AAS 11-137

OPTIMAL FUSION OF VECTOR OBSERVATIONS WITH ANGLE OR GPS PHASE DIFFERENCE OBSERVATIONS FOR THREE-AXIS ATTITUDE DETERMINATION

Sergei Tanygin*

The general problem of determining the three-axis attitude from a combination of vector and angle or GPS phase difference observations is examined. Two cost functions, one for vector observations and one for angle or GPS phase difference observations, are combined into a single maximum likelihood cost function. The initial attitude estimate is found and then updated using the Newton estimation sequence until it converges to the optimal estimate. The numerical examples are presented that demonstrate effectiveness of the new approach.

INTRODUCTION

The three-axis attitude determination for most modern spacecraft is performed using complete vector data. The two types of algorithms employed for this task are deterministic algorithms which solve for attitude using a minimal set of data and optimal algorithms which find the optimal attitude estimate by using additional measurements to minimize appropriate cost functions. The earliest of these algorithms have been introduced several decades ago and many of them have been successfully used on numerous spacecraft but the three-axis attitude determination remains the subject of active research.[1,2,3,4,5] In particular, the emergence of new and improved sensors, and continuous and rapid advancement of computing power provide opportunities for development and application of more sophisticated and capable estimation techniques. One of the relatively recent innovations is the use of GPS phase difference measurements for attitude determination.[6] Unlike vector observations, the GPS phase difference measurements are effectively angle measurements between the direction to a GPS satellite and the antenna baseline formed by two on-board GPS receivers. Since each receiver is typically capable of simultaneously processing signals from multiple GPS satellites, even a single pair of receivers can provide several angle measurements – one for each accessible GPS satellite. Nevertheless, without additional baselines or other types of observations, these measurements are insufficient for a complete three-axis attitude determination. The minimum number and types of observations needed to fully determine attitude are carefully examined by Shuster in Reference 7. In particular, it is shown that the three-axis attitude determination solely from angle observations is especially difficult.[7] Given the sufficient number of angle measurements, it is possible to construct the optimal cost function that should be minimized by the optimal attitude estimate. However, unlike the well-known Wahba cost function for vector observations that is quadratic in quaternions and can be elegantly minimized in closed form,[1-5] the cost function for angle observations is quartic in quaternions and af-

* Senior Astrodynamics Specialist, Analytical Graphics, Inc., 220 Valley Creek Blvd., Exton, PA 19341.

fords no such closed form solution. Instead various iterative or approximate solutions have been proposed. Cohen proposes to iteratively refine the initial attitude estimate by adopting a small angle linearization.[8] This local linearization approach may fail to converge or require many iterations if the initial estimate is poor. Crassidis and Markley[9] as well as Bar-Itzhak et al.[10] propose to solve this problem by constructing a derived set of vector observations and transforming the angle cost function into the Wahba form. This approach works well only when the antenna baselines or the sightlines to GPS satellites are proportional to an orthonormal basis. Otherwise, significant errors may arise culminating in the case when the baselines are coplanar. In Reference 12, Crassidis et al. propose ALLEGRO – an algorithm based on a predictive filtering approach[11] – which converges to the optimal estimate and the corresponding optimal error covariance provided that the observation sampling is fairly frequent.[12] Recently, another approach based on homotopy continuation rather than on local linearization or filtering has been introduced.[13] Given any initial attitude estimate a new set of pseudo-observations is created such that together with the initial estimate they minimize the angle cost function. Then as the pseudo-observations evolve toward the actual observations so do the cost function and its minimizing estimate. In the end, the estimate arrives at the optimal solution just as the pseudo-observations match the actual observations. This method is shown to have excellent convergence properties: faster than gradient based methods and more robust than local linearization techniques. This work focuses on finding the optimal estimate in cases when both vector and angle observations are available. Given at least a single vector observation and a sufficient number of angle observations, the initial estimate can be computed using methods described in Reference 7. The optimal estimate is obtained by minimizing a combined cost function that includes both the classical Wahba form for vector observations and additional terms for angle observations. The algorithm involves updating the initial quaternion estimate via Newton iterations while carefully considering quaternion norm constraint. This paper uses the Small Satellite Technology Initiative (SSTI) Lewis spacecraft also used in several other references to test performance of the proposed algorithm with GPS phase difference measurements. It should be noted however that many other types of angle measurements can be formulated using the same measurement model[14] and therefore be handled by the proposed approach.

MEASUREMENT MODELS

The three-axis attitude determination algorithm proposed in this paper operates using both vector and angle observations. Thus, the measurement models for both types of observations are introduced first. For brevity, the models in this paper are derived assuming uncorrelated measurements, however, correlations between angle measurements can be easily added as shown in Reference 13.

QUEST Measurement Model for Vector Observations

The QUEST measurement model[1] is adopted for vector observations. Let A denote the rotation (or attitude) matrix which is a 3×3 direction cosine matrix mapping vector components from the reference frame to the body-fixed frame and let $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ be 3×1 column-vectors representing the same unit vector known in components in the reference frame and measured in components in the body-fixed frame, respectively. Then according to the QUEST measurement model $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are related via

$$\hat{\mathbf{b}} = \mathsf{A}\hat{\mathbf{a}} + \mathbf{w} \text{ with } \mathbf{w} \sim N(\mathbf{0},\mathsf{R}) \text{ and } \mathsf{R}^{-1} = \sigma^{-2}\left(\mathsf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}^{\mathrm{T}}\right), \tag{1}$$

where R is the measurement error covariance of the white Gaussian noise $\mathbf{w}$.

Measurement Model for Angle Observations

The measurement model for angle observations uses a dot product of two vectors: one is known in the reference frame and the other is known in the body-fixed frame. The reference vectors may represent directions to known celestial bodies, GPS satellites or magnetic field lines. The body-fixed vectors may represent various sensing directions of various sensors, or antenna baselines between GPS receivers. For details see, for example, References 9, 13 and 14. Let $\mathbf{r}$ and $\mathbf{s}$ be 3×1 (not necessarily unit) column-vectors representing the reference and body-fixed directions, respectively. The effective angle measurement is obtained via the dot product of $\mathbf{r}$ and $\mathbf{s}$ which, since they are known in different frames, involves the attitude matrix A :

$$d = \mathbf{s}^{\mathrm{T}}\mathsf{A}\mathbf{r} + w \text{ with } w \sim N(0,\sigma^2), \tag{2}$$

where σ^2 is the variance of the measurement error due to the white Gaussian noise w. Let $\hat{\mathbf{q}}$ denote the quaternion equivalent of A then the effective measurement in terms of $\hat{\mathbf{q}}$ becomes

$$d = \hat{\mathbf{q}}^{\mathrm{T}}\mathsf{K}(\mathbf{r},\mathbf{s})\hat{\mathbf{q}} + w, \tag{3}$$

where $\mathsf{K}(\mathbf{r},\mathbf{s})$ denotes the 4×4 symmetric matrix function which for any 3×1 column-vectors of $\mathbf{r}$ and $\mathbf{s}$ is defined as follows

$$\mathsf{K}(\mathbf{r},\mathbf{s}) = \begin{bmatrix} \mathbf{s}\mathbf{r}^{\mathrm{T}} + \mathbf{r}\mathbf{s}^{\mathrm{T}} - \left(\mathbf{r}^{\mathrm{T}}\mathbf{s}\right)\mathsf{I} & -\left(\mathbf{r}\times\mathbf{s}\right) \\ -\left(\mathbf{r}\times\mathbf{s}\right)^{\mathrm{T}} & \mathbf{r}^{\mathrm{T}}\mathbf{s} \end{bmatrix}. \tag{4}$$

Here and throughout the paper I denotes the identity matrix.

QUATERNION COST FUNCTION

The three-axis attitude determination algorithm produces the optimal maximum-likelihood estimate if it minimizes the appropriate cost function. The data-dependent part of the negative-log-likelihood function must add statistically weighted squares of residuals from all measurements. It is instructive to examine vector and angle observations separately and then combine them into a single cost function.

Cost Function for Vector Observations

The problem of determining the three-axis attitude from vector observations has a long and storied history. The cost function stemming from the Wahba attitude determination problem[15] is presented below:

$$\psi(\hat{\mathbf{q}}) = \frac{1}{2}\sum_{m=1}^{M} a_m \left(\hat{\mathbf{b}}_m - \mathsf{A}\hat{\mathbf{a}}_m\right)^{\mathrm{T}}\left(\hat{\mathbf{b}}_m - \mathsf{A}\hat{\mathbf{a}}_m\right) = 1 - \sum_{m=1}^{M} a_m \hat{\mathbf{b}}_m^{\mathrm{T}}\mathsf{A}\hat{\mathbf{a}}_m = 1 - \hat{\mathbf{q}}^{\mathrm{T}}\tilde{\mathsf{K}}_v\hat{\mathbf{q}} = \hat{\mathbf{q}}^{\mathrm{T}}\left(\mathsf{I} - \tilde{\mathsf{K}}_v\right)\hat{\mathbf{q}} \tag{5}$$

Here the weights a_m add up to 1 by construction because of the way they are determined from the individual measurement variances σ_m^2

$$a_m = \frac{\sigma_v^2}{\sigma_m^2} \text{ and } \frac{1}{\sigma_v^2} = \sum_{m=1}^{M} \frac{1}{\sigma_m^2}. \tag{6}$$

Note that the famed "K-matrix" discovered by Paul Davenport in his celebrated q-Method and denoted here as $\tilde{\mathsf{K}}_v$ is simply a weighted sum of the individual observation "K-matrices" computed according to Eq. (4):

$$\tilde{\mathsf{K}}_v = \sum_{m=1}^{M} a_m \mathsf{K}_m \text{ with } \mathsf{K}_m = \mathsf{K}(\hat{\mathbf{a}}_m, \hat{\mathbf{b}}_m) \tag{7}$$

In the q-Method, the optimal estimate is determined by maximizing $\hat{\mathbf{q}}^{\mathrm{T}}\tilde{\mathsf{K}}_v\hat{\mathbf{q}}$ instead of minimizing $\psi(\hat{\mathbf{q}})$. Subject to the unit norm quaternion constraint the two approaches are clearly equivalent. In either case, the optimal quaternion is parallel to the eigenvector of $\tilde{\mathsf{K}}_v$ associated with its largest positive eigenvalue. This eigenvalue is equal to 1 if observations are perfect but is reduced otherwise by the amount correlated with measurement errors. In practice, it remains very close to 1 for typical errors. It is clear that, when using the last form of the cost function $\psi(\hat{\mathbf{q}})$ from Eq.(5), minimizing $\hat{\mathbf{q}}^{\mathrm{T}}\left(\mathsf{I} - \tilde{\mathsf{K}}_v\right)\hat{\mathbf{q}}$ is equivalent to maximizing $\hat{\mathbf{q}}^{\mathrm{T}}\tilde{\mathsf{K}}_v\hat{\mathbf{q}}$ subject to the unit norm quaternion constraint. Note that $\mathsf{I} - \tilde{\mathsf{K}}_v$ is symmetric and positive semi-definite (positive definite when observations are imperfect). Its smallest eigenvalue is equal to 1 minus the largest eigenvalue of $\tilde{\mathsf{K}}_v$ and it has the same associated eigenvector. Also, note that the cost function $\psi(\hat{\mathbf{q}})$ is not the actual data-dependent part of the negative-log-likelihood function but the two functions are related via a simple constant scaling:

$$J_v(\hat{\mathbf{q}}) = \frac{\psi(\hat{\mathbf{q}})}{\sigma_v^2}. \tag{8}$$

Here $J_v(\hat{\mathbf{q}})$ is the data-dependent part of the negative-log-likelihood function that incorporates all vector observations.

Cost Function for Angle Observations

The three-axis attitude determination from angle observations is a more intricate problem then its vector observations counterpart because each effective angle measurement provides less information about the attitude then each vector measurement and because the cost function related to the data-dependent part of the negative-log-likelihood function becomes quartic in the quaternion representation of attitude. Consider the cost function $\phi(\hat{\mathbf{q}})$ that adds statistically weighted squares of measurement residuals from Eq. (3):

$$\phi(\hat{\mathbf{q}}) = \frac{1}{2}\sum_{n=1}^{N} a_n \left(\hat{\mathbf{q}}^{\mathrm{T}}\mathsf{K}_n\hat{\mathbf{q}} - d_n\right)^2, \tag{9}$$

Here, similar to the vector observation cost function in Eq. (5), the weights a_n add up to 1 by construction because of the way they are determined from the individual measurement variances σ_n^2

$$a_n = \frac{\sigma_a^2}{\sigma_n^2} \text{ and } \frac{1}{\sigma_a^2} = \sum_{n=1}^{N} \frac{1}{\sigma_n^2} . \tag{10}$$

Here, similar to vector observations, K_n are the "K-matrices" defined for individual observations using corresponding reference and body-fixed unit vectors:

$$\mathsf{K}_n = \mathsf{K}(\mathbf{r}_n, \mathbf{s}_n) \tag{11}$$

but, unlike in the cost function for vector observations, in this cost function these matrices cannot be simply combined into one. On the other hand, like the cost function for vector observations, this cost function is not the actual data-dependent part of the negative-log-likelihood function but is related to it via a simple constant scaling:

$$J_a(\hat{\mathbf{q}}) = \frac{\phi(\hat{\mathbf{q}})}{\sigma_a^2} . \tag{12}$$

Here $J_a(\hat{\mathbf{q}})$ is the data-dependent part of the negative-log-likelihood function that incorporates all angle observations.

Combined Cost Function

The combined cost function that incorporates all observations is then simply a weighted combination of the two cost functions listed in Eqs. (5) and (9). Consider

$$\eta(\hat{\mathbf{q}}) = \rho_v \psi(\hat{\mathbf{q}}) + \rho_a \phi(\hat{\mathbf{q}}) , \tag{13}$$

where the weights ρ_v and ρ_a are derived from the overall variances of the vector and angle observations, σ_v^2 and σ_a^2:

$$\rho_v = \frac{\sigma^2}{\sigma_v^2} , \; \rho_a = \frac{\sigma^2}{\sigma_a^2} \text{ and } \frac{1}{\sigma^2} = \frac{1}{\sigma_a^2} + \frac{1}{\sigma_v^2} \tag{14}$$

Then the combined cost function is related to the data-dependent part of the negative-log-likelihood function via

$$J(\hat{\mathbf{q}}) = \frac{\eta(\hat{\mathbf{q}})}{\sigma^2} = J_v(\hat{\mathbf{q}}) + J_a(\hat{\mathbf{q}}) , \tag{15}$$

where $J(\hat{\mathbf{q}})$ incorporates all vector and angle observations. The attitude error covariance associated with the optimal estimate that minimizes $J(\hat{\mathbf{q}})$ can be obtained from the inverse of the Fisher information matrix[16] which itself is straightforward to obtain as a sum of the Fisher information matrices known for $J_v(\hat{\mathbf{q}})$ and $J_a(\hat{\mathbf{q}})$ (see References 16 and 9 for the corresponding derivations). The resulting expression for the Fisher information matrix and equivalently for the inverse of the optimal attitude error covariance becomes:

$$\mathsf{P}_{\xi\xi}^{*-1} = \sum_{m=1}^{M} \frac{1}{\sigma_m^2} \left[\mathsf{I} - \hat{\mathbf{b}}_{m(true)} \hat{\mathbf{b}}_{m(true)}^{\mathrm{T}} \right] + \sum_{n=1}^{N} \frac{1}{\sigma_n^2} \left[\mathbf{s}_n \times \left(\mathsf{A}_{(true)} \mathbf{r}_n \right) \right] \left[\mathbf{s}_n \times \left(\mathsf{A}_{(true)} \mathbf{r}_n \right) \right]^{\mathrm{T}} \tag{16}$$

COST FUNCTION MINIMIZATION

In order to avoid dealing with the unit norm quaternion constraint during minimization of $\eta(\hat{\mathbf{q}})$, the modified Rodrigues parameters $\mathbf{p}$ are introduced as intermediate variables. Relationships between $\mathbf{p}$ and $\hat{\mathbf{q}}$ are well known and computationally inexpensive:[17]

$$\mathbf{p} = \frac{\mathbf{q}}{1+q_4} \quad \text{and} \quad \hat{\mathbf{q}} = \frac{1}{1+p^2} \begin{bmatrix} 2\mathbf{p} \\ 1-p^2 \end{bmatrix}, \tag{17ab}$$

where $p^2 = \mathbf{p}^{\mathrm{T}}\mathbf{p}$, $\mathbf{q}$ is the vector part of the quaternion and q_4 is the scalar part.

Let $\hat{\mathbf{q}}_0$ be the initial quaternion estimate and $\mathbf{p}_0$ be derived from $\hat{\mathbf{q}}_0$ using Eq.(17a). Updating estimates for $\mathbf{p}$ using the Newton estimation sequence requires the gradient $\eta'_{\mathbf{p}}$ and the Hessian $\eta''_{\mathbf{pp}}$. The gradient of η with respect to $\hat{\mathbf{q}}$ is obtained by direct differentiation of Eq.(13) using expressions from Eqs.(5) and (9):

$$\eta'_{\hat{\mathbf{q}}} = 2\left(\rho_v \hat{\mathbf{q}}^{\mathrm{T}} \left(\mathsf{I} - \tilde{\mathsf{K}}_v \right) + \rho_a \sum_{n=1}^{N} a_n \left(\hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \hat{\mathbf{q}} - d_n \right) \hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \right). \tag{18}$$

The gradient of η with respect to $\mathbf{p}$ is obtained via the chain rule

$$\eta'_{\mathbf{p}} = \eta'_{\hat{\mathbf{q}}} \hat{\mathbf{q}}'_{\mathbf{p}} = 2\left(\rho_v \hat{\mathbf{q}}^{\mathrm{T}} \left(\mathsf{I} - \tilde{\mathsf{K}}_v \right) + \rho_a \sum_{n=1}^{N} a_n \left(\hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \hat{\mathbf{q}} - d_n \right) \hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \right) \hat{\mathbf{q}}'_{\mathbf{p}}, \tag{19}$$

where from differentiating Eq.(17b)

$$\hat{\mathbf{q}}'_{\mathbf{p}} = \begin{bmatrix} (1+q_4)\mathsf{I} \\ -\mathbf{q}^{\mathrm{T}} \end{bmatrix} - \hat{\mathbf{q}}\mathbf{q}^{\mathrm{T}}. \tag{20}$$

The Hessian of η with respect to $\hat{\mathbf{q}}$ is obtained by additional differentiation of Eq.(18)

$$\eta''_{\hat{\mathbf{q}}\hat{\mathbf{q}}} = 2\left(\rho_v \left(\mathsf{I} - \tilde{\mathsf{K}}_v \right) + \rho_a \sum_{n=1}^{N} a_n \left[\left(\hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \hat{\mathbf{q}} - d_n \right) \mathsf{K}_n + 2\mathsf{K}_n \hat{\mathbf{q}} \hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \right] \right) \tag{21}$$

and the Hessian with respect to $\mathbf{p}$ is obtained via another application of the chain rule:

$$
\begin{aligned}
\eta''_{\mathbf{pp}} &= \hat{\mathbf{q}}'^{\mathrm{T}}_{\mathbf{p}} \eta''_{\hat{\mathbf{q}}\hat{\mathbf{q}}} \hat{\mathbf{q}}'_{\mathbf{p}} + \sum_{i=1}^{4} \eta'_{q_i} q''_{i\mathbf{pp}} \\
&= 2\hat{\mathbf{q}}'^{\mathrm{T}}_{\mathbf{p}} \left(\rho_v \left(\mathsf{I} - \tilde{\mathsf{K}}_v\right) + \rho_a \sum_{n=1}^{N} a_n \left[\left(\hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \hat{\mathbf{q}} - d_n \right) \mathsf{K}_n + 2\mathsf{K}_n \hat{\mathbf{q}} \hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \right] \right) \hat{\mathbf{q}}'_{\mathbf{p}} \ , \\
&\quad + \sum_{i=1}^{4} \eta'_{q_i} q''_{i\mathbf{pp}}
\end{aligned}
\tag{22}
$$

Here η'_{q_i} is simply the ith element of the gradient row-vector $\eta'_{\hat{\mathbf{q}}}$ from Eq.(18) and $q''_{i\mathbf{pp}}$ are the Hessians of individual quaternion elements derived by additional differentiation from Eq.(20):

$$
q''_{1\mathbf{pp}} = 2q_1 \mathbf{q}\mathbf{q}^{\mathrm{T}} - \left(1 + q_4\right) \begin{bmatrix} 3q_1 & q_2 & q_3 \\ q_2 & q_1 & 0 \\ q_3 & 0 & q_1 \end{bmatrix}, \tag{23}
$$

$$
q''_{2\mathbf{pp}} = 2q_2 \mathbf{q}\mathbf{q}^{\mathrm{T}} - \left(1 + q_4\right) \begin{bmatrix} q_2 & q_1 & 0 \\ q_1 & 3q_2 & q_3 \\ 0 & q_3 & q_2 \end{bmatrix}, \tag{24}
$$

$$
q''_{3\mathbf{pp}} = 2q_3 \mathbf{q}\mathbf{q}^{\mathrm{T}} - \left(1 + q_4\right) \begin{bmatrix} q_3 & 0 & q_1 \\ 0 & q_3 & q_2 \\ q_1 & q_2 & 3q_3 \end{bmatrix}, \tag{25}
$$

$$
q''_{4\mathbf{pp}} = \left(1 + q_4\right)\left(2\mathbf{q}\mathbf{q}^{\mathrm{T}} - \left(1 + q_4\right)\mathsf{I}\right). \tag{26}
$$

The Newton estimation sequence for $\mathbf{p}$ is given by

$$
\mathbf{p}_{k+1} = \mathbf{p}_k - \left[\eta''_{\mathbf{pp}}(\mathbf{p}_k)\right]^{-1} \eta'^{\mathrm{T}}_{\mathbf{p}}(\mathbf{p}_k), \ k = 0,1,2,... \tag{27}
$$

where at every step $\hat{\mathbf{q}}_k$ is obtained from $\mathbf{p}_k$ using Eq.(17a). Assuming that the initial estimate $\hat{\mathbf{q}}_0$ is sufficiently accurate, the Hessian inverse $\left[\eta''_{\mathbf{pp}}\right]^{-1}$ can be approximated by using only the dominant terms, i.e. by using $\left[\eta''_{\mathbf{pp}}\right]^{-1} \approx \mathsf{H}^{-1}$ where

$$
\mathsf{H} = 2\hat{\mathbf{q}}'^{\mathrm{T}}_{\mathbf{p}} \left(\rho_v \left(\mathsf{I} - \tilde{\mathsf{K}}_v\right) + 2\rho_a \sum_{n=1}^{N} a_n \mathsf{K}_n \hat{\mathbf{q}} \hat{\mathbf{q}}^{\mathrm{T}} \mathsf{K}_n \right) \hat{\mathbf{q}}'_{\mathbf{p}}. \tag{28}
$$

This matrix is positive definite (assuming that the attitude is observable).

The error covariance of the optimal estimate $\mathbf{p}^* = \lim_{k\to\infty} \mathbf{p}_k$ is given by the inverse of the Fisher information matrix $J''_{\mathbf{pp}(true)}$ and can be obtained from Eq.(15) as

$$\mathsf{P}^*_{\mathbf{pp}} = \left[J''_{\mathbf{pp}(true)}\right]^{-1} = \sigma^2 \left[\eta''_{\mathbf{pp}(true)}\right]^{-1}. \tag{29}$$

Then the attitude error covariance matrix is given by

$$\mathsf{P}^*_{\xi\xi} = \xi'_{\mathbf{p}(true)} \mathsf{P}^*_{\mathbf{pp}} \xi'^{\mathrm{T}}_{\mathbf{p}(true)} = \sigma^2 \xi'_{\mathbf{p}(true)} \left[\eta''_{\mathbf{pp}(true)}\right]^{-1} \xi'^{\mathrm{T}}_{\mathbf{p}(true)}, \tag{30}$$

where

$$\xi'_{\mathbf{p}} = 2(1+q_4)\frac{(1+q_4)\mathsf{I} + [\![\mathbf{q}]\!]}{(1+q_4)\mathsf{I} - [\![\mathbf{q}]\!]}. \tag{31}$$

Here and throughout the paper $[\![\ ]\!]$ denotes a skew-symmetric matrix defined for any vector $\mathbf{v}$ as

$$[\![\mathbf{v}]\!] = \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix}. \tag{32}$$

This attitude error covariance is equivalent to the one defined by Eq.(16).

INITIAL ESTIMATE

The method by which the initial estimate can be obtained depends on the types and number of available observations.

Multiple Vector Observations

If at least two non-collinear vector observations are available then the initial estimate can be obtained by solving the Wahba problem using all vector observations. In this case the initial quaternion estimate $\hat{\mathbf{q}}_0$ is the unit eigenvector of $\tilde{\mathsf{K}}_v$ associated with the largest eigenvalue.

Single Vector Observation

If only a single non-collinear vector observation is available then there need to be at least two independent angle observations. In this case the initial estimate can be obtained following the approach advocated by Shuster in Reference 7. Let $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{b}}_1$ be the reference unit vector and its body-fixed observation. Then a candidate quaternion $\tilde{\hat{\mathbf{q}}}_0$ and the corresponding matrix $\tilde{\mathsf{A}}_0$ can be found as

$$\tilde{\hat{\mathbf{q}}}_0 = \frac{1}{\sqrt{2\left(1+\hat{\mathbf{a}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1\right)}}\begin{bmatrix}\hat{\mathbf{a}}_1+\hat{\mathbf{b}}_1\\ 0\end{bmatrix} \tag{33}$$

and

$$\tilde{\mathsf{A}}_0 = \frac{\left(\hat{\mathbf{a}}_1+\hat{\mathbf{b}}_1\right)\left(\hat{\mathbf{a}}_1+\hat{\mathbf{b}}_1\right)^{\mathrm{T}}}{1+\hat{\mathbf{a}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1} - \mathbf{I}. \tag{34}$$

These represent a different solution from the one found in Reference 7 but this solution is equally valid and is simpler to compute. If $\hat{\mathbf{a}}_1$ and $\hat{\mathbf{b}}_1$ are antipodal then any spin-axis direction can be chosen as long as the rotation angle is set to π. Let $\hat{\mathbf{r}}_1$ and $\hat{\mathbf{s}}_1$ define the most accurate angle observation d_1. Then there exists some rotation angle θ about $\hat{\mathbf{b}}_1$ such that[7]

$$d_1 = \hat{\mathbf{s}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_\Delta(\hat{\mathbf{b}}_1,\theta)\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1, \tag{35}$$

where $\tilde{\mathsf{A}}_\Delta(\hat{\mathbf{b}}_1,\beta)$ denotes the additional rotation matrix defined as

$$\tilde{\mathsf{A}}_\Delta(\hat{\mathbf{b}}_1,\beta) = \mathbf{I} + \frac{2u}{1+u^2}\left[\!\left[\hat{\mathbf{b}}_1\right]\!\right] + \frac{2u^2}{1+u^2}\left[\!\left[\hat{\mathbf{b}}_1\right]\!\right]^2 \text{ with } u = \tan\frac{\theta}{2}. \tag{36}$$

Substituting Eq.(36) into Eq.(35) and performing some algebra yields the following scalar equation which is quadratic in u:

$$\left(2\hat{\mathbf{s}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1\hat{\mathbf{b}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1 - d_1 - \hat{\mathbf{s}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1\right)u^2 + \left(2\left(\hat{\mathbf{b}}_1\times\hat{\mathbf{s}}_1\right)^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1\right)u + \hat{\mathbf{s}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1 - d_1 = 0. \tag{37}$$

Its solution can be reduced to

$$u_{\Delta 1,2} = 2\frac{\pm\sqrt{1-\left(\hat{\mathbf{s}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1\right)^2 - \left(\hat{\mathbf{b}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1\right)^2 - d_1^2 + 2\left(\hat{\mathbf{s}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1\right)\left(\hat{\mathbf{b}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1\right)d_1} - \left(\hat{\mathbf{b}}_1\times\hat{\mathbf{s}}_1\right)^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1}{2\left(\hat{\mathbf{s}}_1^{\mathrm{T}}\hat{\mathbf{b}}_1\right)\left(\hat{\mathbf{b}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1\right) - d_1 - \hat{\mathbf{s}}_1^{\mathrm{T}}\tilde{\mathsf{A}}_0\hat{\mathbf{r}}_1} \tag{38}$$

which is equivalent to but somewhat more efficient than methods proposed in Reference 7. The two possible solutions $u_{\Delta 1}$ and $u_{\Delta 2}$ can be used to create the corresponding quaternions

$$\hat{\mathbf{q}}_{\Delta i} = \frac{1}{\sqrt{1+u_{\Delta i}^2}}\begin{bmatrix}\hat{\mathbf{b}}_1 u_{\Delta i}\\ 1\end{bmatrix},\ i = 1,2, \tag{39}$$

which in turn can be combined with $\tilde{\hat{\mathbf{q}}}_0$ from Eq.(33) to create the two possible initial quaternion estimates $\hat{\mathbf{q}}_{0i}$. Of the two the one resulting in the smaller angle cost function $\phi(\hat{\mathbf{q}}_{0i})$ should be selected as the actual initial estimate.

No Vector Observations

If no vector observations are available then obviously the problem of combining vector and angle observations reduces to simply using all available angle observations. This problem lies outside of the scope of this paper but there exist various methods for solving this problem in general[8-10,12,13] and for obtaining the initial estimate specifically[7].

CONVERGENCE

It is instructive to examine convergence properties of the Newton sequence in Eq.(27). In general, within the sufficiently small neighborhood its convergence should be quadratic. However, because the original quaternion formulation is constrained and it is projected into a subspace to make it unconstrained, it is important to ensure that this projection does not introduce extraneous critical points that may interfere with convergence. The potential problem is easy to see if the gradient vector $\eta_{\hat{\mathbf{q}}}^{\prime\mathrm{T}}$ becomes collinear with the estimate $\hat{\mathbf{q}}$. If this were to occur then the projected gradient vector $\eta_{\mathbf{p}}^{\prime\mathrm{T}}$ would be reduced to zero and the Newton sequence would cease to update (to verify this, examine Eq.(19) and note that $\hat{\mathbf{q}}^{\mathrm{T}}\hat{\mathbf{q}}_{\mathbf{p}}' = \mathbf{0}^{\mathrm{T}}$). Let λ be the alignment parameter defined as

$$\lambda = \left|\cos\delta\right| = \left\|\eta_{\hat{\mathbf{q}}}'\right\|^{-1}\left|\eta_{\hat{\mathbf{q}}}'\hat{\mathbf{q}}\right| \tag{40}$$

which is the absolute value of the cosine of the angle between the gradient vector $\eta_{\hat{\mathbf{q}}}^{\prime\mathrm{T}}$ and the estimate $\hat{\mathbf{q}}$. Then the alignment occurs when $\lambda = 1$. As $\lambda \to 1$, its increment becomes

$$\Delta\lambda = -\left\|\eta_{\hat{\mathbf{q}}}'\right\|^{-1}\eta_{\hat{\mathbf{q}}}'\hat{\mathbf{q}}_{\mathbf{p}}'\left[\eta_{\mathbf{pp}}''\right]^{-1}\hat{\mathbf{q}}_{\mathbf{p}}'^{\mathrm{T}}\eta_{\hat{\mathbf{q}}}'^{\mathrm{T}}\,. \tag{41}$$

As long as the Hessian $\eta_{\mathbf{pp}}''$ is positive definite, which should be expected in the vicinity of the minimum, this increment $\Delta\lambda$ will be negative. This means that $\lambda = 1$ is an unstable condition which should not interfere with the convergence of the Newton sequence. The same is true when the actual Hessian is replaced by its approximation H from Eq.(29) because H is positive definite by construction.

NUMERICAL EXAMPLES

Consider the simulated orbit and attitude profile of the Small Satellite Technology Initiative (SSTI) Lewis satellite. It was launched on August 22, 1997 with the mission to carry out a GPS attitude determination experiment. The spacecraft was lost due to a malfunction not related to the GPS experiment but its orbit and attitude profile as well as its hardware characteristics remain representative of missions that may employ this type of attitude determination. The orbit parameters are given in Table 1. The attitude is assumed to be nadir pointing (z-axis) and constrained along the ground track (x-axis).

Table 1. SSTI Lewis Orbit Parameters.

Parameter	Value
Semimajor axis a	6901.137 km
Inclination i	97.45 deg
Right ascension of ascending node (RAAN)	-157.1 deg
Eccentricity e Pointing profile	0.0001 Earth pointed
Launch date	Aug. 22, 1997

There are four GPS antennas that form three baselines, directions of which are listed below as unit vectors in components in the body-fixed frame in Table 2.

Table 2. SSTI Lewis Normalized Antenna Baselines.

Baseline	Body-Fixed, s_1	Body-Fixed, s_2	Body-Fixed, s_3
1	0.858265169	0.511838137	-0.037451571
2	0.000000	0.999633807	-0.027060151
3	-0.690401856	0. 690401856	-0.216079970

The directional accuracies of various measurement types are defined by the standard deviations listed in Table 3.

Table 3. Directional Accuracies of Various Measurement Types.

Measurement Type	Standard Deviation, σ
Star Tracker	0.00001
Sun Sensor	0.0001
Magnetometer	0.0005
GPS	0.005

During the simulation there are 15 GPS satellites available on 5 Feb 2011 10:00:00.000 UTCG but only 4 are considered in this study. The lines-of-sight to these satellites as unit vectors in components in ICRF reference frame are listed in Table 4.

Table 4. SSTI Lewis – GPS Satellites Lines-of-Sight on 5 Feb 2011 10:00:00.000 UTCG.

GPS PRN Number	ICRF, r_1	ICRF, r_2	ICRF, r_3
2	-0.811593908	0.385889079	-0.438639883
3	0.579466523	-0.784621653	0.220425522

4	-0.992306080	0.061787935	-0.107288836
5	-0.386180120	0.794366097	-0.468878896

At that moment, there are also several reference directions available for vector observations. They are listed in Table 5 again as unit vectors in components in ICRF reference frame.

Table 5. SSTI Lewis –Vector Reference Directions on 5 Feb 2011 10:00:00.000 UTCG.

Type	ICRF, a_1	ICRF, a_2	ICRF, a_3
Sun	0.720354063	-0.636395902	-0.275844667
IGRF Magnetic Field	0.172838128	-0.370115932	-0.912765676
Star HP-100751	-0.055793181	-0.130315629	0.989901489
Star HP-109268	-0.482668111	-0.632151510	-0.606148466

Finally, the true attitude of the Lewis spacecraft is listed in Table 6 as quaternion components relating ICRF and body-fixed frames.

Table 6. SSTI Lewis True Attitude on 5 Feb 2011 10:00:00.000 UTCG.

ICRF to Body, q_1	ICRF to Body, q_2	ICRF to Body, q_3	ICRF to Body, q_4
0.084752986	-0.049301463	-0.973427007	0.206944822

Covariance Analysis

First the efficacy of adding angle observations is examined by comparing the attitude error covariance matrices in different cases.

Case 1: All Vector Observations + (4 GPS Lines-of-sight x 3 baselines). Consider the case when the Sun, the magnetic field and the two stars are available for vector observations. In addition, assume that signals from 4 GPS satellites are detected by all receivers producing a total of 12 angle observations – 4 observations for each of the 3 baselines. The attitude error covariance based solely on the vector observations is

$$\mathsf{P}^{*}_{\xi\xi(vector)} = \begin{bmatrix} 91.1821 & 9.6425 & -54.3778 \\ 9.6425 & 54.9010 & -2.1866 \\ -54.3778 & -2.1866 & 163.3128 \end{bmatrix} \times 10^{-12} . \tag{42}$$

The attitude error covariance after the inclusion of all of the angle observations is only insignificantly smaller:

$$\mathsf{P}^{*}_{\xi\xi(vector+angle)} = \begin{bmatrix} 91.1813 & 9.6423 & -54.3759 \\ 9.6423 & 54.9009 & -2.1863 \\ -54.3759 & -2.1863 & 163.3073 \end{bmatrix} \times 10^{-12} . \tag{43}$$

The marginal improvement after addition of the angle observations is expected because the accuracy of the vector observations is significantly higher.

Case 2: (1 Magnetometer + 1 Sun) Vector Observations + (4 GPS Lines-of-sight x 3 baselines). The improvement is more significant when only two (relatively poor) vector observations are included: the Sun and the magnetic field. The attitude error covariance when using only these vector observations is about three orders of magnitude larger than when using all vector observations (including the high accuracy star observations):

$$\mathsf{P}^*_{\xi\xi(vector)} = \begin{bmatrix} 54.9692 & -110.0467 & 61.4764 \\ -110.0467 & 276.7700 & -149.4247 \\ 61.4764 & -149.4247 & 93.4317 \end{bmatrix} \times 10^{-9}. \tag{44}$$

When the angle observations are added to the two vector observations, the attitude error covariance improvement is still small but not insignificant (~3 %):

$$\mathsf{P}^*_{\xi\xi(vector+angle)} = \begin{bmatrix} 53.7336 & -107.0480 & 59.6645 \\ -107.0480 & 269.4744 & -145.0175 \\ 59.6645 & -145.0175 & 90.7662 \end{bmatrix} \times 10^{-9}. \tag{45}$$

Case 3: 1 Magnetometer Vector Observation + (4 GPS Lines-of-sight x 3 baselines). The situation is qualitatively different when only a single vector observation is available: the magnetic field. In this case, the complete attitude is not observable without including at least some of the angle observations. The attitude error covariance grows significantly and unevenly compared to the case when two vector observations are included. The largest increases related to rotation about the observed vector which is only detectable via relatively poor angle observations:*

$$\mathsf{P}^*_{\xi\xi(vector+angle)} = \begin{bmatrix} 335.8214 & 189.5209 & -613.4230 \\ 189.5209 & 661.4807 & -1329.7823 \\ -613.4230 & -1329.7823 & 4534.8546 \end{bmatrix} \times 10^{-9} \tag{46}$$

Case 4: 1 Magnetometer Vector Observation + (2 GPS Lines-of-sight x 3 baselines). The diminishing accuracy in measuring angle about the single observed vector becomes even more apparent if the number of angle observations is reduced. Consider the case when only 2 GPS satellites are available generating 6 additional angle observations. The attitude error covariance generally doubles in size with the largest increases still related to the rotation about the observed vector.

$$\mathsf{P}^*_{\xi\xi(vector+angle)} = \begin{bmatrix} 431.1612 & 393.1257 & -1292.1765 \\ 393.1257 & 1100.4411 & -2792.7159 \\ -1292.1765 & -2792.7159 & 9415.2490 \end{bmatrix} \times 10^{-9} \tag{47}$$

---

* At the selected moment in time, the simulated spacecraft position happens to be close to the North pole making the z-axis nearly opposite to the direction of the magnetic field.

Convergence Analysis

The Newton estimation sequence can be stopped when either the change in the estimate is sufficiently small or when the value of the cost function is not significantly reduced. By either measure in all simulated cases, a single iteration was sufficient to converge to the optimal estimate within essentially machine precision. A typical plot of the relative change in the cost function is shown in Figure 1. The initial value is the normalized unit value of the cost function. Its actual value is based on the initial attitude estimate obtained by the methods described in this paper. The subsequent relative changes in the cost function are due to the Newton iterations. For the first two simulated cases the change is too insignificant to be noticeable on the plot. For the other two cases, the relative change is only noticeable after the first Newton iteration after which the convergence is effectively achieved and the optimal estimate is found.

In all cases no appreciable difference is found between using the actual Hessian from Eq.(22) and its approximation from Eq.(28). The relative difference between the two matrices is on the order of 10^{-4}.

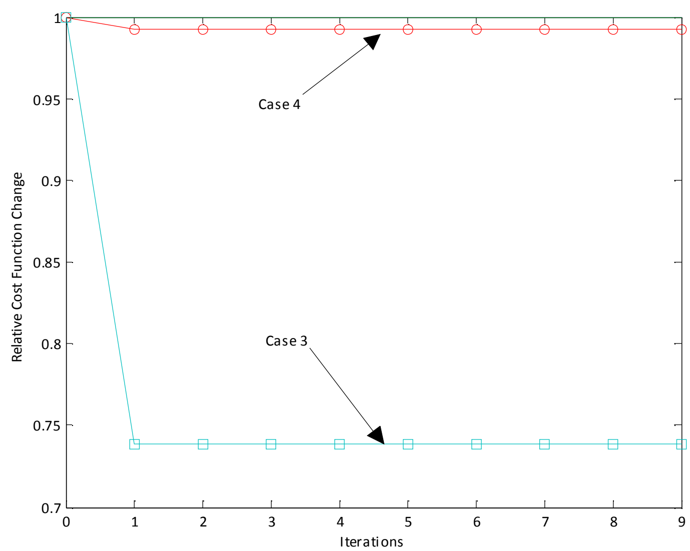

Figure 1. Typical Convergence as Indicated by Relative Change in Cost Function.

Statistical Consistency

The final numerical test involves running multiple trials for each case and verifying that the sampled attitude covariance agrees with the predicted analytical covariance and that the sampled attitude error agrees with the predicted covariance bounds. The sampled attitude error covariance is computed as

$$\mathsf{P}^{*}_{\xi\xi(sampled)} = \frac{1}{N}\sum_{n=1}^{N} \Delta\xi^{*}_{n} \Delta\xi^{*\mathrm{T}}_{n} \tag{48}$$

where N is the number of trials and where $\Delta\boldsymbol{\xi}_n^*$ denotes the attitude error vector for the n-th trial computed using the optimal estimate found once the Newton estimation sequence is converged. For a large number of trials, the difference $\Delta\mathsf{P}^*_{\xi\xi(ij)}$ between the corresponding elements of the sampled and predicted covariances should fall within the Gaussian zero-mean distribution with variance given by

$$\operatorname{var}\left\{\Delta\mathsf{P}^*_{\xi\xi(ij)}\right\} = \frac{1}{N}\left[\left(\mathsf{P}^*_{\xi\xi(ii)}\right)\left(\mathsf{P}^*_{\xi\xi(jj)}\right) + \left(\mathsf{P}^*_{\xi\xi(ij)}\right)^2\right]. \tag{49}$$

The statistical consistency of the attitude error can be examined by computing

$$\mu(\hat{\mathbf{q}}^*) = \Delta\boldsymbol{\xi}^{*\mathrm{T}}\left[\mathsf{P}^*_{\xi\xi}\right]^{-1}\Delta\boldsymbol{\xi}^* \tag{50}$$

which should fall approximately within χ^2-distribution with the mean value of 3 and the variance of 6. Consider all cases using $N = 100$ for each.

Case 1: All Vector Observations + (4 GPS Lines-of-sight x 3 baselines). The sampled covariance shown below along with its element-by-element 1σ bounds

$$\begin{aligned} \mathsf{P}^*_{\xi\xi(sampled)} &= \begin{bmatrix} 90.3440 & 19.9220 & -56.1536 \\ 19.9220 & 54.9010 & -21.4624 \\ -56.1536 & -21.4624 & 169.0648 \end{bmatrix} \times 10^{-12} \\ &\pm \begin{bmatrix} 12.8948 & 7.1406 & 13.3593 \\ 7.1406 & 7.7642 & 9.4713 \\ 13.3593 & 9.4713 & 23.0953 \end{bmatrix} \times 10^{-12} \end{aligned} \tag{51}$$

should be compared with the predicted covariance from Eq.(43) with which it agrees within less than 3σ. The mean and variance of $\mu(\hat{\mathbf{q}}^*)$ in this case are 2.9051 and 4.5154, respectively, not far from the expected χ^2 values. The plot of $\mu(\hat{\mathbf{q}}^*)$ in this case is shown in Figure 2.

Case 2: (1 Magnetometer + 1 Sun) Vector Observations + (4 GPS Lines-of-sight x 3 baselines). In this case, the sampled covariance shown below along with its element-by-element 1σ bounds

$$\begin{aligned} \mathsf{P}^*_{\xi\xi(sampled)} &= \begin{bmatrix} 66.7093 & -125.2200 & 70.5638 \\ -125.2200 & 292.6994 & -154.1668 \\ 70.5638 & -154.1668 & 95.9636 \end{bmatrix} \times 10^{-9} \\ &\pm \begin{bmatrix} 7.5991 & 16.1056 & 9.1853 \\ 16.1056 & 38.1094 & 21.3282 \\ 9.1853 & 21.3282 & 12.8363 \end{bmatrix} \times 10^{-9} \end{aligned} \tag{52}$$

should be compared with the predicted covariance from Eq.(45) with which it agrees within less than 3σ. The mean and variance of $\mu(\hat{\mathbf{q}}^*)$ in this case are 3.3841 and 6.3890, respectively, again not far from the expected χ^2 values. The plot of $\mu(\hat{\mathbf{q}}^*)$ in this case is shown in Figure 3.

Case 3: 1 Magnetometer Vector Observation + (4 GPS Lines-of-sight x 3 baselines). The sampled covariance shown below along with its element-by-element 1σ bounds

$$\begin{aligned} \mathsf{P}^*_{\xi\xi(sampled)} &= \begin{bmatrix} 268.2019 & 76.0540 & -338.3961 \\ 76.0540 & 466.4378 & -885.1999 \\ -338.3961 & -885.1999 & 3302.2010 \end{bmatrix} \times 10^{-9} \\ &\pm \begin{bmatrix} 47.4923 & 50.7994 & 137.8111 \\ 50.7994 & 93.5475 & 218.3584 \\ 137.8111 & 218.3584 & 641.3253 \end{bmatrix} \times 10^{-9} \end{aligned} \tag{53}$$

should be compared with the predicted covariance from Eq.(46) with which it once again agrees within less than 3σ. The mean and variance of $\mu(\hat{\mathbf{q}}^*)$ in this case are 2.5232 and 3.9944, respectively, still not far from the expected χ^2 values. In this case there are two plots shown in Figure 4: one is for $\mu(\hat{\mathbf{q}}_0)$ which is based on the initial (deterministic) estimate computed using a single vector and a single angle observation, the other is for $\mu(\hat{\mathbf{q}}^*)$ which, as in previous cases, is based on the optimal estimate. The two plots confirm that the optimal estimate is significantly more accurate.

Case 4: 1 Magnetometer Vector Observation + (2 GPS Lines-of-sight x 3 baselines). For the final case, the sampled covariance shown below along with its element-by-element 1σ bounds

$$\begin{aligned} \mathsf{P}^*_{\xi\xi(sampled)} &= \begin{bmatrix} 492.5848 & 474.0954 & -1561.8465 \\ 474.0954 & 1039.4064 & -2837.9679 \\ -1561.8465 & -2837.9679 & 10200.0530 \end{bmatrix} \times 10^{-9} \\ &\pm \begin{bmatrix} 60.9754 & 79.3105 & 239.3577 \\ 79.3105 & 155.6259 & 426.1477 \\ 239.3577 & 426.1477 & 1331.5172 \end{bmatrix} \times 10^{-9} \end{aligned} \tag{54}$$

should be compared with the predicted covariance from Eq.(47) with which it agrees within less than 3σ. The mean and variance of $\mu(\hat{\mathbf{q}}^*)$ in this case are 3.0150 and 7.8637, respectively, again not far from the expected χ^2 values. As in the previous case, there are two plots shown (Figure 5): one is for $\mu(\hat{\mathbf{q}}_0)$ and the other is for $\mu(\hat{\mathbf{q}}^*)$. The plots again confirm that the optimal estimate is more accurate than the initial estimate, although in this case by a smaller margin because it includes fewer angle observations (6) compared to the previous case (12).

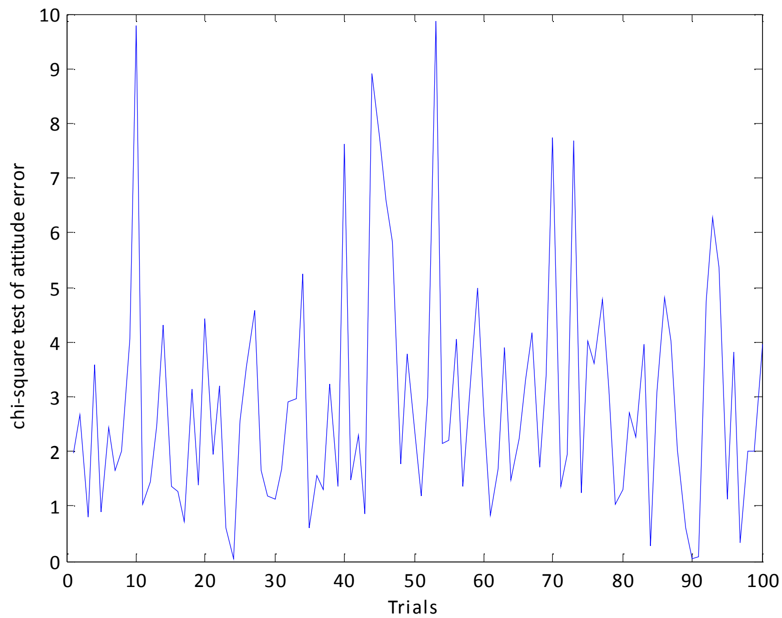

Figure 2. χ^2 **Test of Attitude Error for Case 1.**

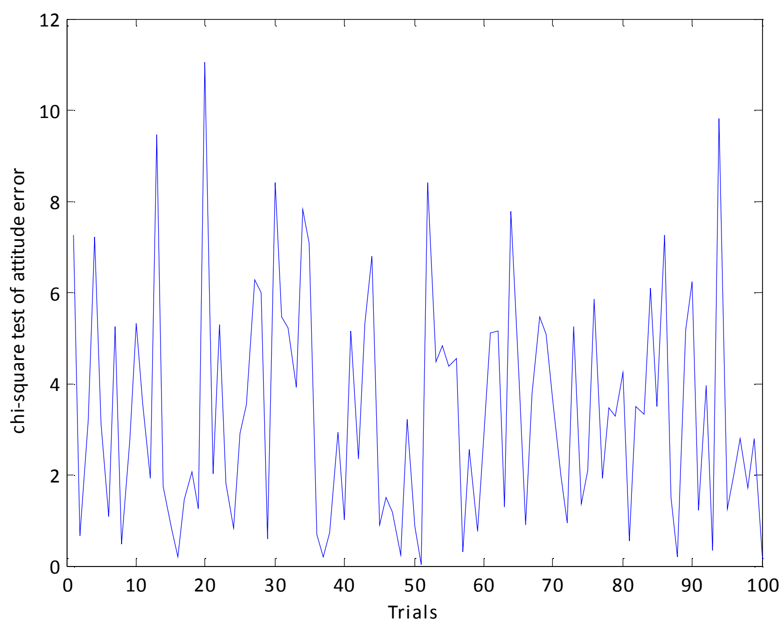

Figure 3. χ^2 **Test of Attitude Error for Case 2.**

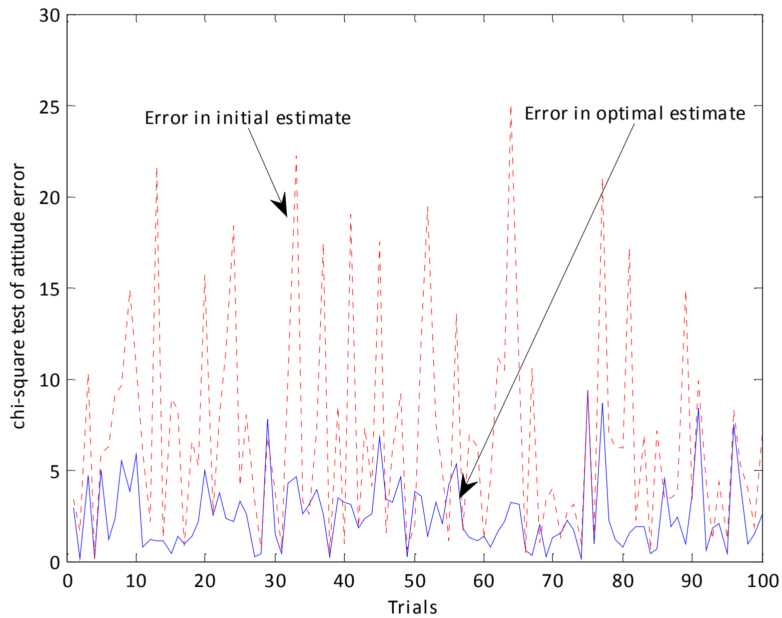

Figure 4. χ^2 **Test of Attitude Error for Case 3.**

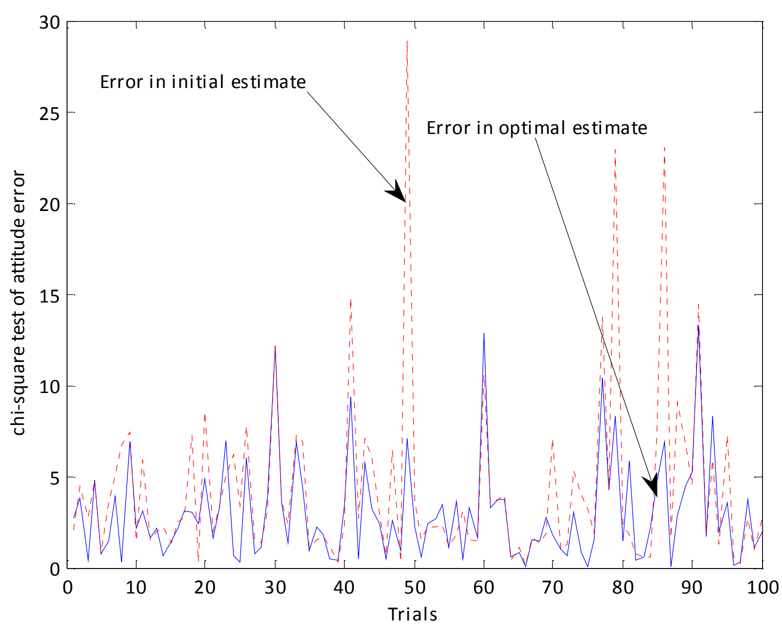

Figure 5. χ^2 **Test of Attitude Error for Case 4.**

CONCLUSION

The three-axis attitude determination can be performed using either vector or angle observations, or both. This paper demonstrates how the two types of observations can be effectively combined in order to generate the optimal attitude estimate. When there are multiple vector observations, addition of angle observations, such as GPS phase difference observations, which are typically less accurate, provides only marginal improvement to the estimate. The improvement is more noticeable when the accuracies of the two types of observations are comparable. In the special case when only a single vector observation is available, additional angle observations become indispensable because without them the complete three-axis attitude is unobservable. In this case, the initial attitude estimate can be significantly improved by employing the Newton estimation sequence that converges to the optimal attitude estimate which incorporates information from all available observations.

REFERENCES

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

[2] F. L. Markley, "Attitude Determination Using Vector Observations and the Singular Value Decomposition." *The Journal of the Astronautical Sciences*, Vol. 36, No. 3, 1988, pp. 245-258.

[3] F. L. Markley, "Attitude Determination Using Vector Observations: A Fast Optimal Matrix Algorithm." *The Journal of the Astronautical Sciences*, Vol. 41, No. 2, 1993, pp. 261-280.

[4] D. Mortari, "Second Estimator for the Optimal Quaternion." *Journal of Guidance, Control and Dynamics*, Vol. 23, No. 5, 2000, pp. 885-888.

[5] F. L. Markley and D. Mortari, "Quaternion Estimation Using Vector Observations." *The Journal of the Astronautical Sciences*, Vol. 48, Nos. 2 and 3, 2000, pp. 359-380.

[6] C. E. Cohen, "Attitude Determination Using GPS." Ph.D. Dissertation, Stanford Univ., 1992.

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

[8] C. E. Cohen, "Attitude Determination." *Global Positioning System: Theory and Applications*, Ch. 19, Vol. 2, ed. by B. W. Parkinson and J. J. Spilker, Progress in Astronautics and Aeronautics, Vol.164, AIAA, Washington, DC, 1996.

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

[10] I. Y. Bar-Itzhack, P. Y. Montgomery and J. C. Garrick, "Algorithms for Attitude Determination Using GPS." Proceedings of the AIAA Guidance, Navigation, and Control Conference, New Orleans, LA, Aug. 1997, AIAA 97-3616, pp. 841-851.

[11] J. L. Crassidis and F. L. Markley, "Predictive Filtering for Nonlinear Systems." *Journal of Guidance, Control and Dynamics*, Vol. 20, No. 3, May-June 1997, pp. 566-572.

[12] J. L. Crassidis, E. G. Lightsey and F. L. Markley, "Efficient and Optimal Attitude Determination Using Recursive Global Positioning System Signal Operations." *Journal of Guidance, Control and Dynamics*, Vol. 22, No. 2, 1999, pp. 193-201.

[13] S. Tanygin, "Angles Only Three-Axis Attitude Determination." AIAA/AAS Astrodynamics Specialist Conference, Toronto, Canada, August 2–5, 2010. Paper No. AIAA 2010–7826.

[14] S. Tanygin and M. D. Shuster, "Spin-Axis Attitude Estimation." *The Journal of the Astronautical Sciences*, Vol. 55, No.1, 2007, pp. 107-139.

[15] G. Wahba, "A Least Squares Estimate of Spacecraft Attitude," Problem 65-1, *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.

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

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