Q-METHOD EXTENDED KALMAN FILTER

Renato Zanetti[*], Thomas Ainscough[†], John Christian[‡] and Pol D. Spanos[§]

A new algorithm is proposed that smoothly integrates non-linear estimation of the attitude quaternion using Davenport's q-method and estimation of non-attitude states through an extended Kalman filter. The new method is compared to a similar existing algorithm showing its similarities and differences. The validity of the proposed approach is confirmed through numerical simulations.

INTRODUCTION

The well-known *Wahba Problem* [1] is a non-linear, weighted least-squares problem that seeks to obtain the optimal attitude matrix from a set of at least two independent vector measurements. The most common technique used to solved the Wahba problem is the so-called q-method, developed by Davenport and documented in [2]. The q-method rearranges the Wahba performance index into a quadratic performance index of the attitude quaternion, which is constrained to have unit norm. The extremals of this performance index are the eigenvalues of the Davenport matrix, and the optimal quaternion is the unit eigenvector corresponding to the largest eigenvalue.

A variety of numerical approaches exist for calculating the maximum eigenvalue and corresponding eigenvector of the Davenport matrix. For example, the QUEST algorithm [3] calculates the eigenvalue using a Newton-Raphson method and the eigenvector by factoring the quaternion as a vector of Rodrigues parameters. To avoid the singularity of the Rodrigues parameters the method of successive rotations is also introduced in [3]. Alternatively, ESOQ [4] avoids the singularity by computing the quaternion as a vector cross product in four dimensions. In a follow-on algorithm, ESOQ-2 [5], the Euler axis is computed as the null space of a 3×3 matrix that is derived from the Davenport matrix.

QUEST, ESOQ, and ESOQ-2 are numerical implementations of Davenport's q-method. Other numerical techniques exist that compute the attitude matrix directly rather than the quaternion. One such technique from Markley is based on the Singular Value Decomposition (SVD) [6]. Here, it should be noted that the original Wahba problem objective function is fundamentally just a special case of the *Orthogonal Procrustes Problem*, which has received a considerable amount of study since the 1950s [7].

---

[*]Senior Member of the Technical Staff, Vehicle Dynamics and Control, The Charles Stark Draper Laboratory, 17629 El Camino Real, Suite 470, Houston, Texas, 77058. rzanetti@draper.com

[†]Second Lieutenant. US Air Force. Rice University Graduate Student and Draper Laboratory Fellow, The Charles Stark Draper Laboratory, 17629 El Camino Real, Suite 470, Houston, Texas, 77058. tainscough@draper.com

[‡]Engineer, GNC Autonomous Flight Systems Branch, NASA Johnson Space Center, Houston, TX 77058.

[§]Professor of Mechanical Engineering and of Civil Engineering, Rice University, 6100 Main Street, Houston, TX 77005, 233 MEB , spanos@rice.edu

One of the reasons that the Wahba problem has received so much attention is that it provides a globally optimal solution and it does not make any linearization or small angle approximations. Conversely, the workhorse of aerospace estimation, the extended Kalman filter (EKF) [8] relies on linearization to obtain an estimate. The solution to the Wahba problem provides single point attitude estimates and requires all the measurements to be synchronized. The EKF and its attitude-specific extensions (most notably additive EKF [9] and multiplicative EKF [10]) in contrast are recursive estimators.

With this in mind, a number of algorithms have been developed to reformulate Davenport's solution into a recursive algorithm. Two of the first such methods are Filter QUEST [11] and REQUEST [12], which are both sub-optimal filters capable of estimating attitude (but not other states, such as biases). Later, Filter QUEST and REQUEST were shown to be two different formulations of mathematically equivalent filters [13]. Subsequently, the Optimal-REQUEST filter [14] addressed the sub-optimality of these filters, but was still not capable of estimating non-attitude states.

Markley [15] shows how to estimate not only attitude, but also other parameters such sensor biases from vector observations. Extended-QUEST also estimates attitude and non-attitude states [16]. This work introduces a novel EKF-based estimation algorithm that integrates the q-method to process attitude vector measurements. The existing algorithm that most closely resembles the present work is the Sequential Optimal Attitude Recursion (SOAR) filter by Christian and Lightsey [17]. The key difference is that SOAR uses the information formulation of the Kalman filter for the measurement update while the proposed method is a covariance formulation. This difference will usually require smaller matrix inversions when the size of the state vector is large. Another difference between the two methods is how the initial condition is introduced into the Wahba problem. This paper uses quaternion averaging [18], while SOAR uses the information matrix approach by Shuster [19]. In spite of these differences, the proposed q-method EKF (qEKF) and the SOAR filter are shown to be equivalent to second-order in the attitude update and first-order in the non-attitude state update. Hence qEKF and SOAR can be considered the covariance and information approaches to the solution of the same problem.

The proposed algorithm smoothly integrates the q-method into the EKF framework. Similar to the SOAR filter and Extended QUEST, the proposed algorithm processes the vector measurements first and the remaining quantities last. However, unlike Extended QUEST, both the SOAR filter and qEKF do not necessitate numerical iterations. Shuster [20] suggests that numerical solutions to the q-method such as QUEST could be used as a pre-processor for the EKF. The proposed algorithm takes this concept one step further by integrating the q-method into the EKF.

THE WAHBA PROBLEM

Re-written in terms of the inertial-to-body quaternion $\bar{\mathbf{q}}$, the Wahba problem consists of minimizing the performance index

$$\min_{\bar{\mathbf{q}}} \mathcal{J}\left(\hat{\bar{\mathbf{q}}}\right) = \frac{1}{2} \sum_{i=1}^{n} a_i \left\| \tilde{\mathbf{y}}_i - \mathbf{T}(\hat{\bar{\mathbf{q}}}) \tilde{\mathbf{n}}_i \right\|^2, \tag{1}$$

where $\tilde{\mathbf{y}}_i$ are vector observations and $\tilde{\mathbf{n}}_i$ are their representation in the reference frame.

In the absence of noise, the perfect measurement is simply given by

$$\mathbf{y}_i = \mathbf{T}(\bar{\mathbf{q}}) \, \mathbf{n}_i \tag{2}$$

In the presence of uncertainty, this becomes (omitting the dependency on $\bar{\mathbf{q}}$)

$$\tilde{\mathbf{y}}_i = \mathbf{T}\tilde{\mathbf{n}}_i + \delta\mathbf{y}_i \tag{3a}$$

$$\tilde{\mathbf{n}}_i = \mathbf{n}_i + \delta\mathbf{n}_i \tag{3b}$$

since $\|\tilde{\mathbf{y}}_i\| = \|\mathbf{y}_i\| = 1$ and $\|\tilde{\mathbf{n}}_i\| = \|\mathbf{n}_i\| = 1$, the following is also true to first order

$$\tilde{\mathbf{y}}_i^T \delta\mathbf{y}_i \approx \mathbf{y}_i^T \delta\mathbf{y}_i \approx 0 \qquad \tilde{\mathbf{n}}_i^T \delta\mathbf{n}_i \approx \mathbf{n}_i^T \delta\mathbf{n}_i \approx 0 \tag{4}$$

This leads directly to the QUEST measurement model [3] for a unit vector observation,

$$\mathbf{R_{nn}} = \sigma_{\mathbf{n}}^2 \left(\mathbf{I}_{3\times3} - \mathbf{nn}^T \right) \tag{5a}$$

$$\mathbf{R_{yy}} = \sigma_{\mathbf{y}}^2 \left(\mathbf{I}_{3\times3} - \mathbf{yy}^T \right) \tag{5b}$$

Substituting this result into Eq. (1) (and assuming that $\delta\mathbf{y}_i$ and $\delta\mathbf{n}_i$ are uncorrelated) shows that for $\hat{\bar{\mathbf{q}}}$ to be a maximum likelihood estimate of the attitude (to first order) the weights a_i should be the following

$$a_i \approx 1/\left(\sigma_{\mathbf{n}}^2 + \sigma_{\mathbf{y}}^2 \right) \tag{6}$$

Returning to Eq. (1), the goal is now to reformulate the problem in terms of the attitude quaternion. Begin by recalling that the coordinate transformation matrix written as a function of the quaternion is given by

$$\mathbf{T} = \mathbf{T}\left(\bar{\mathbf{q}}\right) = \mathbf{I}_{3\times3} - 2q_4\left[\mathbf{q}_v\times\right] + 2\left[\mathbf{q}_v\times\right]^2 \tag{7a}$$

$$= \left(q_4^2 - \mathbf{q}_v^T\mathbf{q}_v\right)\mathbf{I}_{3\times3} - 2q_4\left[\mathbf{q}_v\times\right] + 2\mathbf{q}_v\mathbf{q}_v^T \tag{7b}$$

The minimization of the Wahba performance index in Eq. (1) is now equivalent to the maximization of

$$\max_{\bar{\mathbf{q}}} \mathcal{J}^\star\left(\hat{\bar{\mathbf{q}}}\right) = \text{trace}\left[\mathbf{T}\left(\hat{\bar{\mathbf{q}}}\right)\mathbf{B}^\mathrm{T}\right] = \hat{\bar{\mathbf{q}}}^\mathrm{T}\mathbf{K}\hat{\bar{\mathbf{q}}}, \tag{8}$$

where the 4×4 Davenport matrix $\mathbf{K}$ is obtained as

$$\mathbf{B} = \sum_{i=1}^{n} a_i \tilde{\mathbf{y}}_i \tilde{\mathbf{n}}_i^\mathrm{T} \qquad\qquad \mathbf{z} = \sum_{i=1}^{n} a_i \left(\tilde{\mathbf{y}}_i \times \tilde{\mathbf{n}}_i\right)$$

$$\mathbf{S} = \mathbf{B} + \mathbf{B}^\mathrm{T} \qquad\qquad \sigma = \text{trace}\left(\mathbf{B}\right)$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{S} - \sigma\mathbf{I}_{3\times3} & \mathbf{z} \\ \mathbf{z}^\mathrm{T} & \sigma \end{bmatrix},$$

the optimal quaternion is the unit eigenvector of $\mathbf{K}$ associated with the maximum eigenvalue.

In this work $\mathbf{K}$ is slightly modified to perform covariance analysis. The performance index is equivalently rewritten as

$$\mathcal{J}^\star\left(\hat{\bar{\mathbf{q}}}\right) = \sigma + \hat{\bar{\mathbf{q}}}^\mathrm{T}\begin{bmatrix} \mathbf{M} & \mathbf{z} \\ \mathbf{z}^\mathrm{T} & 0 \end{bmatrix}\hat{\bar{\mathbf{q}}}, \tag{9}$$

hence the optimal quaternion is the unit eigenvector of the matrix in Eq. (9) corresponding to its maximum eigenvalue and

$$\mathbf{M} = \sum_{i=1}^{n} a_i \left([\tilde{\mathbf{y}}_i\times][\tilde{\mathbf{n}}_i\times] + [\tilde{\mathbf{n}}_i\times][\tilde{\mathbf{y}}_i\times]\right) = \mathbf{S} - 2\sigma \mathbf{I}_{3\times3}. \tag{10}$$

Recall that the perfect measurements $\mathbf{y}_i$ are defined as $\mathbf{y}_i = \mathbf{T}\mathbf{n}_i$ where $\mathbf{T}$ is the true attitude matrix and $\mathbf{n}_i$ are error-free reference vectors. By using $\mathbf{y}_i$ and $\mathbf{n}_i$ in place of $\tilde{\mathbf{y}}_i$ and $\tilde{\mathbf{n}}_i$ in the q-method the true quaternion $\bar{\mathbf{q}}$ is obtained. Matrix $\mathbf{B}_{true}$ is computed with the perfect values $\mathbf{y}_i$ and $\mathbf{n}_i$. When the vectors $\mathbf{y}_i$ and $\mathbf{T}(\bar{\mathbf{q}})\mathbf{n}_i$ are used as the inputs in the q-method the identity quaternion is obtained; with this approach we are estimating the deviation from the true body frame which is denoted as $\delta\bar{\mathbf{q}}^*$, the superscript "$*$" indicates the quaternion conjugate. Using $\mathbf{y}_i$ and $\mathbf{T}(\bar{\mathbf{q}})\mathbf{n}_i$ to calculate matrix $\mathbf{B}$ we obtain $\mathbf{T}(\bar{\mathbf{q}})\mathbf{B}_{true}$, hence the performance index is given by

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}^*) = \text{trace}\left[\mathbf{T}(\delta\bar{\mathbf{q}}^*)\mathbf{T}(\bar{\mathbf{q}})\mathbf{B}_{true}^{\mathrm{T}}\right], \tag{11}$$

Notice that the combination of having perfect measurements and replacing $\mathbf{n}_i$ with $\mathbf{T}(\bar{\mathbf{q}})\mathbf{n}_i$ results in $\mathbf{z} = \mathbf{0}$, which makes the performance index

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}^*) = \sigma + \delta\bar{\mathbf{q}}^{*\mathrm{T}} \begin{bmatrix} \mathbf{H}_{true} & \mathbf{0} \\ \mathbf{0}^{\mathbf{T}} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}^*, \tag{12}$$

where

$$\mathbf{H}_{true} = \sum_{i=1}^{n} a_i \left([\mathbf{y}_i\times][(\mathbf{T}\mathbf{n}_i)\times] + [(\mathbf{T}\mathbf{n}_i)\times][\mathbf{y}_i\times]\right)$$
$$= 2\sum_{i=1}^{n} a_i \left(\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}} - \mathbf{I}_{3\times3}\right) = 2\sum_{i=1}^{n} a_i [\mathbf{y}_i\times]^2. \tag{13}$$

$\mathbf{H}_{true}$ has non-positive eigenvalues, therefore the maximum eigenvalue of the modified Davenport matrix is zero and the optimal solution is the identity quaternion.

Re-introducing the error in the measurements and using $\tilde{\mathbf{y}}_i$ and $\mathbf{T}(\bar{\mathbf{q}})\tilde{\mathbf{n}}_i$ in the q-method the algorithm returns the estimation error since the performance index becomes

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}^*) = \text{trace}\left[\mathbf{T}(\delta\bar{\mathbf{q}}^*)\mathbf{T}(\bar{\mathbf{q}})\mathbf{B}^{\mathrm{T}}\right], \tag{14}$$

or equivalently (making use of the definition of the quaternion conjugate $\bar{\mathbf{q}}^* = \begin{bmatrix} -\mathbf{q}_v^{\mathrm{T}} & q_4 \end{bmatrix}^{\mathrm{T}}$)

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}) = \sigma + \delta\bar{\mathbf{q}}^{\mathrm{T}} \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}} & \delta\mathbf{z} \\ \delta\mathbf{z}^{\mathrm{T}} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}, \tag{15}$$

where

$$\mathbf{H}_{\boldsymbol{\theta}} = \sum_{i=1}^{n} a_i \left([\tilde{\mathbf{y}}_i\times][(\mathbf{T}\tilde{\mathbf{n}}_i)\times] + [(\mathbf{T}\tilde{\mathbf{n}}_i)\times][\tilde{\mathbf{y}}_i\times]\right) \tag{16}$$

$$\delta\mathbf{z} = -\sum_{i=1}^{n} a_i \left(\tilde{\mathbf{y}}_i \times \mathbf{T}\tilde{\mathbf{n}}_i\right) \tag{17}$$

In the absence of noise the optimal eigenvalue is equal to zero. With noise, the optimal eigenvalue is a small quantity $\delta\lambda$. The eigenvalue problem requires the following equation to be satisfied

$$\mathbf{H}_{\boldsymbol{\theta}}\,\delta\mathbf{q}_v + \delta q_4\,\delta\mathbf{z} = \delta\lambda\,\delta\mathbf{q}_v.$$

Making a first-order approximation of the quaternion and neglecting terms of order higher than one the estimation error is found to be

$$\delta\mathbf{q}_v = -\mathbf{H}_{\boldsymbol{\theta}}^{-1}\delta\mathbf{z}.$$

Therefore the covariance of the estimation error is given by

$$\mathbf{P}_{\boldsymbol{\theta\theta}} = 4\mathbf{H}_{\boldsymbol{\theta}}^{-1}\,\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\}(\mathbf{H}_{\boldsymbol{\theta}}^{-1})^{\mathrm{T}},$$

which is equivalent to the result by Shuster but derived differently. Clearly since the true attitude is unknown, $\mathbf{H}_{\boldsymbol{\theta}}$ needs to be evaluated at the estimated attitude; the added approximation is a second-order effect. To first-order we have that

$$\delta\mathbf{z} = -\sum_{i=1}^{n} a_i\left\{\mathbf{y}_i \times (\mathbf{T}\delta\mathbf{n}_i) + \delta\mathbf{y}_i \times (\mathbf{T}\mathbf{n}_i)\right\}. \tag{18}$$

therefore assuming each source of error is uncorrelated from the others

$$\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\} = \sum_{i=1}^{n} a_i^2\left\{[\mathbf{y}_i\times]\,\mathbf{T}\,\mathrm{E}\left\{\delta\mathbf{n}_i\delta\mathbf{n}_i^{\mathrm{T}}\right\}\mathbf{T}^{\mathrm{T}}\,[\mathbf{y}_i\times]^{\mathrm{T}} + [(\mathbf{T}\mathbf{n}_i)\times]\,\mathrm{E}\left\{\delta\mathbf{y}_i\delta\mathbf{y}_i^{\mathrm{T}}\right\}[(\mathbf{T}\mathbf{n}_i)\times]^{\mathrm{T}}\right\}.$$
$$\tag{19}$$

To calculate $\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\}$ the unknown quantities $\mathbf{y}_i$, $\mathbf{n}_i$, and $\mathbf{T}$ need to be replaced with the known quantities $\tilde{\mathbf{y}}_i$, $\tilde{\mathbf{n}}_i$, and $\hat{\mathbf{T}}$.

Therefore, for reasons that will become evident in the subsequent section, suppose one defines $\mathbf{R}$ as

$$\mathbf{R} = 4\,\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\} = 4\sum_{i=1}^{n} a_i^2\left\{[\tilde{\mathbf{y}}_i\times]\,\hat{\mathbf{T}}\mathbf{R}_{\mathbf{nn}}\hat{\mathbf{T}}^{\mathrm{T}}\,[\tilde{\mathbf{y}}_i\times]^{\mathrm{T}} + \left[(\hat{\mathbf{T}}\tilde{\mathbf{n}}_i)\times\right]\mathbf{R}_{\mathbf{yy}}\left[(\hat{\mathbf{T}}\tilde{\mathbf{n}}_i)\times\right]^{\mathrm{T}}\right\} \tag{20}$$

After substituting in Eq. (5b) and Eq. (5a), it follows that to first-order

$$\mathbf{R} = -2\mathbf{H}_{\boldsymbol{\theta}}. \tag{21}$$

This relationship will be critical in showing the equivalence between the qEKF and the SOAR filter.

INITIAL CONDITION

Shuster [19] shows one method to introduce initial conditions and quaternion measurements into the Wahba problem. Here a different approach is used. This section only treats the inclusion of the initial condition $\hat{\bar{\mathbf{q}}}_0$, which is equivalent to having a single quaternion measurement available on top of the vector measurements. The extension to multiple quaternion measurements is trivial; quaternion "measurements" are usually obtained from pre-processing vector measurements, in this work it is preferred to process the vector measurements directly.

The initial condition $\hat{\bar{\mathbf{q}}}_0$ is introduced with the quaternion averaging performance index [18]

$$\mathcal{J}(\hat{\bar{\mathbf{q}}}) = \hat{\bar{\mathbf{q}}}^\mathrm{T} \mathbf{\Xi}\left(\hat{\bar{\mathbf{q}}}_0\right) \mathbf{A}_0 \mathbf{\Xi}\left(\hat{\bar{\mathbf{q}}}_0\right)^\mathrm{T} \hat{\bar{\mathbf{q}}} + \frac{1}{2} \sum_{i=1}^{n} a_i \left\| \tilde{\mathbf{y}}_i - \mathbf{T}\left(\hat{\bar{\mathbf{q}}}\right) \tilde{\mathbf{n}}_i \right\|^2 \tag{22}$$

where

$$\mathbf{\Xi}(\bar{\mathbf{q}}) = \begin{bmatrix} q_4 \mathbf{I}_{3\times 3} + [\mathbf{q}_v \times] \\ -\mathbf{q}_v^\mathrm{T} \end{bmatrix}.$$

The initial estimation error is

$$\delta\bar{\mathbf{q}}_0 = \bar{\mathbf{q}} \otimes \hat{\bar{\mathbf{q}}}_0^*,$$

where the quaternion product $\otimes$ is defined such that the quaternions are multiplied in the same order as the attitude matrices.

To perform covariance analysis the same procedure as the previous section is used and instead of estimating the quaternion $\bar{\mathbf{q}}$, the deviation from it is sought. The performance index for the equivalent maximization problem then becomes

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}) = \delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} \mathbf{H_0} & \delta\mathbf{z_0} \\ \delta\mathbf{z_0}^\mathrm{T} & s_0 \end{bmatrix} \delta\bar{\mathbf{q}} + \sigma + \delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}} & \delta\mathbf{z} \\ \delta\mathbf{z}^\mathrm{T} & 0 \end{bmatrix} \delta\bar{\mathbf{q}} \tag{23}$$

$$\begin{aligned} \mathbf{H_0} &= -\mathbf{A}_0 - [\delta\mathbf{q}_{v0}\times]\mathbf{A}_0 + \mathbf{A}_0[\delta\mathbf{q}_{v0}\times] + [\delta\mathbf{q}_{v0}\times]\mathbf{A}_0[\delta\mathbf{q}_{v0}\times] \\ \delta\mathbf{z_0} &= \mathbf{A}_0\delta\mathbf{q}_{v0} + [\delta\mathbf{q}_{v0}\times]\mathbf{A}_0\delta\mathbf{q}_{v0} \\ s_0 &= -\delta\mathbf{q}_{v0}^\mathrm{T}\mathbf{A}_0\delta\mathbf{q}_{v0} \end{aligned}$$

making a first-order approximation and combining the matrices we obtain

$$\mathcal{J}^\star(\delta\bar{\mathbf{q}}) = \sigma + \delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} -\mathbf{A}_0 - [\delta\mathbf{q}_{v0}\times]\mathbf{A}_0 + \mathbf{A}_0[\delta\mathbf{q}_{v0}\times] + \mathbf{H}_{\boldsymbol{\theta}} & \mathbf{A}_0\delta\mathbf{q}_{v0} + \delta\mathbf{z} \\ \delta\mathbf{q}_{v0}^\mathrm{T}\mathbf{A}_0 + \delta\mathbf{z}^\mathrm{T} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}. \tag{24}$$

hence

$$\delta\mathbf{q}_v = -(-\mathbf{A}_0 - [\delta\mathbf{q}_{v0}\times]\mathbf{A}_0 + \mathbf{A}_0[\delta\mathbf{q}_{v0}\times] + \mathbf{H}_{\boldsymbol{\theta}})^{-1}(\mathbf{A}_0\delta\mathbf{q}_{v0} + \delta\mathbf{z}) \tag{25a}$$

$$\simeq -(-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}(\mathbf{A}_0\delta\mathbf{q}_{v0} + \delta\mathbf{z}), \tag{25b}$$

where the approximation holds to first-order. As was noted in Eq. (20), define

$$\mathbf{R} = 4\,\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^\mathrm{T}\right\} \tag{26}$$

assuming $\mathbf{q}_{v0}$ and $\delta\mathbf{z}$ are uncorrelated

$$\mathbf{P}_{\boldsymbol{\theta\theta}} = (-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}(\mathbf{A}_0\mathbf{P}_{\boldsymbol{\theta\theta}0}\mathbf{A}_0 + \mathbf{R})(-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^\mathrm{T} \tag{27a}$$

$$= \mathbf{K}_{\boldsymbol{\theta}}\left(\mathbf{A}_0\mathbf{P}_{\boldsymbol{\theta\theta}0}\mathbf{A}_0 + \mathbf{R}\right)\mathbf{K}_{\boldsymbol{\theta}}^\mathrm{T} \tag{27b}$$

where

$$\mathbf{K}_{\boldsymbol{\theta}} = (-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}.$$

This covariance update equation is now rewritten in the familiar Joseph form [21]

$$(-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}(-\mathbf{A}_0) = (-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}(-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}} - \mathbf{H}_{\boldsymbol{\theta}})$$

$$= \mathbf{I} - (-\mathbf{A}_0 + \mathbf{H}_{\boldsymbol{\theta}})^{-1}\mathbf{H}_{\boldsymbol{\theta}}$$

$$= \mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}$$

applying this result into Eq. (25b)

$$\delta \mathbf{q}_v = (\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}} \mathbf{H}_{\boldsymbol{\theta}}) \delta \mathbf{q}_{v0} - \mathbf{K}_{\boldsymbol{\theta}} \delta \mathbf{z}.$$

it then follows that

$$\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}} = (\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}} \mathbf{H}_{\boldsymbol{\theta}}) \, \mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}0} \, (\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}} \mathbf{H}_{\boldsymbol{\theta}})^{\mathrm{T}} + \mathbf{K}_{\boldsymbol{\theta}} \mathbf{R} \mathbf{K}_{\boldsymbol{\theta}}^{\mathrm{T}} \tag{28}$$

which is the Joseph formula. The initial weight is chosen as $\mathbf{A}_0 = 2\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}0}^{-1}$ because the first term of Eq. (22) does not contain the factor 1/2 and $\delta\hat{\boldsymbol{\theta}} \simeq 2\boldsymbol{\Xi} \left(\hat{\bar{\mathbf{q}}}_0 \right)^{\mathrm{T}} \hat{\mathbf{q}}$.

THE Q-METHOD EXTENDED KALMAN FILTER

Let's start with linear measurements, the extension to the nonlinear case can be readily obtained using standard extended Kalman filter techniques. Let $\mathbf{y}$ be a set of measurements of a state vector $\mathbf{x}$ corrupted by zero mean noise $\boldsymbol{\eta}$ with covariance $\mathbf{R}$

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}, \tag{29}$$

where $\mathbf{H}$ is the measurement mapping (or sensitivity) matrix. Let $\hat{\mathbf{x}}^-$ be an unbiased estimate of $\mathbf{x}$ with corresponding estimation error covariance given by $\mathbf{P}^-$. The *a priori* estimation error is given by

$$\mathbf{e}^- = \mathbf{x} - \hat{\mathbf{x}}^-. \tag{30}$$

The unbiased linear update based upon $\hat{\mathbf{x}}^-$ and $\mathbf{y}$ produces the *a posteriori* estimate given by

$$\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{K} \left(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^- \right) = \hat{\mathbf{x}}^- + \mathbf{K}\boldsymbol{\epsilon}, \tag{31}$$

where $\boldsymbol{\epsilon}$ is called the measurement residual and $\mathbf{K}$ is some deterministic matrix of appropriate dimensions to be determined. The *a posteriori* estimation error is expressed as

$$\mathbf{e}^+ = \mathbf{x} - \hat{\mathbf{x}}^+ = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{e}^- - \mathbf{K}\boldsymbol{\eta}. \tag{32}$$

Assuming that the measurement error $\boldsymbol{\eta}$ and the *a priori* estimation error, $\mathbf{e}^-$ are uncorrelated and each are zero mean, we find that the *a posteriori* estimation error covariance is given by the Joseph formula [21]

$$\mathbf{P}^+ = \mathrm{E} \left\{ \left(\mathbf{x} - \hat{\mathbf{x}}^+ \right) \left(\mathbf{x} - \hat{\mathbf{x}}^+ \right)^{\mathrm{T}} \right\} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^-(\mathbf{I} - \mathbf{K}\mathbf{H})^{\mathrm{T}} + \mathbf{K}\mathbf{R}\mathbf{K}^{\mathrm{T}}, \tag{33}$$

where $\mathbf{I}$ is the identity matrix of appropriate dimension. Notice that no assumptions have been made as to the choice of $\mathbf{K}$ and the Joseph update equation is valid for all $\mathbf{K}$.

Suppose now that we partition $\mathbf{x}$ into 3 attitude states, $\boldsymbol{\theta}$, and $n - 3$ other states, $\mathbf{s}$ as

$$\mathbf{x} = \begin{bmatrix} \mathbf{s} \\ \boldsymbol{\theta} \end{bmatrix}. \tag{34}$$

Partitioning the relevant matrices accordingly it is obtained that

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{\mathbf{s}\mathbf{s}} & \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}} \\ \mathbf{P}_{\boldsymbol{\theta}\mathbf{s}} & \mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix} \tag{35a}$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\mathbf{s}} & \mathbf{H}_{\boldsymbol{\theta}} \end{bmatrix} \tag{35b}$$

$$\mathbf{K}_{opt} = \begin{bmatrix} \mathbf{K}_{\mathbf{s},opt} \\ \mathbf{K}_{\boldsymbol{\theta},opt} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\mathbf{s}\mathbf{s}}^- \mathbf{H}_{\mathbf{s}}^{\mathrm{T}} + \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^- \mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}} \\ \mathbf{P}_{\boldsymbol{\theta}\mathbf{s}}^- \mathbf{H}_{\mathbf{s}}^{\mathrm{T}} + \mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^- \mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}} \end{bmatrix} \mathbf{W}^{-1} \tag{35c}$$

where $\mathbf{W}$ is the residuals covariance matrix

$$
\begin{aligned}
\mathbf{W} &= \mathbf{H}\mathbf{P}^-\mathbf{H}^\mathrm{T} + \mathbf{R} \\
&= \mathbf{H_s}\mathbf{P}^-_{\mathbf{ss}}\mathbf{H}^\mathrm{T}_\mathbf{s} + \mathbf{H_s}\mathbf{P}^-_{\mathbf{s}\theta}\mathbf{H}^\mathrm{T}_\theta + \mathbf{H}_\theta\mathbf{P}^-_{\theta\mathbf{s}}\mathbf{H}^\mathrm{T}_\mathbf{s} + \mathbf{H}_\theta\mathbf{P}^-_{\theta\theta}\mathbf{H}^\mathrm{T}_\theta + \mathbf{R}.
\end{aligned}
\tag{36}
$$

The updated portions of the covariance are

$$
\mathbf{P}^+_{\mathbf{ss}} = \mathbf{P}^-_{\mathbf{ss}} - \mathbf{K_s}\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{ss}}\\\mathbf{P}^-_{\theta\mathbf{s}}\end{bmatrix} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{ss}}\\\mathbf{P}^-_{\theta\mathbf{s}}\end{bmatrix}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T}_\mathbf{s} + \mathbf{K_s}\mathbf{W}\mathbf{K}^\mathrm{T}_\mathbf{s}
\tag{37a}
$$

$$
\mathbf{P}^+_{\mathbf{s}\theta} = \mathbf{P}^-_{\mathbf{s}\theta} - \mathbf{K_s}\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{ss}}\\\mathbf{P}^-_{\theta\mathbf{s}}\end{bmatrix}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T}_\theta + \mathbf{K_s}\mathbf{W}\mathbf{K}^\mathrm{T}_\theta
\tag{37b}
$$

$$
\mathbf{P}^+_{\theta\mathbf{s}} = \mathbf{P}^-_{\theta\mathbf{s}} - \mathbf{K}_\theta\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{ss}}\\\mathbf{P}^-_{\theta\mathbf{s}}\end{bmatrix} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix}^\mathrm{T}\mathbf{H}\mathbf{K}^\mathrm{T}_\mathbf{s} + \mathbf{K}_\theta\mathbf{W}\mathbf{K}^\mathrm{T}_\mathbf{s}
\tag{37c}
$$

$$
\mathbf{P}^+_{\theta\theta} = \mathbf{P}^-_{\theta\theta} - \mathbf{K}_\theta\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T}_\theta + \mathbf{K}_\theta\mathbf{W}\mathbf{K}^\mathrm{T}_\theta
\tag{37d}
$$

These equations are derived from the Joseph formula and are therefore valid for any choice $\mathbf{K_s}$ and $\mathbf{K}_\theta$.

We now choose the $\mathbf{K_s}$ and $\mathbf{K}_\theta$ carefully such that the gain for the states is the optimal value $\mathbf{K}_{\mathbf{s},opt}$ from Eq. (35c) and allowing $\mathbf{K}_\theta$ to be (as yet) unspecified. Thus, substituting for the value of $\mathbf{K}_{\mathbf{s},opt}$ for the three components in Eqs. (37a)–(37c), it becomes

$$
\mathbf{P}^+ = \begin{bmatrix} \mathbf{P}^-_{\mathbf{ss}} - \mathbf{K}_{\mathbf{s},opt}\mathbf{W}\mathbf{K}^\mathrm{T}_{\mathbf{s},opt} & \mathbf{P}^-_{\mathbf{s}\theta} - \mathbf{K}_{\mathbf{s},opt}\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix} \\ \mathbf{P}^-_{\theta\mathbf{s}} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T}_{\mathbf{s},opt} & \mathbf{P}^-_{\theta\theta} - \mathbf{K}_\theta\mathbf{H}\begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix} - \begin{bmatrix}\mathbf{P}^-_{\mathbf{s}\theta}\\\mathbf{P}^-_{\theta\theta}\end{bmatrix}^\mathrm{T}\mathbf{H}^\mathrm{T}\mathbf{K}^\mathrm{T}_\theta + \mathbf{K}_\theta\mathbf{W}\mathbf{K}^\mathrm{T}_\theta \end{bmatrix}
\tag{38}
$$

This equation is valid for any value of $\mathbf{K}_\theta$. Notice that there is no $\mathbf{K}_\theta$ in the cross-covariance between $\mathbf{s}$ and θ. Therefore, what is remarkable about this equation is that once the optimal $\mathbf{K}_{\mathbf{s},opt}$ is chosen, the cross-covariance between $\mathbf{s}$ and θ is independent of the choice of $\mathbf{K}_\theta$. This property is fundamental to the development of the q-method EKF because it allows us to choose $\mathbf{K}_\theta$ as defined in the previous section

$$
\mathbf{K}_\theta = (-\mathbf{A}_0 + \mathbf{H}_\theta)^{-1}.
$$

The measurement residual is obtained as

$$
\epsilon = 2\mathbf{K}^{-1}_\theta\mathrm{vec}\left(\hat{\mathbf{q}}^+ \otimes (\hat{\mathbf{q}}^-)^*\right)
$$

where $\mathrm{vec}(\bar{\mathbf{q}})$ is the function that returns the vector part of the quaternion.

It is assumed that the vector measurements are only functions of the vehicle attitude and do not depend on any other states, i.e. $\mathbf{H_s} = \mathbf{0}$. In summary the proposed algorithm has a propagation phase identical to that of the multiplicative EKF and an update phase as follows

1. Calculate the Davenport matrix $\mathbf{K}$ associated with all attitude vector measurements

2. Calculate $\mathbf{A}_0 = 2(\mathbf{P}_{\boldsymbol{\theta\theta}}^-)^{-1}$

3. Calculate the updated attitude quaternion as the unit eigenvector associated with the maximum eigenvalue of

$$\mathbf{K}_{aug} = -\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^-\right)\mathbf{A}_0\,\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}} + \mathbf{K}$$

4. Calculate $\mathbf{H}_{\boldsymbol{\theta}}$, $\mathbf{R}$, $\mathbf{K}_{\boldsymbol{\theta}}$ using Eqs. (16), (26), (19), and (28)

5. Update the non attitude states as

$$\mathbf{s}^+ = \mathbf{s}^- + \mathbf{K}_{\mathbf{s},opt}\,\boldsymbol{\epsilon}$$
$$\mathbf{K}_{\mathbf{s},opt} = \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}}(\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta\theta}}^-\mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}} + \mathbf{R})^{-1}$$
$$\boldsymbol{\epsilon} = 2\mathbf{K}_{\boldsymbol{\theta}}^{-1}\mathrm{vec}\left(\hat{\bar{\mathbf{q}}}^+ \otimes (\hat{\bar{\mathbf{q}}}^-)^*\right)$$

6. Update the total covariance using Eq. (38) and $\mathbf{H} = \begin{bmatrix}\mathbf{O} & \mathbf{H}_{\boldsymbol{\theta}}\end{bmatrix}$.

7. Process any remaining measurements using the standard MEKF algorithm

COMPARISON WITH THE SOAR FILTER

This section demonstrates the equivalence of the qEKF and the SOAR filter. It begins by making a key observation about the attitude profile matrix, and then proceeds to compare the attitude update and the non-attitude update.

Observations on Computation of the Attitude Profile Matrix

Begin by recalling that the Wahba Problem objective function given in Eq. (8) is the negative log-likelihood function when a_i are chosen as shown in Eq. (6). The attitude may be expanded about the estimate using a Taylor Series expansion truncated to second-order

$$J\left(\boldsymbol{\delta\theta}\right) = -\mathrm{trace}\left[\left(\mathbf{I}_{3\times3} + [-\boldsymbol{\delta\theta}\times] + \frac{1}{2}\left[-\boldsymbol{\delta\theta}\times\right]^2\right)\mathbf{T}\mathbf{B}^{\mathrm{T}}\right] \tag{39}$$

Under mild conditions, the Fisher information matrix, $\mathcal{F}_{\boldsymbol{\theta\theta}}$ is the expected value of the second-order derivative of the negative log-likelihood function. Recall from the Cramèr-Rao inequality that the attitude covariance, $\mathbf{P}_{\boldsymbol{\theta\theta}}$, is related to the Fisher information matrix by [22]

$$\mathbf{P}_{\boldsymbol{\theta\theta}}^{-1} \leq \mathcal{F}_{\boldsymbol{\theta\theta}} = E\left[\frac{\partial^2 J(\boldsymbol{\delta\theta})}{\partial\boldsymbol{\delta\theta}\,\partial\boldsymbol{\delta\theta}}\right] \tag{40}$$

and that $\mathcal{F}_{\boldsymbol{\theta\theta}}$ approaches $\mathbf{P}_{\boldsymbol{\theta\theta}}^{-1}$ as the number of measurements become large.

Because Eq. (40) requires the second derivative of $J\left(\boldsymbol{\delta\theta}\right)$ with respect to $\boldsymbol{\delta\theta}$, terms in $J\left(\boldsymbol{\delta\theta}\right)$ that are independent of $\boldsymbol{\delta\theta}$ or linear in $\boldsymbol{\delta\theta}$ are unimportant in the computation of $\mathcal{F}_{\boldsymbol{\theta\theta}}$. Therefore,

$$\mathcal{F}_{\boldsymbol{\theta\theta}} = E\left[\frac{\partial^2 J(\boldsymbol{\delta\theta})}{\partial\boldsymbol{\delta\theta}\,\partial\boldsymbol{\delta\theta}}\right] = E\left[\frac{\partial^2}{\partial\boldsymbol{\delta\theta}\,\partial\boldsymbol{\delta\theta}}\left(-\mathrm{trace}\left[\frac{1}{2}\left[-\boldsymbol{\delta\theta}\times\right]^2\mathbf{T}\mathbf{B}^{\mathrm{T}}\right]\right)\right] \tag{41}$$

To compact notation, define the matrix $\mathbf{V} = \mathbf{T}\mathbf{B}^{\mathrm{T}}$,

$$\mathcal{F}_{\theta\theta} = \frac{1}{2}E\left[\frac{\partial^2}{\partial\boldsymbol{\delta\theta}\,\partial\boldsymbol{\delta\theta}}\left(-\mathrm{trace}\left[[-\boldsymbol{\delta\theta}\times]^2\,\mathbf{V}\right]\right)\right] \tag{42}$$

Now, making the observation that,

$$[-\boldsymbol{\delta\theta}\times]^2 = \boldsymbol{\delta\theta}\,\boldsymbol{\delta\theta}^{\mathrm{T}} - \boldsymbol{\delta\theta}^{\mathrm{T}}\boldsymbol{\delta\theta}\mathbf{I}_{3\times3} \tag{43}$$

one may directly rewrite Eq. (42) as

$$\mathcal{F}_{\theta\theta} = -\frac{1}{2}E\left[\frac{\partial^2}{\partial\boldsymbol{\delta\theta}\partial\boldsymbol{\delta\theta}}\left(\mathrm{trace}\left[\boldsymbol{\delta\theta}\,\boldsymbol{\delta\theta}^{\mathrm{T}}\mathbf{V}\right] - \boldsymbol{\delta\theta}^{\mathrm{T}}\boldsymbol{\delta\theta}\,\mathrm{trace}\left[\mathbf{V}\right]\right)\right] \tag{44}$$

Taking advantage of the cyclic properties of the trace operator,

$$\mathcal{F}_{\theta\theta} = -\frac{1}{2}E\left[\frac{\partial^2}{\partial\boldsymbol{\delta\theta}\partial\boldsymbol{\delta\theta}}\left(\boldsymbol{\delta\theta}^{\mathrm{T}}\mathbf{V}\boldsymbol{\delta\theta} - \mathrm{trace}\left[\mathbf{V}\right]\boldsymbol{\delta\theta}^{\mathrm{T}}\boldsymbol{\delta\theta}\right)\right] \tag{45}$$

Straightforward differentiation will yield,

$$\mathcal{F}_{\theta\theta} = \mathrm{trace}\left[\mathbf{V}\right]\mathbf{I}_{3\times3} - \frac{1}{2}\left(\mathbf{V} + \mathbf{V}^{\mathrm{T}}\right) \tag{46}$$

In the presence of perfect measurements one may note that $\mathbf{V} = \mathbf{V}^{\mathrm{T}}$, and it is under these conditions that one arrives at the result presented by Shuster in [19]. In general, however, this is not the case and using Shuster's formulation will result in nonsymmetric information and covariance matrices (clearly not correct!). Fortunately, this assumption is not necessary and one may correctly compute the Fisher information matrix in the presence of noise as,

$$\mathbf{P}_{\theta\theta}^{-1} \approx \mathcal{F}_{\theta\theta} = \mathrm{trace}\left[\mathbf{T}\mathbf{B}^{\mathrm{T}}\right]\mathbf{I}_{3\times3} - \frac{1}{2}\left(\mathbf{T}\mathbf{B}^{\mathrm{T}} + \mathbf{B}\mathbf{T}^{\mathrm{T}}\right) \tag{47}$$

The solution provided by Shuster in Ref. [19] to compute $\mathbf{B}$ from $\mathcal{F}_{\theta\theta}$ and $\mathbf{T}$ is still valid. By taking the trace of Eq. (47), note that

$$\mathrm{trace}\left[\mathcal{F}_{\theta\theta}\right] = 3\mathrm{trace}\left[\mathbf{T}\mathbf{B}^{\mathrm{T}}\right] - \mathrm{trace}\left[\mathbf{T}\mathbf{B}^{\mathrm{T}}\right] = 2\mathrm{trace}\left[\mathbf{T}\mathbf{B}^{\mathrm{T}}\right] \tag{48}$$

Substituting this back into Eq. (47),

$$\mathbf{T}\mathbf{B}^{\mathrm{T}} + \mathbf{B}\mathbf{T}^{\mathrm{T}} = \mathrm{trace}\left[\mathcal{F}_{\theta\theta}\right]\mathbf{I}_{3\times3} - 2\mathcal{F}_{\theta\theta} \tag{49}$$

Now, it is straightforward to verify that the following solution originally given by Shuster in [19] is also a solution to this equation,

$$\mathbf{B} = \left[\frac{1}{2}\mathrm{trace}\left[\mathcal{F}_{\theta\theta}\right]\mathbf{I}_{3\times3} - \mathcal{F}_{\theta\theta}\right]\mathbf{T} \tag{50}$$

Equivalence of the Attitude Update

Recall from [17] that the SOAR filter includes the a-priori attitude information through the following term in the objective function

$$-\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\mathbf{K}^{-}\hat{\bar{\mathbf{q}}} = -\mathrm{trace}\left[\mathbf{T}\left(\mathbf{B}^{-}\right)^{\mathrm{T}}\right] \tag{51}$$

In [17] it is also shown that, after a second-order expansion of the matrix exponential of $[-\boldsymbol{\delta\theta}\times]$ about the *a priori* attitude, this objective function may be rewritten as

$$-\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\mathbf{K}^{-}\hat{\bar{\mathbf{q}}} = -\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\mathbf{K}^{-}\hat{\bar{\mathbf{q}}}^{-} + \frac{1}{2}\boldsymbol{\delta\theta}^{\mathrm{T}}\mathcal{F}_{\theta\theta}\boldsymbol{\delta\theta} \tag{52}$$

The first term is a constant (not dependent on the *a posteriori* attitude) and disappears when the first differentials are taken to compute the optimal attitude.

It is now straightforward to show that the *a priori* attitude term introduced in Eq. (22) is equivalent to $1/2\boldsymbol{\delta\theta}^{\mathrm{T}}\mathcal{F}_{\theta\theta}\boldsymbol{\delta\theta}$ to second-order. Thus, both the qEKF and the SOAR filters can be shown to include the *a priori* attitude information in an equivalent manner to second-order.

To show this, begin by noting that

$$\boldsymbol{\delta}\mathbf{q}_{v} = \boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} = \sin\left(\frac{\boldsymbol{\delta\theta}}{2}\right) \tag{53}$$

Taking the Taylor Series expansion of $\sin\left(\boldsymbol{\delta\theta}/2\right)$, one may show that to second-order

$$\boldsymbol{\delta}\mathbf{q}_{v} = \sin\left(\frac{\boldsymbol{\delta\theta}}{2}\right) = \frac{\boldsymbol{\delta\theta}}{2} - \frac{1}{3}\left(\frac{\boldsymbol{\delta\theta}}{2}\right)^{3} + \frac{1}{5}\left(\frac{\boldsymbol{\delta\theta}}{2}\right)^{5}\ldots \approx \frac{\boldsymbol{\delta\theta}}{2} \tag{54}$$

Therefore, the first term in Eq. (22) may be rewritten as,

$$\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)\mathbf{A}_{0}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} \approx \frac{1}{4}\boldsymbol{\delta\theta}^{\mathrm{T}}\mathbf{A}_{0}\boldsymbol{\delta\theta} \tag{55}$$

Noting from before that $\mathbf{A}_{0}$ was chosen as $\mathbf{A}_{0} = 2\left(\mathbf{P}_{\theta\theta}^{-}\right)^{-1} \approx 2\mathcal{F}_{\theta\theta}$, this directly yields

$$\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)\mathbf{A}_{0}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} \approx \frac{1}{2}\boldsymbol{\delta\theta}^{\mathrm{T}}\mathcal{F}_{\theta\theta}\boldsymbol{\delta\theta} \tag{56}$$

Therefore, the *a priori* attitude additions to the objective function for both SOAR and the qEKF are equivalent to second-order.

Equivalence of the Non-Attitude Update

Partition the Fisher information matrix of the full covariance as

$$\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathbf{ss}} & \mathbf{P}_{\mathbf{s}\theta} \\ \mathbf{P}_{\theta\mathbf{s}} & \mathbf{P}_{\theta\theta} \end{bmatrix}^{-1} = \mathbf{F}_{\mathbf{xx}} = \begin{bmatrix} \mathbf{F}_{\mathbf{ss}} & \mathbf{F}_{\mathbf{s}\theta} \\ \mathbf{F}_{\theta\mathbf{s}} & \mathbf{F}_{\theta\theta} \end{bmatrix} \tag{57}$$

The relation between $\mathcal{F}_{\theta\theta}$ and $\mathbf{F}_{\theta\theta}$ stems from the inversion of a partitioned matrix,

$$\mathbf{F}_{\theta\theta} = \mathbf{P}_{\theta\theta}^{-1} + \mathbf{F}_{\theta s}\mathbf{F}_{ss}^{-1}\mathbf{F}_{s\theta} = \mathcal{F}_{\theta\theta} + \mathbf{F}_{\theta s}\mathbf{F}_{ss}^{-1}\mathbf{F}_{s\theta} \tag{58}$$

11

Given this definition, recall from Ref. [17] that the optimal update of the non-attitude states in the SOAR filter is given by

$$\mathbf{s}^+ = \mathbf{s}^- - 2\left(\mathbf{F}_{ss}^-\right)^{-1}\mathbf{F}_{s\theta}^- \, \Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^+ \tag{59a}$$

$$\mathbf{s}^+ \approx \mathbf{s}^- - \left(\mathbf{F}_{ss}^-\right)^{-1}\mathbf{F}_{s\theta}^- \, \delta\boldsymbol{\theta} \tag{59b}$$

The objective is now to show that this is equivalent to the qEFK non-attitude update. Begin by recalling that

$$\mathbf{s}^+ = \mathbf{s}^- + \mathbf{K}_{\mathbf{s},opt}\,\boldsymbol{\epsilon}$$

$$\mathbf{K}_{\mathbf{s},opt} = \mathbf{P}_{s\theta}^-\mathbf{H}_{\theta}^{\mathrm{T}}(\mathbf{H}_{\theta}\mathbf{P}_{\theta\theta}^-\mathbf{H}_{\theta}^{\mathrm{T}} + \mathbf{R})^{-1}$$

$$\mathbf{K}_{\theta} = \left(-2\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} + \mathbf{H}_{\theta}\right)^{-1}$$

$$\boldsymbol{\epsilon} = 2\mathbf{K}_{\theta}^{-1}\mathrm{vec}\left(\hat{\bar{\mathbf{q}}}^+ \otimes (\hat{\bar{\mathbf{q}}}^-)^*\right)$$

and, therefore,

$$\mathbf{s}^+ = \mathbf{s}^- + 2\mathbf{K}_{\mathbf{s},opt}\mathbf{K}_{\theta}^{-1}\Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^+ \tag{60}$$

Looking first at the $\mathbf{K}_{\mathbf{s},opt}$ and recalling the Woodbury identity,

$$\mathbf{K}_{\mathbf{s},opt} = \mathbf{P}_{s\theta}^-\mathbf{H}_{\theta}\left[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{H}_{\theta}\left(\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} + \mathbf{H}_{\theta}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}_{\theta}\right)^{-1}\mathbf{H}_{\theta}^{\mathrm{T}}\mathbf{R}^{-1}\right] \tag{61}$$

Further recognizing from Eq. (21) that $\mathbf{R} = -2\mathbf{H}_{\theta}$ and that $\mathbf{H}_{\theta}$ is symmetric,

$$\mathbf{K}_{\mathbf{s},opt} = \mathbf{P}_{s\theta}^-\mathbf{H}_{\theta}\left[-\frac{1}{2}\mathbf{H}_{\theta}^{-1} + \frac{1}{4}\left(\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} - \frac{1}{2}\mathbf{H}_{\theta}\right)^{-1}\right] \tag{62}$$

And with a bit more simplification,

$$\mathbf{K}_{\mathbf{s},opt} = -\frac{1}{2}\mathbf{P}_{s\theta}^-\left[\mathbf{I}_{3\times3} - \mathbf{H}_{\theta}\left(-2\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} + \mathbf{H}_{\theta}\right)^{-1}\right] = -\frac{1}{2}\mathbf{P}_{s\theta}^-\left[\mathbf{I}_{3\times3} - \mathbf{H}_{\theta}\mathbf{K}_{\theta}\right] \tag{63}$$

Substituting into Eq. (60) for $\mathbf{K}_{\mathbf{s},opt}$,

$$\mathbf{s}^+ = \mathbf{s}^- - \mathbf{P}_{s\theta}^-\left[\mathbf{I}_{3\times3} - \mathbf{H}_{\theta}\mathbf{K}_{\theta}\right]\mathbf{K}_{\theta}^{-1}\,\Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^+ \tag{64}$$

Expanding terms,

$$\mathbf{s}^+ = \mathbf{s}^- - \mathbf{P}_{s\theta}^-\left[\mathbf{K}_{\theta}^{-1} - \mathbf{H}_{\theta}\right]\Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^+ \tag{65a}$$

$$= \mathbf{s}^- + 2\mathbf{P}_{s\theta}^-\left(\mathbf{P}_{\theta\theta}^-\right)^{-1}\Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^+ \tag{65b}$$

Now, recall from the definition of the partitioned matrix inverse that

$$\mathbf{F}_{s\theta}^- = -\mathbf{F}_{ss}^-\mathbf{P}_{s\theta}^-\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} \tag{66a}$$

$$\left(\mathbf{F}_{ss}^-\right)^{-1}\mathbf{F}_{s\theta}^- = -\mathbf{P}_{s\theta}^-\left(\mathbf{P}_{\theta\theta}^-\right)^{-1} \tag{66b}$$

And, substituting this into Eq. (65b),

$$\mathbf{s}^+ = \mathbf{s}^- - 2\left(\mathbf{F}_{\mathbf{ss}}^-\right)^{-1}\mathbf{F}_{\mathbf{s\theta}}^- \, \Xi\left(\hat{\bar{\mathbf{q}}}^-\right)^{\mathrm{T}} \hat{\bar{\mathbf{q}}}^+ \tag{67a}$$

$$\mathbf{s}^+ \approx \mathbf{s}^- - \left(\mathbf{F}_{ss}^-\right)^{-1}\mathbf{F}_{s\theta}^- \, \boldsymbol{\delta\theta} \tag{67b}$$

which are equivalent to the SOAR Filter updates from Eq. (59).

NUMERICAL EXAMPLE

For this numerical example the spacecraft is placed in a circular orbit with an altitude of 622 km and an inclination of 45 degrees. At the beginning of the simulation the Earth is at vernal equinox 20 March 2012 and the spacecraft is at the ascending node. Throughout its orbit the spacecraft is oriented such that the body-fixed X axis is directed in track and the Z axis is Earth-pointing with the Y axis following a right handed coordinate system. As a result the spacecraft has a constant angular velocity equal in magnitude to the orbital mean motion. The sun vector is assumed constant for the duration of the simulation. The magnetic field vector is obtained from the World Magnetic Model in the MATLAB Aerospace toolbox.

A gyro is used to measure the angular velocity of the spacecraft and is defined by the following sensor model [23]

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \boldsymbol{\beta} + \boldsymbol{\eta}_v$$
$$\dot{\boldsymbol{\beta}} = \boldsymbol{\eta}_u$$

where $\boldsymbol{\omega}$ is the true angular velocity, $\tilde{\boldsymbol{\omega}}$ is the measured angular velocity, $\boldsymbol{\beta}$ is the gyro bias vector, and $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are zero-mean Gaussian white-noise processes. Simulated vectors measurements are created by adding noise to the true direction in the spacecraft body frame. The reference vectors remain noise free as the model is assumed perfect for this test case. The scalar weights a_i of the Wahba problem follow the QUEST measurement model and are given by $1/\sigma_{sun}^2$ and $1/\sigma_{mag}^2$ for the sun sensor and magnetometer measurements respectively.

Error Source	Symbol	Value
Sun-sensor noise ($\boldsymbol{\eta}_{sun}$)	σ_{sun}	0.1 deg
Magnetometer noise ($\boldsymbol{\eta}_{mag}$)	σ_{mag}	0.5 deg
Angular Random Walk ($\boldsymbol{\eta}_v$)	σ_v	$\sqrt{10} \times 10^{-7}\,\mathrm{rad/sec}^{1/2}$
Gyro Bias Random Walk ($\boldsymbol{\eta}_u$)	σ_u	$\sqrt{10} \times 10^{-10}\,\mathrm{rad/sec}^{3/2}$

Table 1. Sensors Errors

The state vector consists of the three component gyro bias vector and the three component attitude angle representation $\mathbf{x}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{\beta}^{\mathrm{T}} & \boldsymbol{\theta}^{\mathrm{T}} \end{bmatrix}$. The initial gyro bias covariance is 0.2^2 (deg/hr)^2 in each axis and the initial attitude covariance is 0.1^2 deg^2 in each axis. The initial estimated quaternion is obtained from perturbing the true quaternion according to the initial attitude covariance while the initial estimated gyro bias is always zero. The simulation spans 6000 seconds which is slightly more than one full orbit and uses a step size between observations of 1 second.

Figs. 1 and 2 show the performance of 100 Monte Carlo runs. The red lines are the 100 instances of the estimation error, the black line is the 3-sigma value of the sample standard deviation. Underneath the black lines there are 100 blue lines, which are the 3-sigma filter's prediction of its own

13

uncertainty. Since the predicted uncertainty matches the actual uncertainty we conclude that the filter is consistent. Fig. 3 shows the performance of SOAR under the same circumstances. It can be seen that there is no visible difference between the two algorithms.
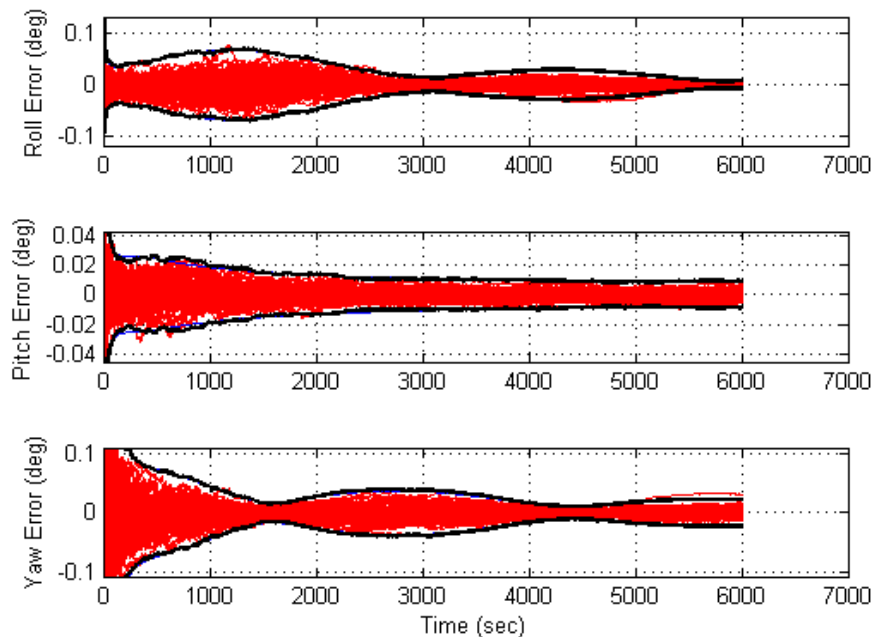

Figure 1. Attitude estimation error of qEKF expressed in body frame

CONCLUSIONS

The q-method was integrated into an EKF-based filter to produce the novel qEKF filter for attitude estimation capable of estimating both attitude and non-attitude states without additional numerical iteration. Within the filter attitude vector measurements are first processed using the q-method which solves the non-linear Wahba problem directly without any linearizing assumptions. Remaining measurements are processed to update the non-attitude states using the standard MEKF algorithm. qEKF was shown to be equivalent to the Sequential Optimal Attitude Recursion (SOAR) filter to second-order in the attitude update and to first-order in the non-attitude state update where each method represents the covariance and information matrix formulation respectively. In qEKF the initial condition is introduced into the Wahba problem through quaternion averaging where the SOAR filter relies on the information matrix approach. The equivalence of qEKF and SOAR was also validated by simulation results in which the filter estimated the attitude and gyro bias.

For this work it was assumed that vector measurements are only functions of the vehicle attitude and not dependent on any other states. A follow-on paper will expand upon this work to extend qEKF to eliminate this assumption and accommodate the inclusion of such states as sensor biases and satellite position. As a result, qEKF will be capable of processing a full range of measurements to estimate both attitude and non-attitude states within a single filter.
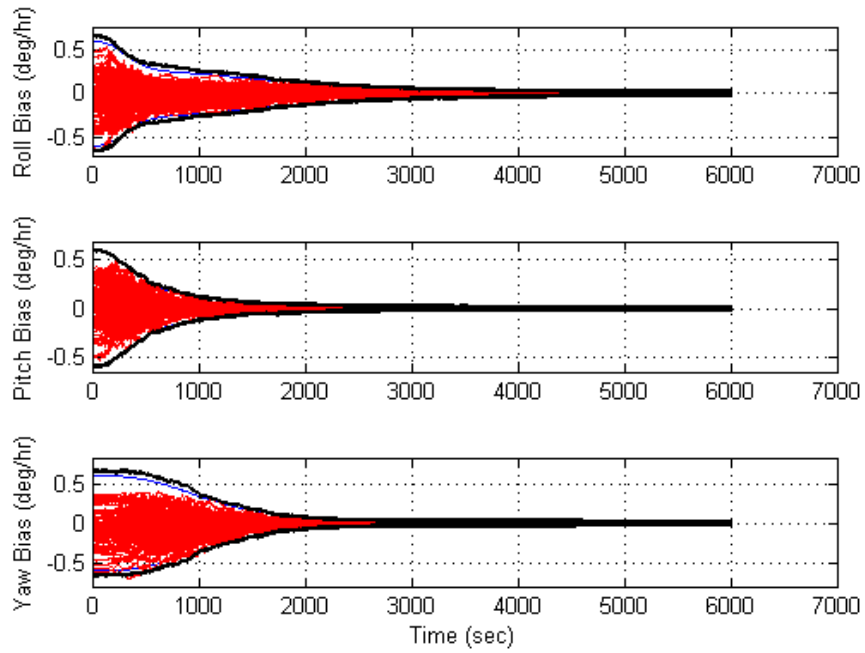
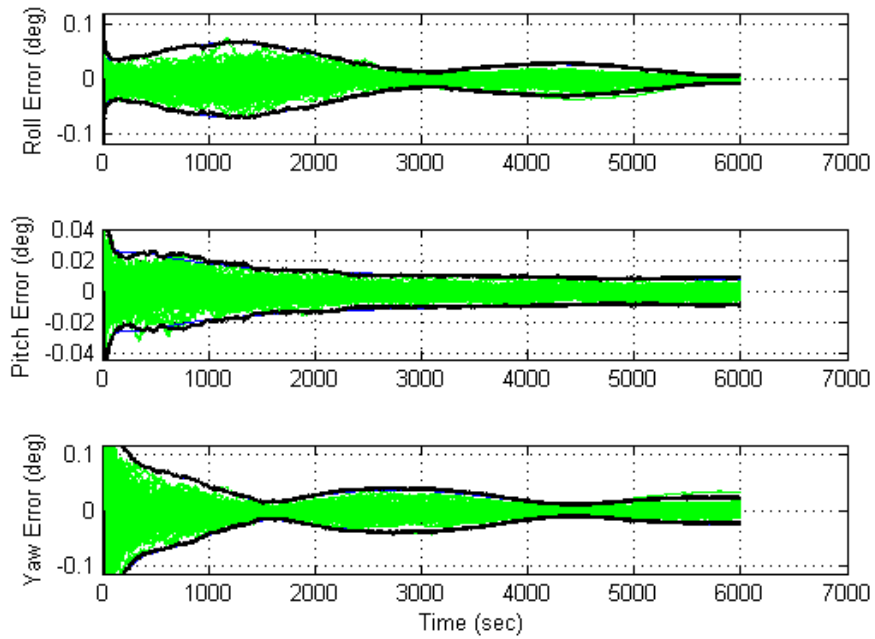**Figure 2. Gyro bias estimation error of qEKF**

Figure 3. Attitude estimation error of SOAR expressed in body frame

REFERENCES

[1] G. Wahba, "A least Square Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, July 1965, p. 409. Problem 65-1.

[2] J. E. Keat, "Analysis of Least-Squares Attitude Determination Routine DOAOP," Tech. Rep. CSC/TM-77/6034, Computer Sciences Corporation, February 1977.

[3] 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.

[4] D. Mortari, "ESOQ: A Closed-Form Solution to the Wahba Problem," *The Journal of the Astronautical Sciences*, Vol. 45, April-June 1997, pp. 195–204.

[5] D. Mortari, "ESOQ-2 Single-Point Algorithm for Fast Optimal Spacecraft Attitude Determination," *Advances in the Astronautical Sciences*, Vol. 95, No. 2, 1997, pp. 817–826.

[6] 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.

[7] Peter H. Schönemann, "A Generalized Solution of the Orthogonal Procrustes Problem," *Psychometrika*, Vol. 31, No. 1, 1966, pp. 1–10.

[8] A. Gelb, ed., *Applied Optimal Estimation*. Cambridge, MA: The MIT press, 1996.

[9] I. Y. Bar-Itzhack and Y. Oshman, "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 21, January 1985, pp. 128–135.

[10] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429.

[11] M. D. Shuster, "A Simple Kalman Filter and Smoother for Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, January-March 1989, pp. 89–106.

[12] I. Y. Bar-Itzhack, "REQUEST: A Recursive QUEST Algorithm for Sequential Attitude Determination," *Journal of Guidance Control and Dynamics*, Vol. 19, September-October 1996, pp. 1034–1038.

[13] M. D. Shuster, "Filter QUEST or REQUEST," *Journal of Guidance, Control, and Dynamics*, Vol. 32, March-April 2009, pp. 643–645.

[14] D. Choukroun, I. Bar-Itzhack, and Y. Oshman, "Optimal-REQUEST Algorithm for Attitude Determination," *Journal of Guidance, Control, and Dynamics*, Vol. 27, May-June 2004, pp. 418–425.

[15] F. L. Markley, "Attitude Determination and Parameter Estimation Using Vector Observations: Theory," *The Journal of the Astronautical Sciences*, Vol. 37, January-March 1989, pp. 41–58.

[16] M. L. Psiaki, "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance Control and Dynamics*, Vol. 23, March-April 2000, pp. 206–214.

[17] J. A. Christian and E. G. Lightsey, "Sequential Optimal Attitude Recursion Filter," *Journal of Guidance Control and Dynamics*, Vol. 33, November-December 2010, pp. 1787–1800.

[18] F. L. Markley, Y. Cheng, J. L. Crassidis, and Y. Oshman, "Averaging Quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 30, July–August 2007, pp. 1193–1197.

[19] M. D. Shuster, "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, January-March 1989, pp. 79–88.

[20] M. D. Shuster, "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences*, Vol. 38, July–September 1990, pp. 377–393.

[21] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processing with Applications to Guidance*. Providence, RI: AMS Chelsea Publishing, 2nd ed., 2005.

[22] H. Sorenson, *Parameter Estimation: Principles and Problems*. New York, NY: Marcel Dekker, Inc, 1980.

[23] R. Farrenkopf, "Analytic Steady State Accuracy Solutions for Two Common Spacecraft Attitude Estimators," *Journal of Guidance, Control, and Dynamics*, Vol. 1, No. 4, 1978, pp. 282–284.