabla g(\overline{\mathbf{z}}_c)$ is parallel to $\nabla f(\overline{\mathbf{z}}_c)$, since all vector components are identical after multiplying by a constant scalar value α.

Since $\overline{\mathbf{z}}_c$ lies on both ∂F and ∂G and since the normal vector fields at $\overline{\mathbf{z}}_c$ are parallel, $\overline{\mathbf{z}}_c$ is a new tangent point appearing when $T_g = T_{g*}$, assuming that $[\mathbf{z}_c]_1 \neq 0$.

If T_g shrinks incrementally, such that $T_g < T_{g*}$, then $\overline{\mathbf{z}}_c$ moves outside *G*. To see this, note that if $T_g < T_{g*}$, then $\left[\boldsymbol{\mu}_g\right]_1 = \psi[\mathbf{z}_c]_1$ where $\psi > 0$. Hence $\left|\left[\overline{\mathbf{z}}_c - \boldsymbol{\mu}_g\right]_i\right| > \left|\left[\mathbf{z}_c - \boldsymbol{\mu}_g\right]_i\right|$ and $g(\overline{\mathbf{z}}_c) > g(\mathbf{z}_c)$. Because $g(\mathbf{z}_c) \in \partial F$, $g(\mathbf{z}_c) = 1$ and $g(\overline{\mathbf{z}}_c) > 1$. Thus $\overline{\mathbf{z}}_c \notin G$ if $T_g < T_{g*}$.

✈

Theorem 8 (Tight Bounding) *Regardless of the location of the contact point* $\mathbf{z}_c$*, the smallest value of* T_g *for which* $G \supset F$ *is* $T_g = T_{g*}$.

Background: This theorem builds on the previous two to prove the conservative bound from Theorem 5 is a tight bound.

Proof: Theorem 6 identified that point on ∂F in the neighborhood of $\mathbf{z}_c$ move outside *G* if $T_g < T_{g*}$ and $[\mathbf{z}_c]_1 = 0$. Theorem 7 identified that a new tangent points appears and moves outside G if $T_g < T_{g*}$ and $[\mathbf{z}_c]_1 \neq 0$. These two cases cover all values $[\mathbf{z}_c]_1$. Hence, *G* does not contain *F* if $T_g < T_{g*}$. By contrast, Theorem 5 proved that $G \supset F$ for $T_g \geq T_{g*}$. Combining these results, the smallest value of T_g for which $G \supset F$ is $T_g = T_{g*}$.

✈

Theorem 9 (Bounding for Unknown Contact Point): *If T is finite and* $T_g = T_*$, *then* $G \supset F$ *regardless of the location of the contact point* $\mathbf{z}_c$.

Background: This theorem extends the previous theorem to cases when the contact point location $\mathbf{z}_c$ is unknown.

Proof: Because the definition of T_{g*}, from (9), depends on $\mathbf{z}_c$, it is only possible to use Theorem 8 (to ensure $G \supset F$ by setting $T_g = T_{g*}$) when the contact point $\mathbf{z}_c$ is known. The requirement that $\mathbf{z}_c$ be known can be relaxed by finding the maximum value of T_{g*} over all $\mathbf{z}_c$. Recalling $\left\|\mathbf{z}_c\right\| = \sqrt{T}$ and that $\lambda_{\max}$ is the largest element of the diagonal covariance matrix Λ,

$$\max_{\mathbf{z}_c}\left(\frac{\mathbf{z}_c^T \mathbf{\Lambda} \mathbf{z}_c}{\lambda_{\min}^2}\right) = T\,\frac{\lambda_{\max}}{\lambda_{\min}^2}\,. \tag{63}$$

The value on the right side of (63) has previously been defined as T_* in (10). If $T_g = T_*$ then $T_g \geq T_{g*} \forall \mathbf{z}_c$ and hence $G \supset F$ regardless of the location of the contact point $\mathbf{z}_c$

✈

References

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

[2] Pope, S., *Algorithms for Ellipsoids*, Report FDA-08-01, 2008.

[3] Phelts, R.E., Walter, T., and Enge, P., “Toward real-time SQM for WAAS: improved detection techniques,” *Proc. ION-GNSS*, 2003, pp. 2739-2749.

[4] Liu, F., Brenner, M., Tang, C.Y., “Signal deformation monitoring scheme implemented in a prototype local area augmentation system ground installation,” *Proc. ION-GNSS*, 2006, pp. 367-380.

[5] Rife, J., and Phelts, R.E., “Formulation of a time-varying maximum allowable error for Ground-Based Augmentation Systems,” *IEEE Trans. Aerospace and Electronic Systems*, Vol. 44, No. 2, 2008, pp. 548-560.

[6] Khanafseh, S., Pullen, S., Warburton, J., “Carrier phase ionospheric gradient ground monitor for GBAS with experimental validation,” *NAVIGATION*, Vol. 59, No. 1, 2012, pp. 51-60.

[7] Reuter, R., Weed, D., and Brenner, M., “Ionosphere gradient detection for Cat III GBAS,” *Proc. ION GNSS 2012*, Nashville, TN, 2012, pp. 2175-2183.

[8] Brown, R.G., “Receiver Autonomous Integrity Monitoring,” In *Global Positioning System: Theory and Applications, Volume II*, B.W. Parkinson and J.J. Spilker, Eds., Washington, D.C.: AIAA, 1996.

[9] Lee, Y., “Performance of Receiver Autonomous Integrity Monitoring (RAIM) in the Presence of Simultaneous Multiple Satellite Faults,” *Proc. ION Annual Meeting*, 2004.

[10] Joerger, M., Chan, F.-C., Langel, S., Pervan, B., “RAIM detector and estimator design to minimize the integrity risk," *Proc. ION GNSS 2012*, Nashville, TN, 2012, pp. 2785-2807.

[11] Osechas, O., Misra, P., and Rife, J., “Carrier-phase acceleration RAIM for GNSS satellite clock fault detection,” *NAVIGATION*, Vol. 59, No. 3, 2012, pp. 221-235.

[12] Rife, J., “Collaboration-enhanced receiver integrity monitoring (CERIM),” *Proc. IEEE Intelligent Transportation Systems (ITSC)*, pp.13-18, 2011.

[13] Rife, J., “Collaboration-enhanced receiver integrity monitoring with common residual estimation,” *Proc. IEEE/ION Position Location and Navigation Symposium (PLANS)*, pp.1042-1053, April 2012.

[14] Osechas, O. and Rife, J., “Distributed ionosphere monitoring by collaborating mobile receivers,” *IEEE Transactions on Aerospace & Electronic Systems*, accepted 2014.

[15] Rife, J.H., “Overbounding missed-detection probability for a chi-square monitor,” *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*, Nashville, TN, September 2012, pp. 1348-1368.

[16] Rife, J.H., “The effect of uncertain covariance on a chi square integrity monitor,” *NAVIGATION*, vol. 60, no. 4, 2013.

[17] Joerger, M. and Pervan, B. “Kalman filter-based integrity monitoring against sensor faults,” *Journal of Guidance, Control, and Dynamics*, Vol. 36, No. 2, 2013, pp. 349-361.