Generalized Attitude Determination with One Dominant Vector Observation

John L. Crassidis*
University at Buffalo, State University of New York, Amherst, New York, 14260-4400

Yang Cheng†
Mississippi State University, Mississippi State, MS 39762

I. Introduction

Attitude determination can be accomplished using a number of sensors, such as star trackers, magnetometers, horizon sensors, magnetometers, and arc-length measurements, e.g. using the Global Positioning System (GPS) [1]. Even though the sensing mechanism for each is different, most attitude observations can be expressed in vector form with the exception of GPS observations, which are arc-length types of observations. Vector-form observations can be used to determine the attitude by solving Wahba's problem [2]. A unique solution is given when at least two non-collinear vector observations exist. If the quaternion [3] is used for the attitude parametrization then the associated loss function in Wahba's problem is quadratic in nature, and the attitude orthogonalization constraint reduces down to a unit-norm vector constraint in the quaternion. Many solutions to this problem, such as Davenport's q Method [1] and the QUaternion ESTimator (QUEST) algorithm [4], can be used.

The loss function using only arc-length observations with the quaternion parametrization is quartic in nature. For the GPS case, if at least three non-planar baselines or sightlines exist then the GPS observations can be expressed in vector form [5]. A transformed loss function can be derived under these conditions, but it is still quartic in the quaternion because it is in fact equivalent to the non-transformed loss function. A suboptimal solution is presented in [5], which reduces the quartic loss function into a quadratic one so that any solution that solves Wahba's problem be employed. The error-attitude covariance is also derived, which can be compared with the maximum likelihood Cramér-Rao lower bound to see the efficiency of the suboptimal solution. The suboptimal solution can be used as a starting point for an iterative algorithm, such as a nonlinear least-squares (NLS) approach, to determine the optimal solution.

An approximate solution using only vector observations with one dominant observation is derived in [6]. The algorithm is very computationally efficient, and does not require iterative calculations or transcendental functions. This Note expands upon the work shown in [6] by deriving an approximate attitude solution involving an accurate line-of-sight (LOS) vector observation and other types of attitude observations, which may include other LOS or arc-length observations. These types of observations represent the most general type used for most attitude determination

*Samuel P. Capen Chair Professor, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Fellow AIAA.
†Associate Professor, Department of Aerospace Engineering. Email: cheng@ae.msstate.edu. Associate Fellow AIAA.

systems. The attitude parametrization chosen in this Note is the quaternion. The suboptimal covariance is derived, as well as a simple scalar expression that can be computed without the attitude solution to check the accuracy of the approximate attitude solution.

II. Algorithm Development

The generalized attitude determination problem minimizes the following loss function:

$$J(A) = \left[\frac{1}{2}\sum_{k=1}^{n}\sigma_k^{-2}\|\mathbf{b}_k - A\mathbf{r}_k\|^2\right] + \left[\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}(\phi_{ij} - \mathbf{c}_i^T A\mathbf{s}_j)^2\right] \tag{1}$$

where A is the attitude matrix, which is a 3×3 proper orthogonal matrix, $\mathbf{b}_k$ is the LOS unit vector expressed in body-frame coordinates, $\mathbf{r}_k$ is the LOS unit vector expressed in reference-frame coordinates, σ_k is the standard deviation of the unit vector observation errors [7], and σ_{ij} is the standard deviation of the arc-length observations. These arc-length observations, denoted by ϕ_{ij}, may come from GPS attitude sensors [1] or other sensors [8], where $\mathbf{c}_i$ is a vector expressed in body-frame coordinates, and $\mathbf{s}_j$ is a vector expressed in reference-frame coordinates. For GPS observations, $\mathbf{c}_i$ is the ith baseline vector, and $\mathbf{s}_j$ is the jth sightline vector. To date a non-iterative solution for the optimal attitude that minimizes Eq. (1) is not available, but the optimal attitude can be found using an iterative NLS solution.

A closed-form approximate solution for the attitude is now derived. Suppose that an accurate LOS observation is provided by $\mathbf{r}_1$ and $\mathbf{b}_1$, whose associated standard deviation σ_1 is much smaller than the other standard deviations. The above optimal attitude determination problem can be approximated by a suboptimal problem that minimizes the loss function

$$J(A) = \left[\frac{1}{2}\sum_{k=2}^{n}\sigma_k^{-2}\|\mathbf{b}_k - A\mathbf{r}_k\|^2\right] + \left[\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}(\phi_{ij} - \mathbf{c}_i^T A\mathbf{s}_j)^2\right] \tag{2}$$

subject to the constraint

$$\mathbf{b}_1 = A\mathbf{r}_1 \tag{3}$$

Note that Eq. (3) contains no errors, which means it is assumed to be a noise-free measurement.

The suboptimal problem can be solved by a computationally efficient algorithm, which is now developed. The attitude is subsequently parameterized by the quaternion $\mathbf{q}$. The quaternion is a four-dimensional vector, defined as

$$\mathbf{q} \equiv \begin{bmatrix} \mathbf{q}_{1:3} \\ q_4 \end{bmatrix} \tag{4}$$

with

$$\mathbf{q}_{1:3} \equiv [q_1 \;\; q_2 \;\; q_3]^T = \mathbf{e}\sin(\vartheta/2) \tag{5a}$$

$$q_4 = \cos(\vartheta/2) \tag{5b}$$

where $\mathbf{e}$ is the unit Euler axis and ϑ is the rotation angle [3]. A quaternion parameterizing an attitude satisfies a single constraint given by $\|\mathbf{q}\| = 1$. In terms of the quaternion, its associated attitude matrix is given by

$$A(\mathbf{q}) = \Xi^T(\mathbf{q})\Psi(\mathbf{q}) \tag{6}$$

with

$$\Xi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_3 + [\mathbf{q}_{1:3}\times] \\ -\mathbf{q}_{1:3}^T \end{bmatrix}, \quad \Psi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_3 - [\mathbf{q}_{1:3}\times] \\ -\mathbf{q}_{1:3}^T \end{bmatrix} \tag{7a}$$

$$[\mathbf{q}_{1:3}\times] \equiv \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \tag{7b}$$

where I_3 is a 3×3 identity matrix and the matrix $[\mathbf{q}_{1:3}\times]$ is the standard cross-product matrix.

Equation (2) can be rewritten in terms of the quaternion as

$$J(\mathbf{q}) = \mathbf{q}^T (K + G)\,\mathbf{q} + \frac{\alpha}{2} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} \sigma_{ij}^{-2}\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2 \tag{8}$$

where

$$K \equiv -\sum_{k=2}^{n} \sigma_k^{-2}\Omega(\mathbf{b}_k)\Gamma(\mathbf{r}_k) \tag{9a}$$

$$G \equiv \sum_{i=1}^{N}\sum_{j=1}^{M} \sigma_{ij}^{-2}\phi_{ij}C_{ij} \tag{9b}$$

$$\alpha \equiv \left[2\sum_{k=1}^{n}\sigma_k^{-2}\right] + \left[\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\phi_{ij}^2\right] \tag{9c}$$

$$C_{ij} \equiv \Omega(\mathbf{c}_i)\Gamma(\mathbf{s}_j) \tag{9d}$$

$$\Omega(\mathbf{d}) \equiv \begin{bmatrix} -[\mathbf{d}\times] & \mathbf{d} \\ -\mathbf{d}^T & 0 \end{bmatrix}, \quad \Gamma(\mathbf{d}) \equiv \begin{bmatrix} [\mathbf{d}\times] & \mathbf{d} \\ -\mathbf{d}^T & 0 \end{bmatrix} \tag{9e}$$

where $\mathbf{d}$ is any 3×1 vector. Unless $\mathbf{b}_1 = -\mathbf{r}_1$, the most general unit quaternion satisfying Eq. (3) is given by

$$\mathbf{q} = \cos(\psi/2)\,\mathbf{q}_{\min} + \sin(\psi/2)\,\mathbf{q}_{180} \tag{10}$$

with

$$\mathbf{q}_{\min} = \frac{1}{\sqrt{2(1+\mathbf{b}_1^T\mathbf{r}_1)}} \begin{bmatrix} \mathbf{b}_1 \times \mathbf{r}_1 \\ 1 + \mathbf{b}_1^T\mathbf{r}_1 \end{bmatrix} \tag{11a}$$

$$\mathbf{q}_{180} = \frac{1}{\sqrt{2(1+\mathbf{b}_1^T\mathbf{r}_1)}} \begin{bmatrix} \mathbf{b}_1 + \mathbf{r}_1 \\ 0 \end{bmatrix} \tag{11b}$$

where ψ is an arbitrary parameter. When $\mathbf{b}_1 = -\mathbf{r}_1$, $\mathbf{q}_{\min}$ and $\mathbf{q}_{180}$ are indeterminate, but this condition can be avoided by solving for the attitude with respect to a reference coordinate frame related to the original reference frame by a 180 degree rotation about one of the coordinate axes [1, 4]. Note that, in the generalized attitude determination problem, both the representations of $\mathbf{r}_k$ and those of $\mathbf{s}_j$ are transformed by reference-frame rotations.

Substituting Eq. (10) into $\mathbf{q}^T(K+G)\mathbf{q}$ leads to

$$\mathbf{q}^T(K+G)\,\mathbf{q} = (\mu\cos\psi + \nu\sin\psi + \kappa)/2 \tag{12}$$

where

$$\mu \equiv \mathbf{q}_{\min}^T(K+G)\,\mathbf{q}_{\min} - \mathbf{q}_{180}^T(K+G)\,\mathbf{q}_{180} \tag{13a}$$

$$\nu \equiv 2\mathbf{q}_{\min}^T(K+G)\,\mathbf{q}_{180} \tag{13b}$$

$$\kappa \equiv \mathbf{q}_{\min}^T(K+G)\,\mathbf{q}_{\min} + \mathbf{q}_{180}^T(K+G)\,\mathbf{q}_{180} \tag{13c}$$

Substituting Eq. (10) into $\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2$ gives

$$\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2 = \left(\bar{\mu}_{ij}\cos\psi + \bar{\nu}_{ij}\cos\psi + \bar{\kappa}_{ij}\right)^2/4 \tag{14}$$

where

$$\bar{\mu}_{ij} \equiv \mathbf{q}_{\min}^T C_{ij} \mathbf{q}_{\min} - \mathbf{q}_{180}^T C_{ij} \mathbf{q}_{180} \tag{15a}$$

$$\bar{\nu}_{ij} \equiv 2\mathbf{q}_{\min}^T C_{ij} \mathbf{q}_{180} \tag{15b}$$

$$\bar{\kappa}_{ij} \equiv \mathbf{q}_{\min}^T C_{ij} \mathbf{q}_{\min} + \mathbf{q}_{180}^T C_{ij} \mathbf{q}_{180} \tag{15c}$$

Performing the multiplications in Eq. (14) yields

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2 = \frac{1}{2}\left(\gamma_1\cos^2\psi + \gamma_2\sin^2\psi + \gamma_3\sin\psi\cos\psi + \gamma_4\cos\psi + \gamma_5\sin\psi + \gamma_6\right) \tag{16}$$

where

$$\gamma_1 \equiv \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\mu}_{ij}^2 \tag{17a}$$

$$\gamma_2 \equiv \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\nu}_{ij}^2 \tag{17b}$$

$$\gamma_3 \equiv \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\mu}_{ij}\bar{\nu}_{ij} \tag{17c}$$

$$\gamma_4 \equiv \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\mu}_{ij}\bar{\kappa}_{ij} \tag{17d}$$

$$\gamma_5 \equiv \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\nu}_{ij}\bar{\kappa}_{ij} \tag{17e}$$

$$\gamma_6 \equiv \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\bar{\kappa}_{ij}^2 \tag{17f}$$

Substituting Eqs. (12) and (16) into Eq. (8) gives

$$\begin{aligned} J(\psi) = &\frac{1}{2}\left[\gamma_1\cos^2\psi + \gamma_2\sin^2\psi + \gamma_3\sin\psi\cos\psi + (\mu+\gamma_4)\cos\psi + (\nu+\gamma_5)\sin\psi\right] \\ &+ (\kappa+\alpha+\gamma_6)/2 \end{aligned} \tag{18}$$

The necessary condition to minimize $J(\psi)$ is given by

$$-(\mu+\gamma_4)\sin\psi + (\nu+\gamma_5)\cos\psi - (\gamma_1-\gamma_2)\sin(2\psi) + \gamma_3\cos(2\psi) = 0 \tag{19}$$

where the identities $\sin(2\psi) = 2\sin\psi\cos\psi$ and $\cos(2\psi) = \cos^2\psi - \sin^2\psi$ have been used. Note that terms involving $\sin(2\psi)$ and $\cos(2\psi)$ now appear that are not in the formulation shown in [6]. If only LOS observations exist then $G = 0$, and all the γ terms are also zero. In this case, the solution reduces to the solution given by [6]. For the general case multiple roots may exist now because of the extra terms.

An iterative approach can be used to find all the roots. But, here a different approach is derived that does not involve an iterative solution. Define $x \equiv \sin\psi$, which gives $\cos\psi = \pm\sqrt{1-x^2}$. Using these definitions, Eq. (19) can be written as

$$\gamma_3(1-2x^2) - (\mu+\gamma_4)x = \pm\sqrt{1-x^2}\,[2(\gamma_1-\gamma_2)x - (\nu+\gamma_5)] \tag{20}$$

Squaring both sides of Eq. (20), and collecting terms gives

$$\delta_4 x^4 + \delta_3 x^3 + \delta_2 x^2 + \delta_1 x + \delta_0 = 0 \tag{21}$$

where

$$\delta_0 \equiv \gamma_3^2 - (\nu+\gamma_5)^2 \tag{22a}$$

$$\delta_1 \equiv 2\,[2(\gamma_1-\gamma_2)(\nu+\gamma_5) - \gamma_3(\mu+\gamma_4)] \tag{22b}$$

$$\delta_2 \equiv (\mu+\gamma_4)^2 + (\nu+\gamma_5)^2 - 4\left[\gamma_3^2 + (\gamma_1-\gamma_2)^2\right] \tag{22c}$$

$$\delta_3 \equiv 4\,[\gamma_3(\mu+\gamma_4) - (\gamma_1-\gamma_2)(\nu+\gamma_5)] \tag{22d}$$

$$\delta_4 \equiv 4\left[\gamma_3^2 + (\gamma_1-\gamma_2)^2\right] \tag{22e}$$

There are two possibilities for x. One is that they are all real, and the other is that two are real and the other two are complex conjugates. Descartes' Rule of Signs can be used to determine the number of real roots. Once these roots have been determined, $\sin\psi$, $\cos\psi$, and ψ are given by

$$\sin\psi = x \tag{23a}$$

$$\cos\psi = \text{sign}\left[\gamma_3(1-2x^2) - (\mu+\gamma_4)x\right]\,\text{sign}\,[2(\gamma_1-\gamma_2)x - (\nu+\gamma_5)]\,\sqrt{1-x^2} \tag{23b}$$

$$\psi = \text{ATAN2}(\sin\psi, \cos\psi) \tag{23c}$$

where sign$(\cdot)$ returns the sign of the argument, and ATAN2 is the four-quadrant inverse tangent function. The loss function given by Eq. (18) is evaluated to find the minimizing ψ. Note that ψ itself is not actually required to compute the loss function because Eqs. (23a) and (23b) can directly be used in Eq. (18). Finally, the minimizing quaternion is calculated using Eq. (10), and using the identities $\cos(\psi/2) = \pm\sqrt{(1+\cos\psi)/2}$ and $\sin(\psi/2) = \pm\sqrt{(1-\cos\psi)/2}$. If $\psi \in [-\pi, \pi]$, then $\psi/2 \in [-\pi/2, \pi/2]$, and $\cos(\psi/2)$ and $\sin(\psi/2)$ have the same sign when $\sin\psi > 0$ and the opposite sign when $\sin\psi < 0$. Then, the minimizing quaternion is given by

$$\mathbf{q} = \sqrt{\frac{1+\cos\psi}{2}}\,\mathbf{q}_{\text{min}} + \text{sign}(\sin\psi)\sqrt{\frac{1-\cos\psi}{2}}\,\mathbf{q}_{180} \tag{24}$$

The minimizing quaternion can also be determined without computing transcendental functions by using Eqs. (23a) and (23b) in Eq. (24). The algorithm (without reference frame rotations) is summarized in Algorithm 1.

Algorithm 1 Generalized Quaternion Estimation with One Dominant Vector

```
Compute matrices K, G, and C_ij and scalar α                    ▷ Eqs. (9)
Compute μ, ν, κ                                                  ▷ Eqs. (13c)
Compute γ_1 through γ_6                                          ▷ Eqs. (15) and (17)
Compute δ_0 through δ_4                                          ▷ Eqs. (22)
Solve for real x (= sin ψ)                                       ▷ Eq. (21)
Compute all cos ψ using real values of x                         ▷ Eq. (23b)
Select the sin ψ and cos ψ that minimize the loss function       ▷ Eq. (18)
Construct q_min and q_180                                        ▷ Eq. (11)
Construct the quaternion estimate q                              ▷ Eq. (24)
```

Now special cases of the coefficients of the quartic equation are discussed. Consider the case when three orthonormal $\mathbf{c}_i$'s exist with $\sigma_{cj}^{-2} \equiv \sigma_{1j}^{-2} = \sigma_{2j}^{-2} = \sigma_{3j}^{-2}$. Then, the double summation on the right side of Eq. (8) is given by

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2 &= \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{M}\sigma_{cj}^{-2}\mathbf{s}_j^T A^T \mathbf{c}_i \mathbf{c}_i^T A\mathbf{s}_j \\ &= \frac{1}{2}\sum_{j=1}^{M}\sigma_{cj}^{-2}\mathbf{s}_j^T A^T A\mathbf{s}_j \\ &= \frac{1}{2}\sum_{j=1}^{M}\sigma_{cj}^{-2}\end{aligned} \tag{25}$$

Thus, this term is now independent of the attitude matrix. Therefore, Eq. (8) is equivalent to solving Wahba's problem [5]. The same is true when three orthonormal $\mathbf{s}_j$'s exist with $\sigma_{is}^{-2} \equiv \sigma_{i1}^{-2} = \sigma_{i2}^{-2} = \sigma_{i3}^{-2}$, which leads to

$$\begin{aligned}\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\left(\mathbf{q}^T C_{ij}\mathbf{q}\right)^2 &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{3}\sigma_{is}^{-2}\mathbf{c}_i^T A\mathbf{s}_j \mathbf{s}_j^T A^T \mathbf{c}_i \\ &= \frac{1}{2}\sum_{i=1}^{N}\sigma_{is}^{-2}\mathbf{c}_i^T A A^T \mathbf{c}_i \\ &= \frac{1}{2}\sum_{i=1}^{N}\sigma_{is}^{-2}\end{aligned} \tag{26}$$

This again is independent of the attitude matrix. In both cases, since the quartic terms in the loss function given by Eq. (8) vanish, then $\bar{\mu}_{ij} = \bar{\nu}_{ij} = \bar{\kappa}_{ij} = 0$ and thus $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_3 = \gamma_5 = 0$. The quartic equation reduces to a quadratic equation $(\mu^2 + \nu^2)x^2 - \nu^2 = 0$.

The attitude solution is nonunique if and only if the loss function in ψ is independent of ψ. The condition is

$$\gamma_1 = \gamma_2, \quad \gamma_3 = \mu + \gamma_4 = \nu + \gamma_5 = 0 \tag{27}$$

In that case, the loss function becomes $J(\psi) = (\gamma_1 + \kappa + \alpha + \gamma_6)/2$ and the coefficients of the quartic equation become $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 0$.

III. Covariance Analysis

This section derives the covariance of the attitude errors for the suboptimal algorithm. A simple expression is also derived that can be used to check the accuracy of the suboptimal algorithm as compared to the optimal solution derived from a maximum likelihood analysis. Two methods are used to derive the relationship between the attitude error and the measurement noise. The first is based on a perturbation approach, while the second is based on a constrained least-squares approach. Both will be shown to yield identical results.

Under the small-error assumption, the attitude matrix is approximated by

$$A = (I_3 - [\boldsymbol{\delta\vartheta}\times])A^{\text{true}} \tag{28}$$

where $\boldsymbol{\delta\vartheta}$ is the vector of the attitude errors. The error-quaternion is approximated by $\boldsymbol{\delta}\mathbf{q} = [\frac{1}{2}\boldsymbol{\delta\vartheta}^T \;\; 1]^T$. The covariance matrix associated with the attitude estimate is

$$P_{\text{sub}} = E\{\boldsymbol{\delta\vartheta}\,\boldsymbol{\delta\vartheta}^T\} \tag{29}$$

Let $\mathbf{b}_k^{\text{true}} = A^{\text{true}}\mathbf{r}_k$, which is the true body-vector, and let ϕ_{ij}^{true} denote the true arc-length. The models for the body and arc-length observations are respectively given by

$$\mathbf{b}_k = \mathbf{b}_k^{\text{true}} + \boldsymbol{\Delta}\mathbf{b}_k \tag{30a}$$

$$\phi_{ij} = \phi_{ij}^{\text{true}} + \Delta\phi_{ij} \tag{30b}$$

where $\boldsymbol{\Delta}\mathbf{b}_k$ is the body-vector error whose covariance is given by the QUEST Measurement Model (QMM) [7], and $\Delta\phi_{ij}$ is the arc-length error whose variance is given by σ_{ij}^2.

The optimal covariance, denoted by P_{opt}, is given by the inverse of the combination of the Fisher information matrix (FIM) associated with the LOS observation plus the FIM associated with other observations:

$$P_{\text{opt}} = F^{-1} \tag{31a}$$

$$
\begin{aligned}
F \equiv \sigma_1^{-2}[I_3 - \mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T] + \sum_{k=2}^{n} \sigma_k^{-2}[I_3 - \mathbf{b}_k^{\text{true}}(\mathbf{b}_k^{\text{true}})^T] \\
+ \sum_{i=1}^{N}\sum_{j=1}^{M} \sigma_{ij}^{-2}[A^{\text{true}}\mathbf{s}_j\times]\mathbf{c}_i\mathbf{c}_i^T[A^{\text{true}}\mathbf{s}_j\times]^T
\end{aligned}
\tag{31b}
$$

where σ_1 is the standard deviation associated with the error in $\mathbf{b}_1$, and F is the FIM. Also, define

$$
\bar{F} \equiv F - \sigma_1^{-2}[I_3 - \mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T] \tag{32a}
$$

$$
\sigma_{\text{eff}}^{-2} \equiv (\mathbf{b}_1^{\text{true}})^T \bar{F}\, \mathbf{b}_1^{\text{true}} \tag{32b}
$$

where $\bar{F}$ is the portion of the FIM associated with the remaining observations outside of $\mathbf{b}_1^{\text{true}}$, and σ_{eff}^2 is an effective error-variance.

A. Perturbation Approach

For this approach, the error-quaternion is given by

$$
\boldsymbol{\delta}\mathbf{q} = \begin{bmatrix} \frac{\boldsymbol{\delta\vartheta}}{2} \\ 1 \end{bmatrix} \approx \boldsymbol{\delta}\mathbf{q}_{\min} + \frac{\hat{\psi}}{2}\boldsymbol{\delta}\mathbf{q}_{180} \tag{33}
$$

where

$$
\boldsymbol{\delta}\mathbf{q}_{\min} = \frac{1}{\sqrt{4 + 2(\mathbf{b}_1^{\text{true}})^T\boldsymbol{\Delta}\mathbf{b}_1}} \begin{bmatrix} -\mathbf{b}_1^{\text{true}} \times \boldsymbol{\Delta}\mathbf{b}_1 \\ 2 + (\mathbf{b}_1^{\text{true}})^T\boldsymbol{\Delta}\mathbf{b}_1 \end{bmatrix} \tag{34a}
$$

$$
\boldsymbol{\delta}\mathbf{q}_{180} = \frac{1}{\sqrt{4 + 2(\mathbf{b}_1^{\text{true}})^T\boldsymbol{\Delta}\mathbf{b}_1}} \begin{bmatrix} 2\mathbf{b}_1^{\text{true}} + \boldsymbol{\Delta}\mathbf{b}_1 \\ 0 \end{bmatrix} \tag{34b}
$$

Also, the notation $\hat{\psi}$ is used to denote ψ with error due to measurement noise. So, $\boldsymbol{\delta\vartheta} = \hat{\psi}\,\mathbf{b}_1^{\text{true}} - \mathbf{b}_1^{\text{true}} \times \boldsymbol{\Delta}\mathbf{b}_1$.

Now an expression for $\hat{\psi}$ in terms of the noise is derived. The necessary condition that $\hat{\psi}$ satisfies is approximated by

$$
-(\mu + \gamma_4)\,\hat{\psi} + (\nu + \gamma_5) - (\gamma_1 - \gamma_2)\,(2\hat{\psi}) + \gamma_3 = 0 \tag{35}
$$

Then $\hat{\psi}$ is given by

$$
\hat{\psi} = \frac{\nu + \gamma_3 + \gamma_5}{\mu + 2\gamma_1 + 2\gamma_2 + \gamma_4} \tag{36}
$$

After significant algebra it can be shown that

$$\hat{\psi} = \frac{1}{d}\left[\left(\sum_{k=2}^{n}\mathbf{k}_k^T\mathbf{\Delta b}_k\right) + \left(\sum_{i=1}^{N}\sum_{j=1}^{M}k_{ij}\Delta\phi_{ij}\right) + \mathbf{k}_1^T\mathbf{\Delta b}_1\right] \tag{37}$$

where

$$\mathbf{k}_k \equiv \sigma_k^{-2}\mathbf{b}_k^{\text{true}} \times \mathbf{b}_1^{\text{true}} \tag{38a}$$

$$k_{ij} \equiv \sigma_{ij}^{-2}\mathbf{c}_i^T[A^{\text{true}}\mathbf{s}_j\times]\mathbf{b}_1^{\text{true}} \tag{38b}$$

$$\mathbf{k}_1 \equiv \left\{\sum_{k=2}^{n}\sigma_k^{-2}[\mathbf{b}_1^{\text{true}}\times][\mathbf{b}_k^{\text{true}}\times]^2\mathbf{b}_1^{\text{true}}\right\} + \left\{\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\left(\mathbf{c}_i^T[A^{\text{true}}\mathbf{s}_j\times]\mathbf{b}_1^{\text{true}}\right)[\mathbf{b}_1^{\text{true}}\times][A^{\text{true}}\mathbf{s}_j\times]\mathbf{c}_i\right\} \tag{38c}$$

$$d \equiv \left\{-\sum_{k=2}^{n}\sigma_k^{-2}(\mathbf{b}_1^{\text{true}})^T[\mathbf{b}_k^{\text{true}}\times]^2\mathbf{b}_1^{\text{true}}\right\} + \left\{\sum_{i=1}^{N}\sum_{j=1}^{M}\sigma_{ij}^{-2}\left(\mathbf{c}_i^T[A^{\text{true}}\mathbf{s}_j\times]\mathbf{b}_1^{\text{true}}\right)^2\right\} \tag{38d}$$

Substituting Eq (37) into the $\boldsymbol{\delta\vartheta}$ expression gives

$$\boldsymbol{\delta\vartheta} = \left(\sum_{k=2}^{n}\frac{\mathbf{b}_1^{\text{true}}\mathbf{k}_k^T}{d}\mathbf{\Delta b}_k\right) + \left(\sum_{i=1}^{N}\sum_{j=1}^{M}\frac{k_{ij}\mathbf{b}_1^{\text{true}}}{d}\Delta\phi_{ij}\right) + \left(-[\mathbf{b}_1^{\text{true}}\times] + \frac{\mathbf{b}_1^{\text{true}}\mathbf{k}_1^T}{d}\right)\mathbf{\Delta b}_1 \tag{39}$$

Thus, the covariance is given by $P_{\text{sub}} = P_v + P_a + P_1$, where

$$P_v \equiv \sum_{k=2}^{n}\frac{\sigma_k^2\mathbf{k}_k^T\mathbf{k}_k}{d^2}\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T \tag{40a}$$

$$P_a \equiv \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{\sigma_{ij}^2k_{ij}^2}{d^2}\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T \tag{40b}$$

$$P_1 \equiv \sigma_1^2\left(-[\mathbf{b}_1^{\text{true}}\times] + \frac{\mathbf{b}_1^{\text{true}}\mathbf{k}_1^T}{d}\right)\left(-[\mathbf{b}_1^{\text{true}}\times] + \frac{\mathbf{b}_1^{\text{true}}\mathbf{k}_1^T}{d}\right)^T \tag{40c}$$

Using the following identities:

$$d = (\mathbf{b}_1^{\text{true}})^T\bar{F}\,\mathbf{b}_1^{\text{true}} = \sigma_{\text{eff}}^{-2} \tag{41a}$$

$$P_v + P_g = \sigma_{\text{eff}}^2\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T \tag{41b}$$

$$\left(-[\mathbf{b}_1^{\text{true}}\times] + \frac{\mathbf{b}_1^{\text{true}}\mathbf{k}_1^T}{d}\right)[\mathbf{b}_1^{\text{true}}\times] = I_3 - \sigma_{\text{eff}}^2\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\bar{F} \tag{41c}$$

then gives

$$P_{\text{sub}} = \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T + \sigma_1^2\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\bar{F}\right]\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\bar{F}\right]^T \tag{42}$$

It is important to note that Eq. (42) is a general expression for any $\bar{F}$. Some special cases for $\bar{F}$ will be shown later. Also, the matrix $\left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right]$ is a projection matrix onto the plane perpendicular to $\mathbf{b}_1^{\text{true}}$, so $\left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \mathbf{b}_1^{\text{true}} = \mathbf{0}$. In addition, substituting Eq. (32a) into Eq. (42) shows that P_{sub} remains unchanged using F in place of $\bar{F}$.

B. Constrained Least Squares Approach

The following measurement model is assumed for the constrained least-squares problem [9]:

$$\mathbf{y}_1 = H_1 \mathbf{x}^{\text{true}} + \mathbf{\Delta y}_1 \tag{43a}$$

$$\mathbf{y}_2 = H_2 \mathbf{x}^{\text{true}} \tag{43b}$$

where $\mathbf{\Delta y}_1$ is the zero-mean measurement noise with the covariance matrix being the identity matrix. The optimal estimate $\mathbf{x}$ minimizes the loss function

$$J(\mathbf{x}) = \frac{1}{2} \|\mathbf{y}_1 - H_1 \mathbf{x}\|^2 \tag{44}$$

subject to

$$\mathbf{y}_2 = H_2 \mathbf{x}$$

The optimality condition is given by

$$\begin{bmatrix} H_1^T H_1 & H_2^T \\ H_2 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} H_1^T \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \tag{45}$$

where $\boldsymbol{\lambda}$ is the vector of Lagrange multipliers.

For this attitude determination problem, the linearized measurement model is

$$\frac{\mathbf{\Delta b}_k}{\sigma_k} = \frac{[\mathbf{b}_k^{\text{true}} \times]}{\sigma_k} \boldsymbol{\delta\vartheta} \tag{46a}$$

$$\frac{\Delta \phi_{ij}}{\sigma_{ij}} = \frac{\mathbf{c}_i^T [A^{\text{true}} \mathbf{s}_j \times]}{\sigma_{ij}} \boldsymbol{\delta\vartheta} \tag{46b}$$

To apply a constrained least-square estimation algorithm, define

$$\mathbf{x} \equiv \boldsymbol{\delta\vartheta} \tag{47a}$$

$$\mathbf{y}_1 \equiv \left[\frac{\mathbf{\Delta b}_2^T}{\sigma_2}, \ldots, \frac{\mathbf{\Delta b}_n^T}{\sigma_n}, \frac{\Delta \phi_{11}}{\sigma_{11}}, \ldots, \frac{\Delta \phi_{NM}}{\sigma_{NM}} \right]^T \tag{47b}$$

$$H_1 \equiv \left[\frac{[\mathbf{b}_2^{\text{true}} \times]^T}{\sigma_2}, \ldots, \frac{[\mathbf{b}_n^{\text{true}} \times]^T}{\sigma_n}, \frac{[A^{\text{true}} \mathbf{s}_1 \times]^T \mathbf{c}_1}{\sigma_{11}}, \ldots, \frac{[A^{\text{true}} \mathbf{s}_M \times]^T \mathbf{c}_N}{\sigma_{NM}} \right]^T \tag{47c}$$

$$\mathbf{y}_2 \equiv \mathcal{M}\frac{\boldsymbol{\Delta}\mathbf{b}_1}{\sigma_1} \tag{47d}$$

$$H_2 \equiv \mathcal{M} = [-\mathbf{h}_2 \ \ \mathbf{h}_1]^T \tag{47e}$$

where $\mathbf{h}_1$ is any vector perpendicular to $\mathbf{b}_1^{\text{true}}$, and $\mathbf{h}_2 = \mathbf{b}_1^{\text{true}} \times \mathbf{h}_1$. This reduced form for H_2 is chosen because the full form leads to a non-invertible matrix. It still provides the correct properties for the constraint because it can be verified that

$$\mathcal{M}\mathbf{b}_1^{\text{true}} = \mathbf{0} \tag{48a}$$

$$\mathcal{M}\mathcal{M}^T = I_2 \tag{48b}$$

$$\mathcal{M}^T\mathcal{M} = I_3 - \mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T \tag{48c}$$

and $H_1^T H_1 = \bar{F}$. The noise vector satisfies

$$E\left\{\begin{bmatrix} H_1^T\mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}\right\} = \mathbf{0}, \quad E\left\{\begin{bmatrix} H_1^T\mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}\begin{bmatrix} H_1^T\mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}^T\right\} = \begin{bmatrix} \bar{F} & 0_{3\times 2} \\ 0_{2\times 3} & \sigma_1^2 I_2 \end{bmatrix} \tag{49}$$

The optimal estimates are

$$\begin{bmatrix} \boldsymbol{\delta\vartheta} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \bar{F} & \mathcal{M}^T \\ \mathcal{M} & 0_{2\times 2} \end{bmatrix}^{-1}\begin{bmatrix} H_1^T\mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \tag{50}$$

Note that $\bar{F}$ may be singular. Let

$$\begin{bmatrix} \bar{F} & \mathcal{M}^T \\ \mathcal{M} & 0_{2\times 2} \end{bmatrix}^{-1} = \begin{bmatrix} \mathcal{B} & C^T \\ C & \mathcal{D} \end{bmatrix} \tag{51}$$

where the following quantities can be easily derived:

$$\mathcal{B} = \frac{\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T}{(\mathbf{b}_1^{\text{true}})^T\bar{F}\,\mathbf{b}_1^{\text{true}}} = \sigma_{\text{eff}}^2\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T \tag{52a}$$

$$C = \mathcal{M} - \mathcal{M}\bar{F}\mathcal{B} \tag{52b}$$

$$\mathcal{D} = -\mathcal{M}\bar{F}C^T \tag{52c}$$

Note that Eq. (52a) follows from $\mathcal{M}\mathcal{B} = 0_{2\times3}$. The error is given by

$$\boldsymbol{\delta\vartheta} = \begin{bmatrix} \mathcal{B} & C^T \end{bmatrix} \begin{bmatrix} H_1^T \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \tag{53}$$

which is equivalent to Eq. (39). Then the covariance is

$$P_{\text{sub}} = \begin{bmatrix} \mathcal{B} & C^T \end{bmatrix} \begin{bmatrix} \bar{F} & 0 \\ 0 & \sigma_1^2 I_2 \end{bmatrix} \begin{bmatrix} \mathcal{B} & C^T \end{bmatrix}^T = \mathcal{B}\bar{F}\mathcal{B}^T + \sigma_1^2 C C^T \tag{54}$$

It can be shown that

$$\mathcal{B}\bar{F}\mathcal{B}^T = \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \tag{55a}$$

$$C C^T = \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right]^T \tag{55b}$$

Hence, substituting the $\mathcal{B}$ and C expressions from Eq. (52) into Eq. (54) gives back exactly the same suboptimal covariance in Eq. (42).

C. Optimality Condition

Assume P_{sub} and F are nonsingular and $\sigma_{\text{eff}}^{-2} > 0$. A condition required to make $P_{\text{sub}}\, F$ close to the identity matrix is now derived. Carrying out the multiplication gives

$$\begin{aligned} P_{\text{sub}}\, F &= \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T F + \sigma_1^2 F \\ &\quad - \sigma_{\text{eff}}^2 \sigma_1^2 \left[\mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F} F + \bar{F}\, \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T F - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T F\right] \end{aligned} \tag{56}$$

Solving Eq. (32a) for F, and substituting the resulting expression into Eq. (56) leads to

$$\begin{aligned} P_{\text{sub}}\, F &= \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F} - \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F} \left[I_3 - \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T\right] + I_3 \\ &\quad + \sigma_1^2 \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \bar{F} \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \end{aligned} \tag{57}$$

It is straightforward to show that the first three terms on the right side Eq. (57) reduce down to zero. Therefore,

$$P_{\text{sub}}\, F = I_3 + \sigma_1^2 \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \bar{F} \left[I_3 - \sigma_{\text{eff}}^2 \mathbf{b}_1^{\text{true}} (\mathbf{b}_1^{\text{true}})^T \bar{F}\right] \tag{58}$$

Clearly, $P_{\text{sub}}\,F = I_3$ only when the second term vanishes, but P_{sub} is close to the optimal covariance F^{-1} when a dominant vector exists, that is, $\sigma_1 \ll \sigma_k$ and $\sigma_1 \ll \sigma_{ij}$.

Taking the trace of $P_{\text{sub}}\,F$ gives

$$\text{Tr}(P_{\text{sub}}\,F) = 3 + \sigma_1^2\,\text{Tr}\left\{\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}\right]\bar{F}\right\} \tag{59}$$

where the idempotent property of the projection matrix has been used. Then a normalized condition to ensure that a good solution is provided is given by

$$\epsilon \equiv \frac{1}{3}\sigma_1^2\text{Tr}\left\{\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}\right]\bar{F}\right\} << 1 \tag{60}$$

The trace in Eq. (60) can be shown to be greater than or equal to zero. Since $\sigma_{\text{eff}}^2 = 1/\text{Tr}[\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}] > 0$, then it is simply required to show

$$\text{Tr}[\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}]\ \text{Tr}(\bar{F}) - \text{Tr}[\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}^2] \geq 0 \tag{61}$$

If $\mathcal{A}$ and $\mathcal{B}$ are positive semi-definite matrices, then $0 \leq \text{Tr}(\mathcal{A}\,\mathcal{B}) \leq \text{Tr}(\mathcal{A})\text{Tr}(\mathcal{B})$ [10]. Defining $\mathcal{A} \equiv \mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\,\bar{F}$ and $\mathcal{B} \equiv \bar{F}$ shows that the inequality in Eq. (61) is valid. This proves $\epsilon \geq 0$. Clearly, P_{sub} is not optimal if $\epsilon \neq 0$ or equivalently

$$\bar{\epsilon} \equiv \left[(\mathbf{b}_1^{\text{true}})^T\,\bar{F}\,\mathbf{b}_1^{\text{true}}\right]\text{Tr}(\bar{F}) - (\mathbf{b}_1^{\text{true}})^T\,\bar{F}^2\,\mathbf{b}_1^{\text{true}} \neq 0 \tag{62}$$

Note also that $\bar{\epsilon} = 0$ is not a sufficient condition for P_{sub} to be optimal.

D. Special Cases

As stated previously, Eq. (42) is valid for any $\bar{F}$. If the QMM model with unit vectors $\mathbf{b}_i^{\text{true}}$ is used for $\bar{F}$, then

$$\bar{F} = \sum_{k=2}^{n}\sigma_k^2\left[I_3 - \mathbf{b}_k^{\text{true}}(\mathbf{b}_k^{\text{true}})^T\right] = -\sum_{k=2}^{n}\sigma_k^2\,[\mathbf{b}_k^{\text{true}}\times]^2 \tag{63}$$

Then σ_{eff}^2 is given by

$$\begin{aligned}\sigma_{\text{eff}}^2 &= \left\{\sum_{k=2}^{n}\sigma_k^2\left[I_3 - (\mathbf{b}_1^{\text{true}})^T\mathbf{b}_k^{\text{true}}(\mathbf{b}_k^{\text{true}})^T\mathbf{b}_1^{\text{true}}\right]\right\}^{-1} \\ &= \left[\sum_{k=2}^{n}\sigma_k^2\|\mathbf{b}_1^{\text{true}}\times\mathbf{b}_k^{\text{true}}\|^2\right]^{-1}\end{aligned} \tag{64}$$

where the identity $\|\mathbf{a}\times\mathbf{b}\|^2 = \|\mathbf{a}\|^2\|\mathbf{b}\|^2 - (\mathbf{a}^T\mathbf{b})^2$ for any 3×1 vectors $\mathbf{a}$ and $\mathbf{b}$ has been used. This is equivalent to the σ_{eff}^2 given in [6]. Define $\mathbf{w} \equiv \sigma_{\text{eff}}^2 \bar{F}\,\mathbf{b}_1^{\text{true}}$. Substituting Eq. (63) into $\mathbf{w}$ gives

$$\begin{aligned}\mathbf{w} &= -\sigma_{\text{eff}}^2 \sum_{k=2}^{n} \sigma_k^2\,[\mathbf{b}_k^{\text{true}}\times]^2\,\mathbf{b}_1^{\text{true}} \\ &= \sigma_{\text{eff}}^2 \sum_{k=2}^{n} \mathbf{b}_k^{\text{true}} \times (\mathbf{b}_1^{\text{true}} \times \mathbf{b}_k^{\text{true}})\end{aligned} \tag{65}$$

This is equivalent to the $\mathbf{w}$ expression given in [6]. Therefore, P_{sub} is equivalent to the expression derived in [6]. Define $c_{kl} \equiv (\mathbf{b}_k^{\text{true}})^T(\mathbf{b}_l^{\text{true}})$. Using

$$\text{Tr}(\bar{F}) = 2\sum_{k=2}^{n}\sigma_k^{-2} \tag{66a}$$

$$(\mathbf{b}_1^{\text{true}})^T\,\bar{F}\,\mathbf{b}_1^{\text{true}} = \sum_{k=2}^{n}\sigma_k^{-2}(1-c_{1k}^2) \tag{66b}$$

$$(\mathbf{b}_1^{\text{true}})^T\,\bar{F}^2\,\mathbf{b}_1^{\text{true}} = \sum_{k=2}^{n}\sum_{l=2}^{n}\sigma_k^{-2}\sigma_l^{-2}\left(1-c_{1k}^2-c_{1l}^2+c_{1k}\,c_{1l}\,c_{kl}\right) \tag{66c}$$

gives

$$\bar{\epsilon} = \sum_{k=2}^{n}\sum_{l=2}^{n}\sigma_k^{-2}\sigma_l^{-2}\left(1-c_{1k}^2+c_{1l}^2-c_{1k}\,c_{1l}\,c_{kl}\right) \tag{67}$$

This equation can be rewritten as

$$\begin{aligned}\bar{\epsilon} &= \sum_{k=2}^{n}\sigma_k^{-4}(1-c_k^2) + \sum_{k=2}^{n}\sum_{l=k+1}^{n}\sigma_k^{-2}\sigma_l^{-2}\left(c_{1k}^2+c_{1l}^2-2c_{1k}c_{1l}c_{kl}\right) \\ &\geq \sum_{k=2}^{n}\sigma_k^{-4}(1-c_k^2) + \sum_{k=2}^{n}\sum_{l=k+1}^{n}\sigma_k^{-2}\sigma_l^{-2}(|c_{1k}|-|c_{1l}|)^2\end{aligned} \tag{68}$$

Clearly, $\bar{\epsilon} \neq 0$ unless all $c_k = 0$, which in turn requires that all $\mathbf{b}_k^{\text{true}}$ be coaligned with $\mathbf{b}_1^{\text{true}}$, a degenerate case with nonunique solutions. Hence, P_{sub} is not optimal for Wahba's problem.

When $\bar{F}$ is constructed from arc-lengths only, it takes the form

$$\bar{F} = \sum_{p=1}^{MN}\sigma_p^{-2}\mathbf{b}_p^{\text{true}}(\mathbf{b}_p^{\text{true}})^T \tag{69}$$

where $\mathbf{b}_p^{\text{true}}$ is a unit vector, σ_p is the associated standard deviation, and MN is the number of arc-lengths. It can be shown that

$$\bar{\epsilon} = \sum_{p}^{MN}\sum_{q}^{MN}\sigma_p^2\sigma_q^2\left(c_{1p}^2-c_{1p}c_{1q}c_{pq}\right) \tag{70}$$

When there is only one arc-length, that is, $MN = 1$, then $\bar{\epsilon} = 0$. In other cases, $\bar{\epsilon} \neq 0$ and P_{sub} is not optimal.

It is now shown that $P_{\text{sub}} F = I_3$ when the attitude is determined from a direction and an arc-length [8]. Without loss in generality it is assumed that this arc-length is given by ϕ_{11}. The FIM for this case is simply given by

$$F = \sigma_1^{-2}[I_3 - \mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T] + \bar{F}_{11} \tag{71}$$

with

$$\bar{F}_{11} \equiv \sigma_{11}^{-2}[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T \tag{72}$$

Also, P_{sub} is given by Eq. (42) with $\bar{F} = \bar{F}_{11}$. Replacing $\bar{F}$ with $\bar{F}_{11}$ in second term of the right side of Eq. (58) leads to the following requirement:

$$\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\bar{F}_{11}\right]\bar{F}_{11}\left[I_3 - \sigma_{\text{eff}}^2\,\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T\bar{F}_{11}\right] = 0_3 \tag{73}$$

Since the projection matrix and $\bar{F}_{11}$ are both singular, then the validation of this requirement must be done using a brute-force approach. Substituting $\sigma_{\text{eff}}^2 \equiv [(\mathbf{b}_1^{\text{true}})^T\bar{F}_{11}\mathbf{b}_1^{\text{true}}]^{-1}$ and Eq. (72) into Eq. (73), and after some simple algebraic manipulations leads to the following condition:

$$\begin{aligned}&\left\{(\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\mathbf{b}_1^{\text{true}}\right\}[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\\&= [A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\end{aligned} \tag{74}$$

The right side of Eq. (74) can be rewritten as

$$\begin{aligned}&[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\mathbf{b}_1^{\text{true}}(\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\\&= \left\{(\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\right\}^2[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\end{aligned} \tag{75}$$

Since $\left\{(\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\right\}^2 = (\mathbf{b}_1^{\text{true}})^T[A^{\text{true}}\mathbf{s}_1\times]\mathbf{c}_1\mathbf{c}_1^T[A^{\text{true}}\mathbf{s}_1\times]^T\mathbf{b}^{\text{true}}$ then Eq. (74) is satisfied. Therefore, P_{sub} is the optimal covariance in this case. Conditions to obtain a deterministic attitude solution for this case are discussed in [8].

IV. Simulation Results

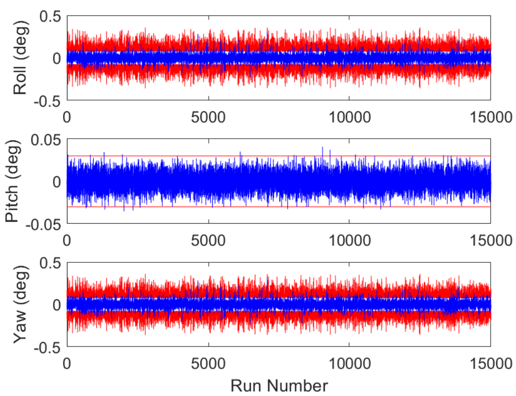
A simulation is performed using the following vectors for the LOS and GPS observations with no other LOS observations:

$$\mathbf{b}_1^{\text{true}} = (\sqrt{2}/2)[1\ \ 0\ \ 1]^T$$

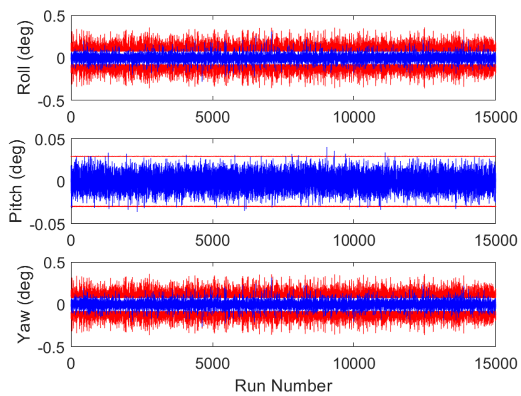
$$\mathbf{c}_1 = (\sqrt{2}/2)[0\ \ 1\ \ 1]^T, \quad \mathbf{c}_2 = [0\ \ 1\ \ 0]^T, \quad \mathbf{c}_3 = [0\ \ 0\ \ 1]^T$$

$$\mathbf{s}_1 = (\sqrt{3}/3)[1\ \ 1\ \ 1]^T, \quad \mathbf{s}_2 = (\sqrt{2}/2)[0\ \ 1\ \ 1]^T$$

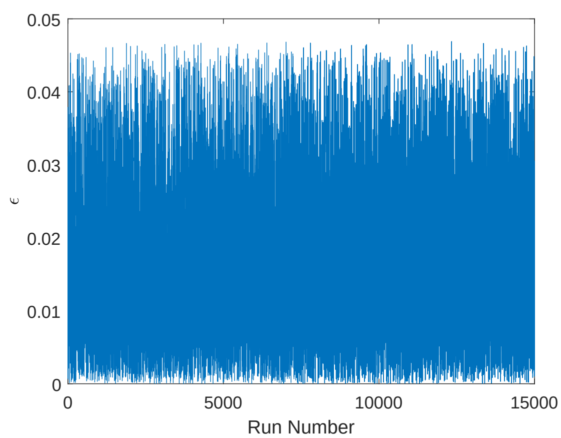
Note that the baselines are co-planer, which leads to an indeterminant solution using the approximate approach in [5]. As in the simulation example shown in [6] 15,000 test cases with uniformly distributed random attitudes are generated. The true body-vector $\mathbf{b}_1^{\text{true}}$ is corrupted by Gaussian random noise with standard deviation of 0.01 degree per axis, which simulates a fine Sun sensor. The true GPS observations are corrupted by Gaussian random noise with a normalized standard deviation of 0.001, corresponding to an attitude error of about 0.5 degrees [5].



(a) Attitude errors using the approximate solution



(b) Attitude errors using the optimal solution



(c) Trace condition

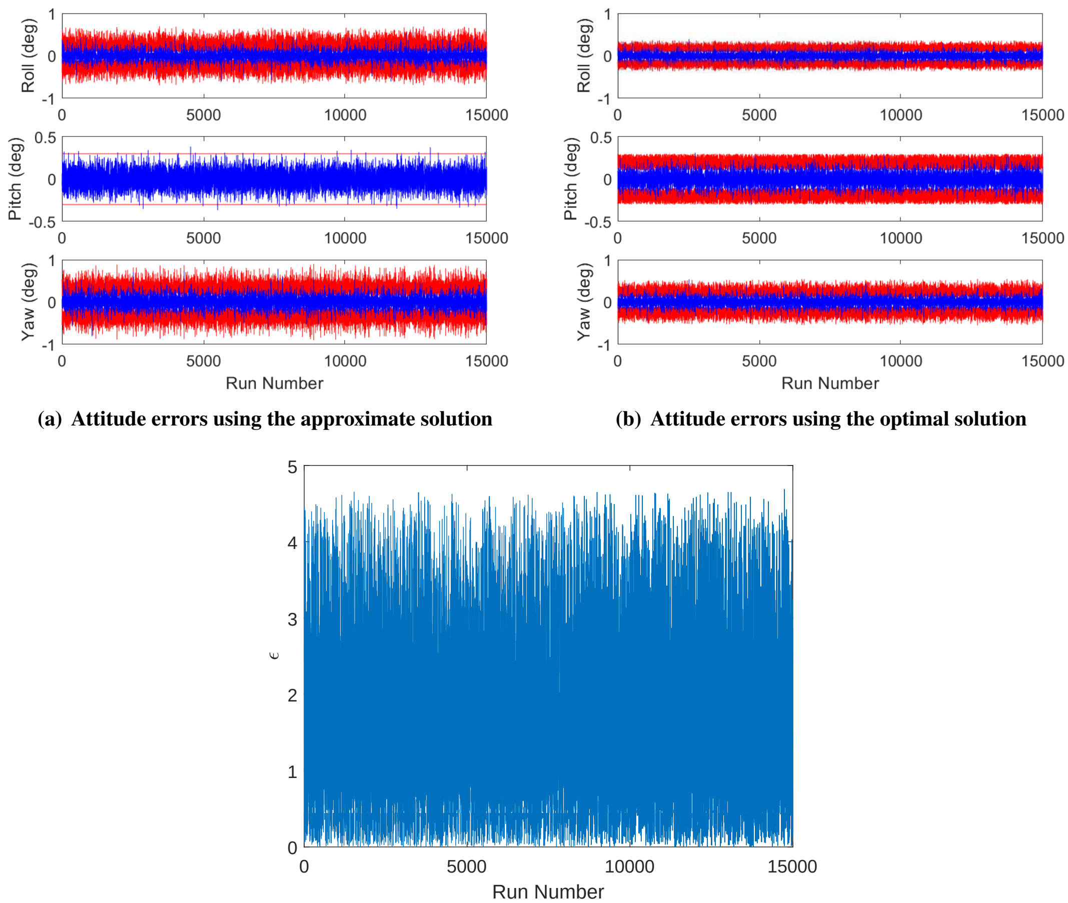
Fig. 1 GPS simulation results with fine Sun sensor.

A plot of the attitude errors using the approximate solution in Algorithm 1, along with their respective 3σ bounds, for the 15,000 test runs is shown in Fig. 1(a). The 3σ bounds change with each run because the GPS covariance is a function of the attitude. Clearly, the attitude errors are consistent with the 3σ bounds derived from P_{sub}. A plot of the attitude errors using a NLS algorithm that minimizes the optimal loss function in Eq. (1), along with their respective 3σ bounds, for the 15,000 test runs is shown in Fig. 1(b). Clearly, the attitude errors are consistent with the 3σ bounds derived from P_{opt}. Comparing Fig. 1(a) to Fig. 1(b) shows that the errors are nearly identical. A plot of the trace condition given by Eq. (60) is shown in Fig. 1(c). This indicates that the attitude solution is very close to being optimal

Table 1 Number of real roots

Case	Two Real Roots	Four Real Roots
Fine Sun Sensor Case	14,547	453
Coarse Sun Sensor Case	14,562	438

for all random attitudes. Table 1 shows the number of root solutions for the polynomial in Eq. (21) that give two real roots and four real roots for this case. A large number of cases involve only two roots.



(a) Attitude errors using the approximate solution **(b) Attitude errors using the optimal solution**

(c) Trace condition

Fig. 2 GPS Simulation results with coarse Sun sensor.

The simulation is rerun where the true body-vector $\mathbf{b}_1^{\text{true}}$ is corrupted by Gaussian random noise with standard deviation of 0.1 degree per axis, which simulates a coarser Sun sensor. The standard deviation of the GPS observations is the same as before. A plot of the attitude errors using the approximate solution in Algorithm 1, along with their respective 3σ bounds, for the 15,000 test runs is shown in Fig. 2(a). Clearly, the attitude errors are again consistent with the 3σ bounds derived from P_{sub}. But the errors are much larger than the previous case, as seen by comparing Fig. 1(a) with Fig. 2(a). This intuitively is correct because a coarse Sun sensor is used in place of a fine Sun sensor in this case. A plot of the attitude errors using a NLS algorithm that minimizes the optimal loss function in Eq. (1), along with their respective 3σ bounds, for the 15,000 test runs is shown in Fig. 2(b). Clearly, the attitude errors are consistent with the 3σ bounds derived from P_{opt}. The errors are larger than those shown in Fig. 1(b) because of the use of a coarse Sun sensor in this case. Comparing Fig. 2(a) to Fig. 2(b) shows that the errors are larger using the approximate solution than using a NLS solution that minimizes Eq. (1), with the exception of the pitch errors in some runs. The sub-optimality for the coarse Sun sensor case is confirmed by the trace condition shown in Fig. 2(c). This is now much larger than 1, which indicates that the suboptimal covariance is not close to the optimal one. The suboptimal solution can be used as a starting guess for the NLS squares algorithm to determine the optimal one, which is how the optimal solution has been determined. Table 1 shows the number of root solutions for the polynomial in Eq. (21) that give two real roots and four real roots for this case. Once again, a large number of cases involve only two roots.

V. Conclusions

This Note presented a generalized attitude determination algorithm when one dominant vector observation is provided. The algorithm involves solving a quartic polynomial, which is known to have closed-form solutions. It is also a non-iterative algorithm, and no transcendental functions are required by using the polynomial solution variable. A simple scalar expression was derived using the suboptimal covariance that can be computed without determining the attitude. This scalar quantity can be used to check the accuracy of the derived attitude solution with respect to the optimal solution. When the case of only vector observations exists, then the approximate attitude solution simplifies greatly to a previously derived solution involving this case. For the case of one vector observation and one arc-length, then the approximate attitude solution was shown to be equivalent to the optimal attitude solution. Simulation results show that a good approximate attitude solution can be provided using the derived algorithm for a realistic scenario involving one fine Sun sensor observation and several Global Positioning System (GPS) arc-length observations. The approximate attitude solution was shown to be worse than the optimal solution when a coarse Sun sensor was used in place of the fine Sun sensor. An iterative nonlinear least-squares (NLS) algorithm needs to be employed to determine the optimal attitude, which may converge to a local minimum depending on the initial guess. Using the approximate solution for the initial guess in the NLS algorithm converged to the optimal solution for every case in the simulated trails, which demonstrates the usefulness of the approximate solution even when it does not approximate the optimal solution well.

Acknowledgments

The authors would like to acknowledge many valuable conversations with F. Landis Markley.

References

[1] Markley, F. L., and Crassidis, J. L., *Fundamentals of Spacecraft Attitude Determination and Control*, Springer, New York, NY, 2014, pp. 223–224, 249. Chapters 4-5, Appendix A.

[2] Wahba, G., "A Least Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, p. 409. doi:10.1137/1007077.

[3] Shuster, M., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439–517.

[4] Shuster, M. D., and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, 1981, pp. 70–77. doi:10.2514/3.19717.

[5] Crassidis, J. L., and Markley, F. L., "New Algorithm for Attitude Determination Using Global Positioning System Signals," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, 1997, pp. 891–896. doi:10.2514/2.4162.

[6] Markley, F. L., and Cheng, Y., "Wahba's Problem with One Dominant Observation," *Journal of Guidance, Control, and Dynamics*, Vol. 41, No. 10, 2018, pp. 2315–2320. doi:10.2514/1.G003619.

[7] Shuster, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, 1989, pp. 79–88.

[8] Shuster, M. D., "Deterministic Three-Axis Attitude Determination," *Journal of the Astronautical Sciences*, Vol. 52, No. 3, 2004, pp. 405–419.

[9] Crassidis, J. L., and Junkins, J. L., *Optimal Estimation of Dynamic Systems*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL, 2012, pp. 82–83.

[10] Ulukök, Z., and Türkmen, R., "On Some Matrix Trace Inequalities," *Journal of Inequalities and Applications*, Article ID 201486, 2010. doi:10.1155/2010/201486.