JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 33, No. 6, November–December 2010

Sequential Optimal Attitude Recursion Filter

John A. Christian* and E. Glenn Lightsey†
University of Texas at Austin, Austin, Texas 78758

DOI: 10.2514/1.49561

A new nonlinear attitude filter called the sequential optimal attitude recursion filter is developed. This routine is based on maximum likelihood estimation and a sequentialization of the Wahba problem that has been extended to include nonattitude states. The algorithm can accept either individual unit vector measurements or quaternion measurements. This new algorithm is compared with existing attitude filtering routines, including the multiplicative extended Kalman filter, filter QUEST, REQUEST, optimal REQUEST, and extended QUEST.

Nomenclature

$\mathbf{B}$	=	attitude profile matrix
$E[\,]$	=	expected value operator
$\mathbf{e}_i$	=	ith unit vector observation
$\mathcal{F}$	=	Fisher information matrix
$\mathbf{F}$	=	partitioned segment of Fisher information matrix
J	=	objective function
$\mathbf{K}$	=	Davenport matrix
$\mathbf{P}$	=	state covariance matrix
$p()$	=	probability density function
$\bar{\mathbf{q}}$	=	attitude quaternion
$\mathbf{R}$	=	measurement covariance matrix
$\mathbf{T}_B^A$	=	rotation matrix from A to B
$\mathrm{tr}[\,]$	=	trace operator
$\mathbf{x}$	=	state vector
$\mathbf{y}$	=	measurement vector
$\boldsymbol{\beta}$	=	parameter vector
$\boldsymbol{\delta\theta}$	=	angle error vector
$\boldsymbol{\epsilon}$	=	line-of-sight measurement error
λ	=	Lagrange multiplier
$\boldsymbol{\theta}$	=	angle vector
σ	=	standard deviation
$\boldsymbol{\omega}$	=	angular velocity vector
$\hat{}$	=	estimated quantity
$\tilde{}$	=	measured quantity

Superscripts

$-$	=	a priori quantity
$+$	=	a posteriori quantity

I. Introduction

THE problem of estimating spacecraft attitude from a set of unit vector observations has received much attention since the 1960s. The most well-known formulation of this problem was posed by Wahba [1] in 1965 and is commonly referred to as the *Wahba problem*. If the Wahba problem is written in terms of the attitude quaternion, an analytic solution exists in the form of the solution to an eigenvalue–eigenvector problem. The optimal attitude is simply given by the eigenvector associated with the largest eigenvalue [2]. Numerous algorithms have been developed over the years to find fast and robust solutions to this problem [3]. Some of the most well-known algorithms include the quaternion estimator (QUEST) [4], the estimator of the optimal quaternion (ESOQ) [5,6], and singular-value decomposition [7].

Traditional solutions to the Wahba problem are batch estimators that assume all the unit vector observations occur at the same time. The Wahba problem is a weighted least-squares problem. If the scalar weights in this problem are chosen to be the inverse of the measurement noise variance, then the optimal solution is a maximum likelihood estimate of the attitude to first order [8]. The objective of the present research is to extend the traditional Wahba problem into a framework that allows for the creation of an optimal attitude filter.

One of the most common methods of attitude filtering in modern systems is the multiplicative extended Kalman filter (MEKF) [9,10]. This method uses the four-component attitude quaternion to represent the attitude, but a three-component representation of the attitude in the filter. The MEKF is structured this way because although the attitude quaternion is globally nonsingular, making it a good choice for representing spacecraft attitude, it must obey a unity norm constraint, making direct implementation of the quaternion in a normal Kalman filter difficult. Additionally, a small-angle assumption allows the three-component attitude representation to be cast in a form consistent with the additive nature of the Kalman filter (i.e., $\mathbf{x} = \hat{\mathbf{x}} + \delta\mathbf{x}$).

Unlike the MEKF, sequential Wahba problem filters estimate the full quaternion without requiring solutions to be computed as small-angle deviations from a reference attitude. The batch solutions to the Wahba problem have also been shown to be extremely robust. These are some of the advantages that have led researchers to investigate sequential solutions to the Wahba problem for attitude filtering. Shuster provided the earliest known work in this topic with the development of filter QUEST in a number of papers [11,12]. Subsequently, Bar-Itzhack [13] introduced an alternate sequentialization approach known as the REQUEST algorithm. Several years after the introduction of the REQUEST algorithm, Shuster [14] demonstrated that although the filter QUEST and REQUEST algorithms approach the problem from different perspectives, they are mathematically equivalent. Both filter QUEST and REQUEST algorithms are examples of suboptimal fading-memory filters. Shuster [11,12] demonstrates that under a specific set of conditions, the optimal fading-memory factor, α_k, may be analytically computed. The equation for α_k is simple, fast to compute, and provides an excellent approximation of the more general optimal fading-memory factor for many practical cases. The assumptions required to arrive at this result are rarely met, however, and a better value of α_k may be computed at the expense of a little more computation time.

An approach for the optimal blending of the new attitude measurements with old attitude estimates directly in the Wahba problem solution framework was proposed in 2004 by Choukroun et al. [15] with the optimal REQUEST algorithm. This algorithm requires access to individual unit vector measurements, however. This may

Presented at the 20th AAS/AIAA Space Flight Mechanics Meeting, San Diego, CA, 14–17 February 2010; received 23 February 2010; revision received 10 May 2010; accepted for publication 11 May 2010.

*Graduate Teaching Assistant, Department of Aerospace Engineering and Engineering Mechanics. Member AIAA.

†Professor, Department of Aerospace Engineering and Engineering Mechanics. Associate Fellow AIAA.

cause difficulties if an instrument provides a measurement in the form of an attitude quaternion, as is the case with most commercially available star trackers. Additionally, like filter QUEST and REQUEST, optimal REQUEST is only capable of estimating attitude. None of the sequential Wahba problem filters discussed above are capable of incorporating estimates of other parameters, such as sensor biases.

Taking a different approach, Psiaki [16] developed the extended QUEST algorithm. The extended QUEST algorithm offers an optimal estimate of the attitude and is capable of also estimating nonattitude states. This algorithm is formulated as a square-root information filter that contains a propagation phase and a state update phase. The formulation of extended QUEST requires the *QR* factorization of the propagated information matrix for use in the filter. It also requires the solution to a more general objective function than required for filter QUEST or REQUEST. Despite these increases in complexity and computations, extended QUEST provides an optimal estimate of both attitude and nonattitude states. It has been shown to be robust with convergence properties that are substantially better than seen in the MEKF [16].

Other algorithms have also been developed that are capable of estimating both attitude states and nonattitude states. Most notable among these is an algorithm introduced by Markley [17,18]. While many other nonlinear attitude filtering approaches exist, they are not discussed here. Thorough discussions of these other methods may be found in the literature [19].

It is within this context that a new nonlinear attitude filtering algorithm is proposed: the sequential optimal attitude recursion (SOAR) filter. The derivation of the SOAR filter begins by creating a maximum likelihood estimate of the state vector through Bayesian estimation. A fresh look at the sequentialization of the Wahba problem allows the classic Wahba problem to be recast in a maximum likelihood estimate framework. Additionally, the Bayesian estimation approach allows for the seamless inclusion of nonattitude states in the SOAR filter. Using ideas from the extended Kalman filter (EKF), the covariance inflation associated with the propagation of the state is examined. Treating the covariance in this manner allows for the straightforward inclusion of process noise and addresses the problem of the suboptimal fading-memory factor seen in filter QUEST and REQUEST. If the attitude covariance and measurement covariance are known at the time of the update, the optimal estimate of the attitude may be determined.

The SOAR filter shares many similarities with some of the previous attitude filtering methods. As subsequent discussions will show, the SOAR filter is most similar to extended QUEST. The paper concludes with a side-by-side comparison of the SOAR filter with a number of the other attitude filtering techniques.

II. Review of the Classical Wahba Problem

The problem of estimating vehicle attitude from a set of unit vector observations has been studied in great detail. The objective function of the Wahba problem is one of the most common ways of describing this problem [1]. The Wahba problem objective function is given by

$$\operatorname{Min} J(\mathbf{T}_B^I) = \frac{1}{2}\sum_{i=1}^{m} w_i \|(\tilde{\mathbf{e}}_i)_B - \mathbf{T}_B^I(\mathbf{e}_i)_I\|^2 \tag{1}$$

where m is the number of observed unit vectors, w_i is a positive weighting on the ith observation, $(\tilde{\mathbf{e}}_i)_B$ is the measured unit vector for the ith observation as expressed in the body frame, $(\mathbf{e}_i)_I$ is the known reference unit vector for the ith observation as expressed in a known inertial frame, and $\mathbf{T}_B^I$ is the rotation matrix that transforms a vector from the inertial frame, I, to the body frame, B.

Measurement noise will prevent the rotation matrix $\mathbf{T}_B^I$ from being able to exactly rotate all of the reference unit vectors to match the corresponding observed unit vectors. The error in each line-of-sight unit vector should formally be represented as a rotation:

$$(\tilde{\mathbf{e}}_i)_B = \mathbf{T}_{\tilde{B}_i}^B \mathbf{T}_B^I (\mathbf{e}_i)_I \tag{2}$$

where $\tilde{B}_i$ is some measurement body frame that would produce the measured unit vector. Because the measurement noise is different for each unit vector observation, there will be i different $\tilde{B}$ frames (one for each of the i observed unit vectors), hence the use of a subscript i in the frame definition. Recent work by Mortari and Majji [20] and Zanetti [21] investigates the idea of a multiplicative error in more detail, but these approaches are not pursued further here. Although the measurement error for a unit vector should formally be described by a rotation as in Eq. (2), it is also frequently described by an additive error if the error is small:

$$(\tilde{\mathbf{e}}_i)_B = \mathbf{T}_B^I(\mathbf{e}_i)_I + \boldsymbol{\epsilon}_i = (\mathbf{e}_i)_B + \boldsymbol{\epsilon}_i \tag{3}$$

where $\boldsymbol{\epsilon}_i$ is the measurement error in the observed unit vector as expressed in the body frame. Note that the measurement error $\boldsymbol{\epsilon}_i$ must be small and lie in a plane perpendicular to $(\mathbf{e}_i)_B$ for the measurement $(\tilde{\mathbf{e}}_i)_B$ to remain a unit vector to first order. Therefore,

$$(\mathbf{e}_i)_B^T \boldsymbol{\epsilon}_i \approx 0 \tag{4}$$

Additionally assume that these unit vector measurements are unbiased, such that $E[\boldsymbol{\epsilon}_i] = 0$. Further, it has been shown that the line-of-sight unit vector observations have a covariance that can be described to first order by [8,22]

$$\mathbf{R}_i = E[\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i^T] = \sigma_{\phi,i}^2 [\mathbf{I}_{3\times 3} - (\mathbf{e}_i)_B (\mathbf{e}_i)_B^T] \tag{5}$$

where $\sigma_{\phi,i}^2$ is the angular variance (in radians) of the measurement error. The covariance matrix given in Eq. (5) describes what is frequently referred to in the literature as the QUEST measurement model.

Wahba's problem from Eq. (1) may be rewritten as is commonly seen in the literature [3]:

$$\operatorname{Min} J(\mathbf{T}_B^I) = \lambda_0 - \sum_{i=1}^{m} w_i \operatorname{tr}[\mathbf{T}_B^I (\mathbf{e}_i)_I (\tilde{\mathbf{e}}_i)_B^T] = \lambda_0 - \operatorname{tr}[\mathbf{T}_B^I \mathbf{B}^T] \tag{6}$$

where

$$\lambda_0 = \sum_{i=1}^{m} w_i$$

and $\mathbf{B}$ is called the attitude profile matrix and is given by

$$\mathbf{B} = \sum_{i=1}^{m} w_i (\tilde{\mathbf{e}}_i)_B (\mathbf{e}_i)_I^T \tag{7}$$

Alternatively, the Wahba problem may be rewritten in terms of the attitude quaternion $\bar{\mathbf{q}}$, instead of the rotation matrix $\mathbf{T}_B^I$, using [23]

$$\mathbf{T}(\bar{\mathbf{q}}) = (q_4^2 - \mathbf{q}^T\mathbf{q})\mathbf{I}_{3\times 3} + 2\mathbf{q}\mathbf{q}^T - 2q_4[\mathbf{q}\times] \tag{8}$$

where $\mathbf{q}$ and q_4 are the vector and scalar parts of the attitude quaternion, respectively. Therefore, define the attitude quaternion $\bar{\mathbf{q}}$ as

$$\bar{\mathbf{q}} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} = \begin{bmatrix} \sin\left(\frac{\theta}{2}\right)\mathbf{e}_\theta \\ \cos\left(\frac{\theta}{2}\right) \end{bmatrix} \tag{9}$$

where the unit vector $\mathbf{e}_\theta$ defines the axis of rotation and θ is the angle of rotation about this axis. Additionally, the nomenclature $[\boldsymbol{\alpha}\times]$ is defined as the skew-symmetric cross-product matrix such that $\boldsymbol{\alpha} \times \boldsymbol{\beta} = [\boldsymbol{\alpha}\times]\boldsymbol{\beta}$:

$$[\boldsymbol{\alpha}\times] = \begin{bmatrix} 0 & -\alpha_3 & \alpha_2 \\ \alpha_3 & 0 & -\alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{bmatrix} \tag{10}$$

Making the appropriate substitutions, the results of Eq. (1) or Eq. (6) may be expressed in the following well-known quadratic form [2,4,24]:

$$\operatorname{Min} J(\bar{\mathbf{q}}) = \lambda_0 - \bar{\mathbf{q}}^T \mathbf{K} \bar{\mathbf{q}} \tag{11}$$

where $\mathbf{K}$ is called the Davenport matrix and is defined as

$$\mathbf{K} = \begin{bmatrix} \mathbf{S} - \mu \mathbf{I}_{3\times3} & \mathbf{z} \\ \mathbf{z}^T & \mu \end{bmatrix} \tag{12}$$

and

$$\mathbf{S} = \mathbf{B} + \mathbf{B}^T \qquad \mu = \mathrm{tr}[\mathbf{B}] \qquad [\mathbf{z}\times] = \mathbf{B}^T - \mathbf{B} \tag{13}$$

The attitude that minimizes the objective function in Eq. (11) may be found by the straightforward application of concepts from optimal control theory [25]. Begin by recalling the quaternion unity norm constraint $\|\bar{\mathbf{q}}\| = 1$ and adjoin this constraint to the objective function through a Lagrange multiplier, λ. Therefore, from Eq. (11),

$$\operatorname{Min} \mathcal{J}(\bar{\mathbf{q}}, \lambda) = \lambda_0 - \bar{\mathbf{q}}^T \mathbf{K} \bar{\mathbf{q}} + \lambda(\bar{\mathbf{q}}^T \bar{\mathbf{q}} - 1) \tag{14}$$

The optimal attitude is found by setting the first differential of $\mathcal{J}(\bar{\mathbf{q}}, \lambda)$ to zero. The optimal attitude is given by the solution to [2,4,24]

$$\mathbf{K}\bar{\mathbf{q}} = \lambda \bar{\mathbf{q}} \tag{15}$$

Note that Eq. (15) is an eigenvalue–eigenvector problem. It is from this point that various solutions have emerged. Despite all the discussion circulating about the various attitude determination algorithms, most of these are simply different methods for solving the eigenvalue–eigenvector problem in Eq. (15). Davenport's q method, for example, directly solves Eq. (15) as if it were an ordinary eigenvalue–eigenvector problem [3], thus making it one of the more robust algorithms because it relies on relatively robust but slow numerical methods for solving the eigenvalue–eigenvector problem. Further, by substituting Eq. (15) into Eq. (11), the original objective function becomes

$$J(\bar{\mathbf{q}}) = \lambda_0 - \bar{\mathbf{q}}^T \lambda \bar{\mathbf{q}} = \lambda_0 - \lambda \bar{\mathbf{q}}^T \bar{\mathbf{q}} = \lambda_0 - \lambda \tag{16}$$

Therefore, it may be noted that the optimal attitude is achieved with the largest eigenvalue.

The QUEST algorithm, developed by Shuster and Oh [4], solves this problem by Newton–Raphson iteration on the characteristic equation to find only the largest eigenvalue and associated eigenvector. While faster than Davenport's q method, the solution approach introduces a singularity for a 180 deg rotation that must be addressed in the algorithm implementation (typically through logic checks and sequential rotations).

Estimators of the optimal quaternion (ESOQ and ESOQ-2), developed by Mortari [5,6], are also fundamentally based on the eigenvalue problem described in Eq. (15). Recognizing that

$$(\mathbf{K} - \lambda \mathbf{I}_{4\times4})\bar{\mathbf{q}} = \mathbf{H}\bar{\mathbf{q}} = \mathbf{0} \tag{17}$$

it is clear that $\bar{\mathbf{q}}$ must lie in the null space of $\mathbf{H}$. Further, if $\mathbf{H}^T = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3 \quad \mathbf{h}_4]$, then $\bar{\mathbf{q}}$ must be perpendicular to all $\mathbf{h}_i$. From here, the ESOQ algorithm computes the optimal quaternion through a four-dimensional cross-product operation (a special case of the n-dimensional cross-product) [26]. This operation uses any three rows of $\mathbf{H}$ to find the four-dimensional vector that is perpendicular to the three-dimensional hyperplane in which all four rows of $\mathbf{H}$ exist. Mortari [6] also proposed a follow-on algorithm called ESOQ-2. This algorithm starts with Eq. (15) and computes a symmetric 3×3 matrix of rank 2, $\mathbf{M}$, such that $\mathbf{M}\mathbf{e}_\theta = \mathbf{0}$. Using a similar approach to that seen in the original ESOQ, the attitude may now be determined by taking the regular three-dimensional cross-product of two rows of $\mathbf{M}$.

III. Attitude Error and Covariance Relations

A. Attitude Error Quaternion

Let the error in the attitude be described by the small-error quaternion $\boldsymbol{\delta}\bar{\mathbf{q}}$:

$$\boldsymbol{\delta}\bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes \hat{\bar{\mathbf{q}}}^{-1} \tag{18}$$

where $\otimes$ denotes the quaternion product operator. Physically, $\boldsymbol{\delta}\bar{\mathbf{q}}$ may be interpreted as the quaternion that rotates the best estimated attitude to the true attitude. If one assumes small angles,

$$\boldsymbol{\delta}\bar{\mathbf{q}} = \begin{bmatrix} \boldsymbol{\delta}\mathbf{q} \\ \delta q_4 \end{bmatrix} \approx \begin{bmatrix} \boldsymbol{\delta\theta}/2 \\ 1 \end{bmatrix} \tag{19}$$

where $\boldsymbol{\delta\theta}$ is a three-dimensional parameterization of the attitude given by $\mathbf{e}_{\delta\theta}\delta\theta$. Here, $\mathbf{e}_{\delta\theta}$ is a unit vector denoting the axis of rotation and $\delta\theta$ is the magnitude of the rotation about that axis in radians. Therefore, $2\boldsymbol{\delta}\mathbf{q} \approx \boldsymbol{\delta\theta}$, which yields

$$\mathbf{P}_{\theta\theta} = E[\boldsymbol{\delta\theta}\boldsymbol{\delta\theta}^T] = 4E[\boldsymbol{\delta}\mathbf{q}\boldsymbol{\delta}\mathbf{q}^T] \tag{20}$$

This choice means that $\mathbf{P}_{\theta\theta}$ is the attitude covariance matrix as expressed in the body frame. Additionally, assume that the quaternion attitude estimate is unbiased:

$$E[\boldsymbol{\delta}\bar{\mathbf{q}}] = E\left[\begin{bmatrix} \boldsymbol{\delta}\mathbf{q} \\ \delta q_4 \end{bmatrix}\right] = [0 \quad 0 \quad 0 \quad 1]^T \tag{21}$$

The definition in Eq. (18) also means that the small-angle error vector is defined such that

$$\mathbf{T}(\boldsymbol{\theta}) = \mathbf{T}(\boldsymbol{\delta\theta})\mathbf{T}(\hat{\boldsymbol{\theta}}) \tag{22}$$

If the three-dimensional parameterization of the attitude given by $\boldsymbol{\theta}$ is defined with respect to the attitude estimate given by $\hat{\bar{\mathbf{q}}}$, then $\hat{\boldsymbol{\theta}} = \mathbf{0}$ and $\mathbf{T}(\hat{\boldsymbol{\theta}}) = \mathbf{I}_{3\times3}$ by definition. Therefore,

$$\boldsymbol{\delta\theta} = \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} = \boldsymbol{\theta} - \mathbf{0} = \boldsymbol{\theta} \tag{23}$$

B. Important Attitude Covariance Relations

Next, a few key relations that relate the attitude profile matrix to the covariance matrix are derived. Begin by recalling the Cramér–Rao inequality [27], which states

$$\mathbf{P}_{\mathbf{xx}} = E[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \geq (\mathcal{F}_{\mathbf{xx}})^{-1} \tag{24}$$

where $\mathcal{F}_{\mathbf{xx}}$ is the Fisher information matrix. The Fisher information matrix is further defined as

$$\mathcal{F}_{\mathbf{xx}} = E\left[\frac{\partial^2 J(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}}\right] \tag{25}$$

where the scalar objective function J is the negative log-likelihood function (i.e., $J(\mathbf{x}) = -\ln p(\mathbf{x}|\mathbf{y})$, where $p(\mathbf{x}|\mathbf{y})$ is the probability of observing the state $\mathbf{x}$ given the observations $\mathbf{y}$). The equality in Eq. (24) is obtained if and only if $p(\mathbf{x}|\mathbf{y})$ is Gaussian and the following are both true [27]:

$$\left[\frac{\partial}{\partial \mathbf{x}} J(\mathbf{x})\right]^T = -\left[\frac{\partial}{\partial \mathbf{x}} \ln p(\mathbf{x}|\mathbf{y})\right]^T = \mathbf{C}[\mathbf{x} - \hat{\mathbf{x}}] \tag{26}$$

where $\mathbf{C}$ is a matrix that is independent of $\mathbf{x}$ and $\mathbf{y}$. If the condition of Eq. (26) is applied to a Gaussian $p(\mathbf{x}|\mathbf{y})$, then it may be shown that the observations must be a linear function of $\mathbf{x}$. This observation was also made by Shuster [8]. Therefore, if these conditions are both satisfied then $\mathcal{F}_{\mathbf{xx}} = \mathbf{P}_{\mathbf{xx}}^{-1}$. Unfortunately, this is not the case in the present problem. When the above conditions are not met, $\mathcal{F}_{\mathbf{xx}}$ approaches $\mathbf{P}_{\mathbf{xx}}^{-1}$ as the number of measurements becomes infinite [8]:

$$\mathbf{P}_{\mathbf{xx}}^{-1} = \lim_{n\to\infty} \mathcal{F}_{\mathbf{xx}} \tag{27}$$

where n is the number of observations.

To be consistent with earlier notation, let $\boldsymbol{\delta\theta}$ be the small-angle rotation that rotates the estimate of the vehicle attitude into the true attitude. Therefore, assuming small angles,

$$\mathbf{T}_B^I = \exp\{[-\boldsymbol{\delta\theta}\times]\}\mathbf{T}_{\hat{B}}^I \tag{28}$$

Using this relation, the Wahba problem objective function of Eq. (6) evaluated at the true attitude may be rewritten as a function of $\boldsymbol{\delta\theta}$

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \mathrm{tr}[\exp\{[-\boldsymbol{\delta\theta}\times]\}\mathbf{T}_{\hat{B}}^I\mathbf{B}^T] \tag{29}$$

Expanding the matrix exponential of $[-\boldsymbol{\delta\theta}\times]$ to second order (because the equation for $J(\boldsymbol{\delta\theta})$ must be differentiated twice),

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \mathrm{tr}\left[\left(\mathbf{I}_{3\times3} + [-\boldsymbol{\delta\theta}\times] + \frac{1}{2}[-\boldsymbol{\delta\theta}\times]^2\right)\mathbf{T}_{\hat{B}}^I\mathbf{B}^T\right] \tag{30}$$

Because the Wahba problem has been shown to be equivalent to the negative log-likelihood function (to first order) if the measurement covariance is chosen to be Eq. (5) [8], the Fisher information matrix may be computed by taking the partial derivative of $J(\boldsymbol{\delta\theta})$ twice with respect to $\boldsymbol{\delta\theta}$:

$$\boldsymbol{\mathcal{F}}_{\theta\theta} = \mathrm{tr}[\mathbf{T}_{\hat{B}}^I\mathbf{B}^T]\mathbf{I}_{3\times3} - \mathbf{T}_{\hat{B}}^I\mathbf{B}^T \tag{31}$$

Therefore,

$$\mathbf{P}_{\theta\theta}^{-1} \approx \boldsymbol{\mathcal{F}}_{\theta\theta} = \mathrm{tr}[\mathbf{T}_{\hat{B}}^I\mathbf{B}^T]\mathbf{I}_{3\times3} - \mathbf{T}_{\hat{B}}^I\mathbf{B}^T \tag{32}$$

Rearranging the above equation for the attitude profile matrix $\mathbf{B}$,

$$\mathbf{B} = \left[\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}]\mathbf{I}_{3\times3} - \boldsymbol{\mathcal{F}}_{\theta\theta}\right]\mathbf{T}_{\hat{B}}^I \tag{33}$$

The relationships of Eqs. (32) and (33) were first shown by Shuster in 1989 [8]. From Eqs. (32) and (33), one may freely move back and forth between $\mathbf{B}$ and $\mathbf{P}_{\theta\theta}$.

C. Wahba Problem Objective Function in Terms of Fisher Information

In preparation for the development of the SOAR algorithm, it is useful to find the value of the Wahba problem objective function in terms of the Fisher information matrix. Begin by substituting Eq. (33) into Eq. (30):

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \mathrm{tr}\left[\left(\mathbf{I}_{3\times3} + [-\boldsymbol{\delta\theta}\times] + \frac{1}{2}[-\boldsymbol{\delta\theta}\times]^2\right)\mathbf{T}_{\hat{B}}^I\mathbf{T}_I^{\hat{B}} \times\left(\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}]\mathbf{I}_{3\times3} - \boldsymbol{\mathcal{F}}_{\theta\theta}^T\right)\right] \tag{34}$$

Expanding the trace operator, this may be rewritten as

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}] - \mathrm{tr}\left[[-\boldsymbol{\delta\theta}\times]\left(\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}]\mathbf{I}_{3\times3} - \boldsymbol{\mathcal{F}}_{\theta\theta}^T\right)\right] - \frac{1}{2}\,\mathrm{tr}\left[[-\boldsymbol{\delta\theta}\times]^2\left(\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}]\mathbf{I}_{3\times3} - \boldsymbol{\mathcal{F}}_{\theta\theta}^T\right)\right] \tag{35}$$

As a brief aside, note that for a 3×3 symmetric matrix, $\mathbf{D}$, the following is true:

$$\mathrm{tr}\,[[\boldsymbol{\alpha}\times]\mathbf{D}] = 0 \tag{36}$$

Therefore, noting that both $\mathbf{P}_{\theta\theta}$ and $\boldsymbol{\mathcal{F}}_{\theta\theta}$ are symmetric,

$$\mathrm{tr}\,[[-\boldsymbol{\delta\theta}\times]\mathbf{P}_{\theta\theta}] = 0 \tag{37}$$

$$\mathrm{tr}\,[[-\boldsymbol{\delta\theta}\times]\boldsymbol{\mathcal{F}}_{\theta\theta}] = 0 \tag{38}$$

Applying this result, the third term in Eq. (35) is zero:

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}] - \frac{1}{2}\,\mathrm{tr}\left[[-\boldsymbol{\delta\theta}\times]^2\left(\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}]\mathbf{I}_{3\times3} - \boldsymbol{\mathcal{F}}_{\theta\theta}^T\right)\right] \tag{39}$$

Expanding out the last term will show that this is also equivalent to

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}] + \frac{1}{2}\boldsymbol{\delta\theta}^T\boldsymbol{\mathcal{F}}_{\theta\theta}\boldsymbol{\delta\theta} \tag{40}$$

Taking closer look at the second term in Eq. (40) and substituting Eq. (32) for $\boldsymbol{\mathcal{F}}_{\theta\theta}$,

$$\frac{1}{2}\,\mathrm{tr}[\boldsymbol{\mathcal{F}}_{\theta\theta}] = \frac{1}{2}\,\mathrm{tr}[\mathrm{tr}[\mathbf{T}_{\hat{B}}^I\mathbf{B}^T]\mathbf{I}_{3\times3} - \mathbf{T}_{\hat{B}}^I\mathbf{B}^T] = \mathrm{tr}[\mathbf{T}_{\hat{B}}^I\mathbf{B}^T] = \hat{\bar{\mathbf{q}}}^T\mathbf{K}\hat{\bar{\mathbf{q}}} \tag{41}$$

Therefore, substituting this into Eq. (40) means that the Wahba problem objective function evaluated at the true attitude may be written in terms of the Fisher information matrix as

$$J(\boldsymbol{\delta\theta}) = \lambda_0 - \hat{\bar{\mathbf{q}}}^T\mathbf{K}\hat{\bar{\mathbf{q}}} + \frac{1}{2}\boldsymbol{\delta\theta}^T\boldsymbol{\mathcal{F}}_{\theta\theta}\boldsymbol{\delta\theta} \tag{42}$$

Looking again at Eq. (16), it follows that

$$J(\boldsymbol{\delta\theta}) = J(\hat{\bar{\mathbf{q}}}) + \frac{1}{2}\boldsymbol{\delta\theta}^T\boldsymbol{\mathcal{F}}_{\theta\theta}\boldsymbol{\delta\theta} \tag{43}$$

The form of Eq. (43) is not surprising given that the solution to the Wahba problem is known to produce a maximum likelihood estimate of the attitude. It is also interesting to note that the difference in the value of the objective function at the true attitude and the value of the objective function at the best estimate of the attitude is related to the square of the Mahalanobis distance [28] between these two attitudes.

The relation in Eq. (43) allows for a number of interesting observations. First, this relation will be critical in the derivation of the SOAR filter. Second, the optimal attitude in terms of $\boldsymbol{\delta\theta}$ may be found by taking the differential of Eq. (43) and setting the result equal to zero,

$$dJ(\boldsymbol{\delta\theta}) = [\boldsymbol{\delta\theta}^T\boldsymbol{\mathcal{F}}_{\theta\theta}]d\boldsymbol{\theta} = 0 \tag{44}$$

Because $d\boldsymbol{\theta}$ is an independent differential it is arbitrary, and its coefficient must be equal to zero if $dJ(\boldsymbol{\delta\theta})$ is to be zero. Further, because $\boldsymbol{\mathcal{F}}_{\theta\theta}$ is full rank, it has no null space. Therefore, it is clear that the optimal solution is at $\boldsymbol{\delta\theta} = \mathbf{0}$. Because $\boldsymbol{\delta\theta}$ represents a rotation with respect to the estimated attitude, it is no surprise that the optimal solution should occur at $\boldsymbol{\delta\theta} = \mathbf{0}$. Further, if the second differential of Eq. (43) is greater than zero, then the optimal solution minimizes the objective function. Therefore, taking the second differential yields

$$d^2J(\boldsymbol{\delta\theta}) = d\boldsymbol{\theta}^T\boldsymbol{\mathcal{F}}_{\theta\theta}d\boldsymbol{\theta} > 0 \tag{45}$$

The Fisher information matrix is positive definite, confirming that the optimal attitude does indeed minimize the objective function.

IV. Development of the SOAR Filter

A. Maximum Likelihood State Estimation

The objective is to achieve a maximum likelihood estimate of the state vector at time t_k, $\mathbf{x}_k$, given an a priori estimate and some set of observations, $\mathbf{y}$. Begin by recalling the Bayes theorem:

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})} \tag{46}$$

where $p(\mathbf{x}|\mathbf{y})$ is the probability density function (PDF) that describes the probability of observing the state vector $\mathbf{x}$ conditioned on the knowledge of the observation $\mathbf{y}$.

In the present optimization problem, suppose the objective is to maximize the likelihood of the state estimate at time t_k conditioned on m new measurements and an a priori estimate of the state (with the a priori estimate conditioned on r old measurements). Using the Bayes theorem, this may be mathematically expressed as

$$\mathrm{Max}\,p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{r+m}) = \frac{p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m}|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_r)}{p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m})} \tag{47}$$

Proceed by assuming that each of the new observation are independent,

$$p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m}|\mathbf{x}_k)=\prod_{i=r+1}^{r+m}p(\mathbf{y}_i|\mathbf{x}_k) \tag{48}$$

such that the optimization problem in Eq. (47) may be rewritten as

$$\begin{aligned}&\mathrm{Max}\,p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{r+m})\\&=\frac{p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_r)\prod_{i=r+1}^{r+m}p(\mathbf{y}_i|\mathbf{x}_k)}{p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m})}\end{aligned} \tag{49}$$

This formulation, of course, makes no assumption about the actual distributions of the state vector or the measurements. If the a priori estimate of the state vector at time t_k is assumed to have a Gaussian distribution, then

$$p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_r)=C_{1,k}\exp\left\{-\frac{1}{2}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)^T(\mathbf{P}_k^-)^{-1}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)\right\} \tag{50}$$

where $C_{1,k}$ is a scalar that makes the integral of $p(\cdot)$ equal to unity. This scalar is a function of the covariance matrix. Moving forward, assume that the covariance, and hence $C_{1,k}$, is independent of the state.

Similarly, if the measurement errors are assumed to have a Gaussian distribution,

$$p(\mathbf{y}_i|\mathbf{x}_k)=C_{2,i}\exp\left\{-\frac{1}{2}(\mathbf{y}_i-h_i(\hat{\mathbf{x}}_k^-))^T\mathbf{R}_i^{-1}(\mathbf{y}_i-h_i(\hat{\mathbf{x}}_k^-))\right\} \tag{51}$$

For the moment, assume the measurements are line-of-sight unit vector observations such that $\mathbf{y}_i=(\tilde{\mathbf{e}}_i)_B$ and $h_i(\hat{\mathbf{x}}_k^-)=\mathbf{T}_B^I(\mathbf{e}_i)_I$. Note that because $(\tilde{\mathbf{e}}_i)_B$ is a unit vector, the domain of $\mathbf{y}_i$ is the unit sphere. A more detailed discussion of the PDF for a unit vector measurement, including the complete expression for the scalar $C_{2,i}$, is given by Shuster in an earlier work [8]. As before, assume that $C_{2,i}$ is independent of the state. Making these substitutions,

$$p(\mathbf{y}_i|\mathbf{x}_k)=C_{2,i}\exp\left\{-\frac{1}{2}[(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I]^T\mathbf{R}_i^{-1}[(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I]\right\} \tag{52}$$

Unfortunately, the measurement covariance matrix $\mathbf{R}_i$ of Eq. (5) is singular and the regular inverse may not be computed ($\mathbf{R}_i$ is a 3×3 matrix with a rank of two). Therefore, computing the Moore–Penrose pseudoinverse of $\mathbf{R}_i$ yields

$$\mathbf{R}_i^{\dagger}=\frac{1}{\sigma_{\phi,i}^2}[\mathbf{I}_{3\times3}-(\mathbf{e}_i)_B(\mathbf{e}_i)_B^T] \tag{53}$$

where the † symbol indicates the pseudoinverse. The use of the pseudoinverse in this application has been investigated in more detail by Shuster [29]. Substituting the pseudoinverse for $\mathbf{R}^{-1}$ in Eq. (52) and recalling the identity from Eq. (4),

$$p(\mathbf{y}_i|\mathbf{x}_k)=C_{2,i}\exp\left\{-\frac{1}{2}\frac{1}{\sigma_{\phi,i}^2}[(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I]^T[(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I]\right\} \tag{54}$$

Therefore, substituting the result of Eq. (54) into Eq. (48),

$$\begin{aligned}p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m}|\mathbf{x}_k)&=\prod_{i=r+1}^{r+m}p(\mathbf{y}_i|\mathbf{x}_k)\\&=C_{3,k}\exp\left\{-\frac{1}{2}\sum_{i=r+1}^{r+m}\frac{1}{\sigma_{\phi,i}^2}\|(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I\|^2\right\}\end{aligned} \tag{55}$$

where

$$C_{3,k}=\prod_{i=r+1}^{r+m}C_{2,i}$$

Note that the term in the exponent of Eq. (55) is equivalent to the traditional Wahba problem objective function to first order if $w_i=1/\sigma_{\phi,i}^2$. This demonstrates that solutions to the Wahba problem produce a maximum likelihood estimate of the attitude if $w_i=1/\sigma_{\phi,i}^2$.

Substituting Eqs. (50) and (55) into Eq. (49), a maximum likelihood estimate of the state exists at

$$\begin{aligned}\mathrm{Max}\,p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{r+m})&=\frac{C_{1,k}C_{3,k}}{p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m})}\\&\times\exp\left\{-\frac{1}{2}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)^T(\mathbf{P}_k^-)^{-1}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)\right\}\\&\times\exp\left\{-\frac{1}{2}\sum_{i=r+1}^{r+m}\frac{1}{\sigma_{\phi,i}^2}\|(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I\|^2\right\}\end{aligned} \tag{56}$$

Note that the use of Eq. (53) means that the optimal solution obtained from Eq. (56) is the result of a first-order approximation of the line-of-sight unit vector errors. Further, recognizing that the maximum of $p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{r+m})$ occurs at the same value of the state as the maximum of $\ln[p(\mathbf{x}_k|\mathbf{y}_1,\mathbf{y}_2,\ldots,\mathbf{y}_{r+m})]$, a maximum likelihood estimate of the state may also be found by finding

$$\begin{aligned}&\mathrm{Max}\,\ln[C_{1,k}]+\ln[C_{3,k}]-\ln[p(\mathbf{y}_{r+1},\ldots,\mathbf{y}_{r+m})]\\&-\frac{1}{2}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)^T(\mathbf{P}_k^-)^{-1}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)\\&-\frac{1}{2}\sum_{i=r+1}^{r+m}\frac{1}{\sigma_{\phi,i}^2}\|(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I\|^2\end{aligned} \tag{57}$$

Because the first three terms do not depend on the estimate of the state, they may be dropped from the objective function in determining the maximum likelihood estimate. Dropping the these terms and rewriting as a minimization problem yields the following form of the objective function:

$$\begin{aligned}\mathrm{Min}L&=\frac{1}{2}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)^T(\mathbf{P}_k^-)^{-1}(\mathbf{x}_k-\hat{\mathbf{x}}_k^-)\\&+\frac{1}{2}\sum_{i=r+1}^{r+m}\frac{1}{\sigma_{\phi,i}^2}\|(\tilde{\mathbf{e}}_i)_B-\mathbf{T}_B^I(\mathbf{e}_i)_I\|^2\end{aligned} \tag{58}$$

B. Discussion of the SOAR Filter State Vector

The objective of the present filter is to estimate the attitude of a vehicle along with a number of other vehicle parameters. To express the attitude, it is desired to use a minimal degree representation (a number of different three-dimensional parameterizations would work here) instead of the constrained four-dimensional quaternion. Here, the attitude will be represented by the vector $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}=\mathbf{e}_\theta\theta \tag{59}$$

where $\boldsymbol{\theta}$ is expressed in the body frame and represents the rotation from the current best estimate of the attitude to the true attitude. As was observed earlier, this means that $\hat{\boldsymbol{\theta}}^-\equiv\mathbf{0}$. The implications of this selection are critical to the proper treatment of attitude in the filter and will be discussed in detail in the following sections.

In addition to estimating the attitude, the SOAR filter also estimates other vehicle parameters that are defined by the parameter vector $\boldsymbol{\beta}$. Therefore, the state vector for the SOAR filter is given by

$$\mathbf{x}=\begin{bmatrix}\boldsymbol{\theta}\\\boldsymbol{\beta}\end{bmatrix} \tag{60}$$

C. Reformulation of the SOAR Objective Function

Recalling the relation between the covariance matrix and the Fisher information matrix from Eq. (27), and define the components of the partitioned Fisher information matrix as

$$\mathbf{P}^{-1} \approx \mathbf{F} = \begin{bmatrix} \mathbf{F}_{\theta\theta} & \mathbf{F}_{\theta\beta} \\ \mathbf{F}_{\beta\theta} & \mathbf{F}_{\beta\beta} \end{bmatrix} \tag{61}$$

Substituting this result at time k into the objective function from Eq. (58) yields

$$\begin{aligned} \mathrm{Min}L = \frac{1}{2}[&\delta\boldsymbol{\theta}_k^T \mathbf{F}_{\theta\theta,k}^- \delta\boldsymbol{\theta}_k + \delta\boldsymbol{\theta}_k^T \mathbf{F}_{\theta\beta,k}^- \delta\boldsymbol{\beta}_k + \delta\boldsymbol{\beta}_k^T \mathbf{F}_{\beta\theta,k}^- \delta\boldsymbol{\theta}_k \\ &+ \delta\boldsymbol{\beta}_k^T \mathbf{F}_{\beta\beta,k}^- \delta\boldsymbol{\beta}_k] + \frac{1}{2}\sum_{i=r+1}^{r+m} \frac{1}{\sigma_{\phi,i}^2} \|(\tilde{\mathbf{e}}_i)_B - \mathbf{T}_B^I (\mathbf{e}_i)_I\|^2 \end{aligned} \tag{62}$$

The subscript k is temporarily dropped in the following derivation to abbreviate notation. This may be done because all terms in Eq. (62) are expressed at time t_k. Note, however, that the a priori terms (denoted with a superscript $-$) may contain information from past measurements and state estimates that has been mapped to time t_k.

Additionally, care must be taken here not to confuse the $\mathcal{F}_{\theta\theta}$ generated from the inverse of the 3×3 attitude covariance matrix in Eq. (32) with the $\mathbf{F}_{\theta\theta}$ described in Eqs. (61) and (62). The $\mathbf{F}_{\theta\theta}$ in Eqs. (61) and (62) is the upper 3×3 matrix of the inverse of the entire covariance matrix, which contains both attitude and nonattitude states. The different font is chosen to help keep this distinction clear.

Further, recognizing that $\mathbf{F}_{\theta\beta} = \mathbf{F}_{\beta\theta}^T$ and using the form of the Wahba problem cost function from Eq. (11),

$$\mathrm{Min}\, L = \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\theta}^- \delta\boldsymbol{\theta} + \delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\beta}^- \delta\boldsymbol{\beta} + \frac{1}{2}\delta\boldsymbol{\beta}^T \mathbf{F}_{\beta\beta}^- \delta\boldsymbol{\beta} + \lambda_0^m - \bar{\mathbf{q}}^T \mathbf{K}^m \bar{\mathbf{q}} \tag{63}$$

where $\mathbf{K}^m$ is the measurement Davenport matrix constructed using only the new measurements. If the measurements are unit vectors, then $\mathbf{K}^m$ is computed as in the regular Wahba problem. If the measurements are quaternions (as is sometimes the case for commercially available star trackers that provide a quaternion estimate instead of the raw measurements), then the measured attitude and measurement covariance matrix are used to create $\mathbf{B}^m$ from Eq. (33). Now $\mathbf{B}^m$ may be used to compute $\mathbf{K}^m$ from Eqs. (12) and (13).

Note that $\delta\boldsymbol{\theta}$ and $\bar{\mathbf{q}}$ in Eq. (63) are just two different ways of expressing the same attitude. The attitude described by these variables is not generally equal to the true attitude. It is the attitude at which the objective function, L, is evaluated and is a variable in the minimization problem.

To obtain a more useful expression for Eq. (63), begin by looking at the first term on the right hand side. Recalling the following relation for the inverse of a partitioned matrix,

$$\begin{aligned} &\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} [\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]^{-1} & -[\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}]^{-1}\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}[\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}]^{-1} & [\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}]^{-1} \end{bmatrix} \end{aligned} \tag{64}$$

it may be shown through a little algebra that

$$\mathbf{F}_{\theta\theta} = \mathbf{P}_{\theta\theta}^{-1} + \mathbf{F}_{\theta\beta}\mathbf{F}_{\beta\beta}^{-1}\mathbf{F}_{\beta\theta} = \mathcal{F}_{\theta\theta} + \mathbf{F}_{\theta\beta}\mathbf{F}_{\beta\beta}^{-1}\mathbf{F}_{\beta\theta} \tag{65}$$

Therefore, the first term in Eq. (63) may be rewritten as

$$\frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\theta}^- \delta\boldsymbol{\theta} = \frac{1}{2}\delta\boldsymbol{\theta}^T \mathcal{F}_{\theta\theta}^- \delta\boldsymbol{\theta} + \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\beta}^- (\mathbf{F}_{\beta\beta}^-)^{-1} \mathbf{F}_{\beta\theta}^- \delta\boldsymbol{\theta} \tag{66}$$

Now, from Eq. (43), it is clear that

$$\frac{1}{2}\delta\boldsymbol{\theta}^T \mathcal{F}_{\theta\theta}^- \delta\boldsymbol{\theta} = J(\delta\boldsymbol{\theta}) - J(\hat{\bar{\mathbf{q}}}^-) = J(\bar{\mathbf{q}}) - J(\hat{\bar{\mathbf{q}}}^-) \tag{67}$$

and substituting the result from Eq. (11) into this expression yields

$$\frac{1}{2}\delta\boldsymbol{\theta}^T \mathcal{F}_{\theta\theta}^- \delta\boldsymbol{\theta} = \lambda_0 - \bar{\mathbf{q}}^T \mathbf{K}^- \bar{\mathbf{q}} - J(\hat{\bar{\mathbf{q}}}^-) \tag{68}$$

Therefore, substituting the result of Eq. (68) into Eq. (66), it may be shown that

$$\frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\theta}^- \delta\boldsymbol{\theta} = \lambda_0^- - \bar{\mathbf{q}}^T \mathbf{K}^- \bar{\mathbf{q}} - J(\hat{\bar{\mathbf{q}}}^-) + \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\beta}^- (\mathbf{F}_{\beta\beta}^-)^{-1} \mathbf{F}_{\beta\theta}^- \delta\boldsymbol{\theta} \tag{69}$$

As a brief aside, the a priori Davenport matrix, $\mathbf{K}^-$, may be found by using Eqs. (12) and (13) and the a priori attitude profile matrix, $\mathbf{B}^-$. The a priori attitude and covariance may be used to find $\mathbf{B}^-$ from Eq. (33):

$$\mathbf{B}^- = \left[\frac{1}{2}\,\mathrm{tr}[(\mathbf{P}_{\theta\theta}^-)^{-1}]\mathbf{I}_{3\times3} - (\mathbf{P}_{\theta\theta}^-)^{-1}\right]\mathbf{T}_{\hat{B}}^I \tag{70}$$

where $\mathbf{T}_{\hat{B}}^I$ is found through Eq. (8) using $\hat{\bar{\mathbf{q}}}^-$ as the attitude.

Substituting the result of Eq. (69) into the objective function from Eq. (63) and combining like terms,

$$\begin{aligned} \mathrm{Min}L = &(\lambda_0^m + \lambda_0^-) - J(\hat{\bar{\mathbf{q}}}^-) - \bar{\mathbf{q}}^T(\mathbf{K}^m + \mathbf{K}^-)\bar{\mathbf{q}} \\ &+ \frac{1}{2}\delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\beta}^- (\mathbf{F}_{\beta\beta}^-)^{-1} \mathbf{F}_{\beta\theta}^- \delta\boldsymbol{\theta} + \delta\boldsymbol{\theta}^T \mathbf{F}_{\theta\beta}^- \delta\boldsymbol{\beta} + \frac{1}{2}\delta\boldsymbol{\beta}^T \mathbf{F}_{\beta\beta}^- \delta\boldsymbol{\beta} \end{aligned} \tag{71}$$

Note that this expression contains both $\bar{\mathbf{q}}$ and $\delta\boldsymbol{\theta}$. These are simply two different ways of expressing the attitude. Therefore, in order to simplify this expression, note the following relations associated with rotations and quaternion multiplication:

$$\begin{bmatrix} \delta\boldsymbol{\theta}/2 \\ 1 \end{bmatrix} \approx \delta\bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes (\hat{\bar{\mathbf{q}}}^-)^{-1} = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} \otimes \begin{bmatrix} -\hat{\mathbf{q}}^- \\ \hat{q}_4^- \end{bmatrix} \tag{72}$$

Using the definition of the quaternion product, proceed by showing that

$$\begin{aligned} \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} \otimes \begin{bmatrix} -\hat{\mathbf{q}}^- \\ \hat{q}_4^- \end{bmatrix} &= \begin{bmatrix} \hat{q}_4^- \mathbf{q} - q_4 \hat{\mathbf{q}}^- + \mathbf{q} \times \hat{\mathbf{q}}^- \\ q_4 \hat{q}_4^- + \mathbf{q}^T \hat{\mathbf{q}}^- \end{bmatrix} \\ &= \begin{bmatrix} \hat{q}_4^- \mathbf{I}_{3\times3} - [\hat{\mathbf{q}}^- \times] & -\hat{\mathbf{q}}^- \\ (\hat{\mathbf{q}}^-)^T & \hat{q}_4^- \end{bmatrix} \bar{\mathbf{q}} \end{aligned} \tag{73}$$

Therefore, define the 3×4 matrix $\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}_k^-)$ as

$$\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) = [\,\hat{q}_4^- \mathbf{I}_{3\times3} - [\hat{\mathbf{q}}^- \times] \quad -\hat{\mathbf{q}}^-\,] \tag{74}$$

such that

$$\delta\boldsymbol{\theta} = 2\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)\bar{\mathbf{q}} \tag{75}$$

Now, substituting this identity into Eq. (71) allows the objective function to be written as

$$\begin{aligned} \mathrm{Min}L = &(\lambda_0^m + \lambda_0^-) - J(\hat{\bar{\mathbf{q}}}^-) - \bar{\mathbf{q}}^T(\mathbf{K}^m + \mathbf{K}^-)\bar{\mathbf{q}} \\ &+ 2\bar{\mathbf{q}}^T \boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T \mathbf{F}_{\theta\beta}^- (\mathbf{F}_{\beta\beta}^-)^{-1} \mathbf{F}_{\beta\theta}^- \boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)\bar{\mathbf{q}} + 2\bar{\mathbf{q}}^T \boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T \mathbf{F}_{\theta\beta}^- \delta\boldsymbol{\beta} \\ &+ \frac{1}{2}\delta\boldsymbol{\beta}^T \mathbf{F}_{\beta\beta}^- \delta\boldsymbol{\beta} \end{aligned} \tag{76}$$

D. Finding the Optimal State Update in SOAR

Using a similar approach as for the optimal solution for the regular Wahba problem, adjoin the attitude quaternion unity norm constraint to the objective function using a Lagrange multiplier, λ, such that

$$\begin{aligned}\operatorname{Min}\mathcal{L}(\bar{\mathbf{q}},\boldsymbol{\beta},\lambda) &= (\lambda_0^m+\lambda_0^-) - J(\hat{\bar{\mathbf{q}}}^-) - \bar{\mathbf{q}}^T(\mathbf{K}^m+\mathbf{K}^-)\bar{\mathbf{q}} \\ &\quad + 2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)\bar{\mathbf{q}} + 2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-\delta\boldsymbol{\beta} \\ &\quad + \frac{1}{2}\delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\beta}^-\delta\boldsymbol{\beta} + \lambda(\bar{\mathbf{q}}^T\bar{\mathbf{q}}-1)\end{aligned} \tag{77}$$

The first differential necessary condition states that the first differential must vanish at a minimal point [25]. Therefore, taking the differential of $\mathcal{L}(\bar{\mathbf{q}},\boldsymbol{\beta},\lambda)$ and setting the result equal to zero,

$$\begin{aligned}d\mathcal{L}(\bar{\mathbf{q}},\boldsymbol{\beta},\lambda) &= \{-2\bar{\mathbf{q}}^T(\mathbf{K}^m+\mathbf{K}^-) \\ &\quad + 4\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) + 2\delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) \\ &\quad + 2\lambda\bar{\mathbf{q}}^T\}d\bar{\mathbf{q}} + \{2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^- + \delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\beta}^-\}d\boldsymbol{\beta} \\ &\quad + \{\bar{\mathbf{q}}^T\bar{\mathbf{q}}-1\}d\lambda = 0\end{aligned} \tag{78}$$

Now look at the coefficient of the $d\lambda$ term. This is simply the quaternion unity norm constraint and, if the constraint is satisfied, this coefficient is equal to zero:

$$\bar{\mathbf{q}}^T\bar{\mathbf{q}} - 1 = 0 \tag{79}$$

Next, look at the coefficient of the $d\boldsymbol{\beta}$ term in Eq. (78). Assuming no constraints on the nonattitude states, $d\boldsymbol{\beta}$ is an independent differential. Therefore, if $d\mathcal{J}(\bar{\mathbf{q}},\boldsymbol{\beta},\lambda)$ is to be zero for an arbitrary $d\boldsymbol{\beta}$, the coefficient of $d\boldsymbol{\beta}$ must be zero:

$$2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^- + \delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\beta}^- = 0 \tag{80}$$

Rearranging and solving for $\delta\boldsymbol{\beta}$ (again recall that $\mathbf{F}_{\beta\beta}$ is symmetric),

$$\delta\boldsymbol{\beta} = -2(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)\bar{\mathbf{q}} \tag{81}$$

Finally, look at the term involving $d\bar{\mathbf{q}} = [d\mathbf{q}^T \quad dq_4]^T$ in Eq. (78). Because of the unity norm constraint, not all four components of $d\bar{\mathbf{q}}$ are independent differentials. Therefore, let the first three components of $d\bar{\mathbf{q}}$ be independent differentials and let the fourth component be the dependent differential. Because the Lagrange multiplier λ is arbitrary, choose it such that the coefficient of the dependent differential dq_4 is zero [25]. The remaining differentials in $d\bar{\mathbf{q}}$ are independent and their coefficients must be zero if $d\mathcal{J}(\bar{\mathbf{q}},\boldsymbol{\beta},\lambda)$ is to be zero for an arbitrary $d\mathbf{q}$. Therefore, the coefficient of $d\bar{\mathbf{q}}$ in Eq. (78) must be zero:

$$\begin{aligned}&-2\bar{\mathbf{q}}^T(\mathbf{K}^m+\mathbf{K}^-) + 4\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) \\ &\quad + 2\delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) + 2\lambda\bar{\mathbf{q}}^T = 0\end{aligned} \tag{82}$$

Substituting the optimal value of $\delta\boldsymbol{\beta}$ from Eq. (81),

$$\begin{aligned}&-\bar{\mathbf{q}}^T(\mathbf{K}^m+\mathbf{K}^-) + 2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) \\ &\quad - 2\bar{\mathbf{q}}^T\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-) + \lambda\bar{\mathbf{q}}^T = 0\end{aligned} \tag{83}$$

The middle two terms cancel, leaving

$$-\bar{\mathbf{q}}^T(\mathbf{K}^m+\mathbf{K}^-) + \lambda\bar{\mathbf{q}}^T = 0 \tag{84}$$

Therefore, if $\mathbf{K}^+$ is given by

$$\mathbf{K}^+ = \mathbf{K}^m + \mathbf{K}^- \tag{85}$$

then Eq. (84) may be rewritten as

$$-\bar{\mathbf{q}}^T\mathbf{K}^+ + \lambda\bar{\mathbf{q}}^T = 0 \tag{86}$$

Recognizing that $\mathbf{K}^+$ is symmetric and that the solution to Eq. (86) is the a posteriori attitude estimate, which allows $\bar{\mathbf{q}}$ to be replaced by $\hat{\bar{\mathbf{q}}}^+$:

$$\mathbf{K}^+\hat{\bar{\mathbf{q}}}^+ = \lambda\hat{\bar{\mathbf{q}}}^+ \tag{87}$$

The optimal $\hat{\bar{\mathbf{q}}}^+$ may be found using any solution method to the normal Wahba problem. Once $\hat{\bar{\mathbf{q}}}^+$ is known, $\delta\boldsymbol{\beta}$ may be found using Eq. (81). The updated parameter vector, therefore, is simply given by

$$\boldsymbol{\beta}^+ = \boldsymbol{\beta}^- + \delta\boldsymbol{\beta} = \boldsymbol{\beta}^- - 2(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\boldsymbol{\Psi}(\hat{\bar{\mathbf{q}}}^-)\hat{\bar{\mathbf{q}}}^+ \tag{88}$$

E. Finding the Optimal Covariance Update in SOAR

Now that the optimal state update has been derived, it is necessary to compute the corresponding update to the state covariance matrix. Begin by rewriting the objective function from Eq. (71) as

$$\begin{aligned}\operatorname{Min}L &= (\lambda_0^m+\lambda_0^-) - J(\hat{\bar{\mathbf{q}}}^-) - \operatorname{tr}[\mathbf{T}_B^I(\mathbf{B}^m+\mathbf{B}^-)^T] \\ &\quad + \frac{1}{2}\delta\boldsymbol{\theta}^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\delta\boldsymbol{\theta} + \delta\boldsymbol{\theta}^T\mathbf{F}_{\theta\beta}^-\delta\boldsymbol{\beta} + \frac{1}{2}\delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\beta}^-\delta\boldsymbol{\beta}\end{aligned} \tag{89}$$

As was done in finding the covariance for the regular Wahba problem in Eq. (30), rewrite the attitude term in terms of $\delta\boldsymbol{\theta}$, where $\delta\boldsymbol{\theta}$ is a small-angle rotation from the best estimate of the body frame:

$$\begin{aligned}\operatorname{Min}L &= (\lambda_0^m+\lambda_0^-) - J(\hat{\bar{\mathbf{q}}}^-) \\ &\quad - \operatorname{tr}\left[\left(\mathbf{I}_{3\times 3} + [-\delta\boldsymbol{\theta}\times] + \frac{1}{2}[-\delta\boldsymbol{\theta}\times]^2\right)\mathbf{T}_{\hat{B}}^I(\mathbf{B}^m+\mathbf{B}^-)^T\right] \\ &\quad + \frac{1}{2}\delta\boldsymbol{\theta}^T\mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\delta\boldsymbol{\theta} + \delta\boldsymbol{\theta}^T\mathbf{F}_{\theta\beta}^-\delta\boldsymbol{\beta} + \frac{1}{2}\delta\boldsymbol{\beta}^T\mathbf{F}_{\beta\beta}^-\delta\boldsymbol{\beta}\end{aligned} \tag{90}$$

If the relation in Eq. (85) is true, then the following must also be true:

$$\mathbf{B}^+ = \mathbf{B}^m + \mathbf{B}^- \tag{91}$$

Substituting this result into Eq. (90), the a posteriori Fisher information matrix may be computed using the relation given in Eq. (25). Therefore, the straightforward differentiation of Eq. (90) will yield the following results:

$$\mathbf{F}_{\theta\theta}^+ = \operatorname{tr}[\mathbf{T}_{\hat{B}}^I(\mathbf{B}^+)^T]\mathbf{I}_{3\times 3} - \mathbf{T}_{\hat{B}}^I(\mathbf{B}^+)^T + \mathbf{F}_{\theta\beta}^-(\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^- \tag{92}$$

$$\mathbf{F}_{\theta\beta}^+ = (\mathbf{F}_{\beta\theta}^+)^T = \mathbf{F}_{\theta\beta}^- \tag{93}$$

$$\mathbf{F}_{\beta\beta}^+ = \mathbf{F}_{\beta\beta}^- \tag{94}$$

This makes sense in that the only new information available is in the attitude estimate. The present implementation of the SOAR filter only allows line-of-sight unit vector measurements and attitude measurements, therefore improved estimates of the parameter vector $\boldsymbol{\beta}$ is only possible through knowledge of the dynamics and correlation between the estimates of the attitude and parameters. Further, it is interesting to note the relationship between Eqs. (92) and (32). The first two terms in Eq. (92) are what would be expected from Eq. (32) to go from $\mathbf{B}^+$ to $\mathbf{F}_{\theta\theta}^+$ in a system in which only attitude states are estimated. The third term, which is new, follows from the inclusion of nonattitude states.

To compute the a posteriori covariance matrix, recall the relation for the inverse of a segmented matrix given in Eq. (64). From this relation, it is straightforward to see that

$$\mathbf{P}_{\theta\theta}^+ = [\mathbf{F}_{\theta\theta}^+ - \mathbf{F}_{\theta\beta}^+(\mathbf{F}_{\beta\beta}^+)^{-1}\mathbf{F}_{\beta\theta}^+]^{-1} \tag{95}$$

$$\mathbf{P}_{\beta\beta}^+ = [\mathbf{F}_{\beta\beta}^+ - \mathbf{F}_{\beta\theta}^+(\mathbf{F}_{\theta\theta}^+)^{-1}\mathbf{F}_{\theta\beta}^+]^{-1} \tag{96}$$

$$\mathbf{P}_{\beta\theta}^+ = (\mathbf{F}_{\beta\beta}^-)^{-1}\mathbf{F}_{\beta\theta}^-\mathbf{P}_{\theta\theta}^+ \tag{97}$$

The same results may also be found using the standard definition of the covariance matrix (i.e., a covariance/expected value approach rather than an information approach). The details of this equivalence are not presented here, for brevity.

F. Propagation of State and Covariance for SOAR

The only remaining step is to propagate the a posteriori state estimate and covariance at time t_k forward in time to create the a priori state estimate and covariance at time t_{k+1}. Consider a system with the following nonlinear dynamic model:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) + \mathbf{G}(t)\mathbf{w}(t) \tag{98}$$

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t), t) \tag{99}$$

Proceed by taking the Taylor series expansion of $\mathbf{f}(\mathbf{x}, \mathbf{u}, t)$ about the estimated trajectory:

$$\dot{\mathbf{x}} = \left[\mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t), t) + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}(t)}\right)(\mathbf{x}(t) - \hat{\mathbf{x}}(t)) + \text{H.O.T.}\right] + \mathbf{G}(t)\mathbf{w}(t) \tag{100}$$

where H.O.T. stands for higher-order terms.

Defining the matrix $\mathbf{F}(\hat{\mathbf{x}}(t), t)$ as

$$\mathbf{F}(\hat{\mathbf{x}}(t), t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\bigg|_{\mathbf{x}=\hat{\mathbf{x}}(t)} \tag{101}$$

and establishing the following state error vectors,

$$\boldsymbol{\delta}\,\mathbf{x} = \mathbf{x} - \hat{\mathbf{x}} \qquad \boldsymbol{\delta}\dot{\mathbf{x}} = \dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} \tag{102}$$

allows for Eq. (100) to be rewritten as the following linearized model:

$$\boldsymbol{\delta}\,\dot{\mathbf{x}} = \mathbf{F}(\hat{\mathbf{x}}, t)\boldsymbol{\delta}\mathbf{x} + \mathbf{G}(t)\mathbf{w}(t) \tag{103}$$

This is a standard result that seen in the implementation of the EKF; it is well-known and discussed thoroughly in the literature [30–32]. Such a linear model is known to admit a solution of the form

$$\boldsymbol{\delta}\,\mathbf{x}(t) = \boldsymbol{\Phi}(t, t_0)\boldsymbol{\delta}\mathbf{x}(t_0) + \int_{t_0}^{t} \boldsymbol{\Phi}(t, \tau)\mathbf{G}(\tau)\mathbf{w}(\tau)\,\mathrm{d}\tau \tag{104}$$

where $\boldsymbol{\Phi}(t, t_0)$ is the state transition matrix from t_0 to t. Therefore, assuming that the state error and the process noise are uncorrelated, i.e., $E[\boldsymbol{\delta}\mathbf{x}(t)\mathbf{w}(t)^T] = 0$, it may be shown that

$$\mathbf{P} = E[\boldsymbol{\delta}\mathbf{x}\boldsymbol{\delta}\mathbf{x}^T] = \boldsymbol{\Phi}(t, t_0)\mathbf{P}_0\boldsymbol{\Phi}(t, t_0)^T + \int_{t_0}^{t} \boldsymbol{\Phi}(t, \tau)\mathbf{G}(\tau)\mathbf{Q}(\tau)\mathbf{G}^T\boldsymbol{\Phi}(t, \tau)^T\,\mathrm{d}\tau \tag{105}$$

where

$$\mathbf{Q}(\tau) = E[\mathbf{w}(\tau)\mathbf{w}(\tau)^T] \tag{106}$$

Taking the time derivative of Eq. (105) (which requires the application of Leibniz's rule) and recognizing that $\dot{\boldsymbol{\Phi}} = \mathbf{F}\boldsymbol{\Phi}$, yields the well-known equation for the propagation of the covariance matrix typically seen in the discrete-continuous EKF:

$$\dot{\mathbf{P}}(t) = \mathbf{F}(\hat{\mathbf{x}}, t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{F}(\hat{\mathbf{x}}, t)^T + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}(t)^T \tag{107}$$

This expression is used to propagate the covariance between measurement updates in SOAR. This is also the same expression used to propagate the covariance in the MEKF. Earlier work by Markley [10] provides a more detailed discussion of covariance propagation in the MEKF.

Note that all of the SOAR filter update equations make the assumption that the a priori state error and the measurement error have Gaussian distributions. The equations of motion for attitude propagation are nonlinear and it is well-known Gaussian distributions undergoing a nonlinear transformation do not necessarily remain Gaussian. Therefore, it is unlikely that the a priori state error at any given time is truly Gaussian. Despite the fact that the true underlying distribution may not actually be Gaussian, many filters still assume a Gaussian distribution. Empirical evidence has shown that these filters still perform well in many scenarios, even though the simplifying assumption of a Gaussian distribution may not strictly be true.

V. Construction of the SOAR Algorithm

All the theoretical components required to create the SOAR filter are now in place. A brief summary of the SOAR filter is provided in the following paragraphs. A flowchart of the SOAR filter is shown in Fig. 1.

Given an a priori attitude estimate and covariance matrix, compute the a priori attitude profile matrix from Eq. (33). Use this to create the a priori Davenport matrix using Eqs. (12) and (13).

Next, compute the measurement attitude profile matrix. If the measurements are from unit vectors, create the measurement attitude profile matrix using Eq. (7). If, on the other hand, a measured attitude is available, compute the measurement attitude profile matrix using Eq. (33). Because the Davenport matrix $\mathbf{K}$ is a linear homogeneous function of $\mathbf{B}$, the measurement attitude profile matrices for all measurement sources (whether they come from unit vector measurements or quaternion measurements) may be added before using Eqs. (12) and (13) to get the measurement Davenport matrix at any given epoch.

The SOAR Davenport matrix may now be computed through Eq. (85). Use any solution method to the Wahba problem to solve for the optimal attitude. With the optimal attitude found, use Eq. (81) to compute the update to the parameter vector $\boldsymbol{\beta}$. Finally, compute the updated covariance matrix by taking the inverse of the Fisher information matrix found using Eqs. (92–94).

The state vector may be propagated to the next measurement epoch using Eq. (99), and the covariance matrix may be propagated using Eq. (107). Once at the next epoch, the procedure is repeated.

VI. Comparison of SOAR with Other Attitude Filters

A. Comparison of the SOAR Filter with the Multiplicative Extended Kalman Filter

The central idea behind the MEKF is to use the unconstrained three-parameter representation $\delta\boldsymbol{\theta}$ instead of the constrained four-parameter attitude quaternion in the state vector of a regular EKF [9,10,19]. This means that the MEKF estimates $\delta\hat{\boldsymbol{\theta}}$ and then updates the attitude quaternion by

$$\bar{\mathbf{q}}^+ \approx \begin{bmatrix} \delta\hat{\boldsymbol{\theta}}/2 \\ 1 \end{bmatrix} \otimes \bar{\mathbf{q}}^- \tag{108}$$

This update only maintains the quaternion unity norm constraint to first order, making it common practice to normalize the a posteriori quaternion, $\bar{\mathbf{q}}^+ = \bar{\mathbf{q}}^+/\|\bar{\mathbf{q}}^+\|$.

An important similarity between the MEKF and the SOAR filter is the objective function used to find the optimal update. Recall that the SOAR filter is a maximum likelihood estimator. Under the assumption that errors are Gaussian, the SOAR filter may also be thought of as a minimum variance filter. Therefore, when the angles are small, the solution to both the MEKF objective function and the SOAR filter objective function are approximately the same, i.e., they both approximately minimize $J = \mathrm{tr}[\mathbf{P}^+]$. Therefore, the MEKF and the SOAR filter are expected to behave in a similar fashion when the errors are small and the MEKF linearization assumptions are good. Simulation confirms that this is the case. A second important similarity is that both the MEKF and SOAR can estimate nonattitude states.

Unlike the MEKF, the SOAR filter deals directly with the nonlinear nature of attitude representations and the Wahba problem. Rather than linearizing the update equation as is required in the MEKF, the SOAR filter explicitly solves the constrained quadratic cost function. This allows the SOAR filter to directly compute the attitude quaternion.

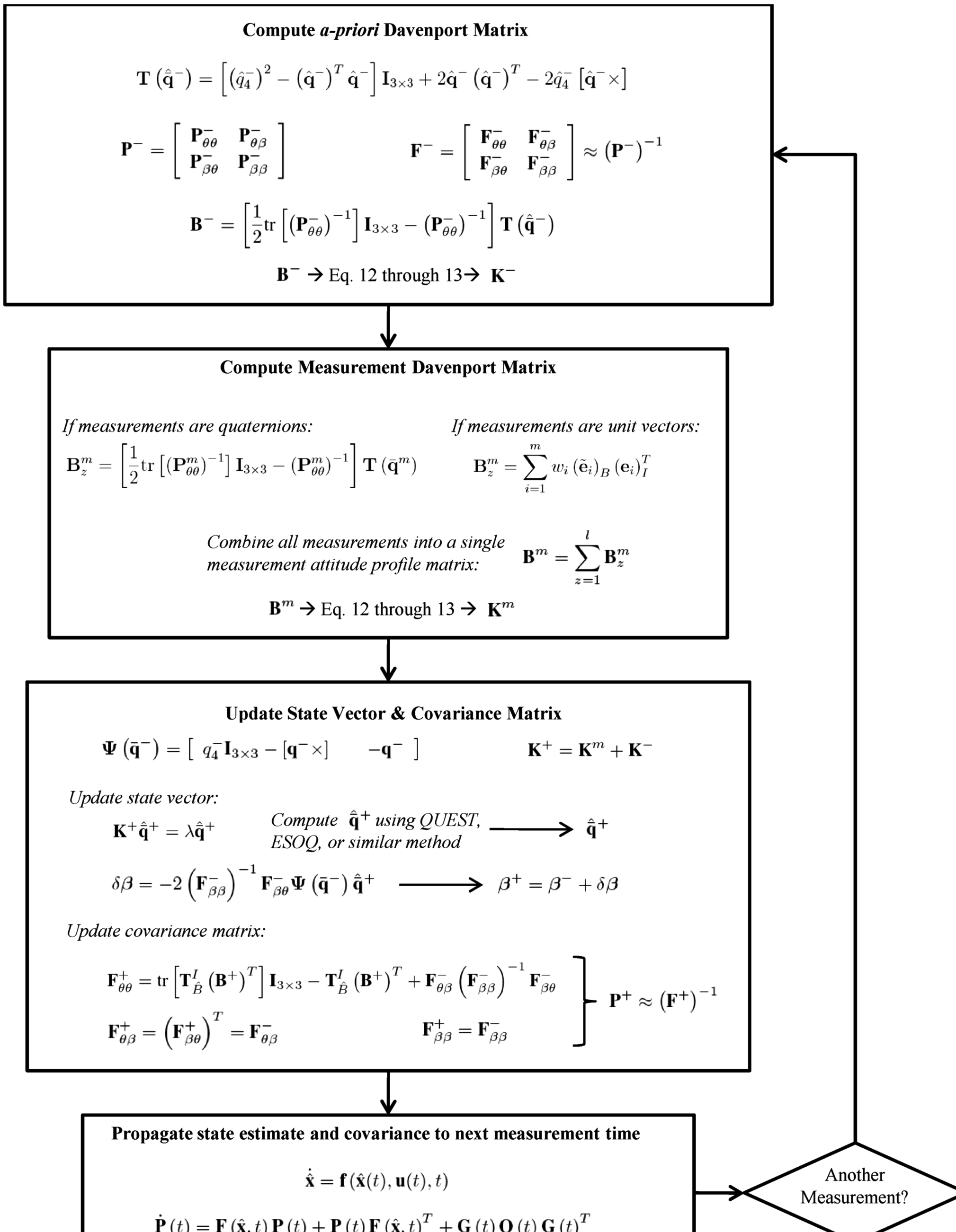


Fig. 1 SOAR filter flowchart.

B. Comparison of SOAR with Filter QUEST and REQUEST

Filter QUEST [11,12] and REQUEST [13] are both based on sequentialization of the Wahba problem. Performing this sequentialization produces the following recursive filter equation (using the same nomenclature as for the SOAR filter) that is called filter QUEST in the literature [11]:

$$\mathbf{B}_{k+1}^{+} = \mathbf{B}_{k+1}^{m} + \alpha_{k+1}\boldsymbol{\Phi}(t_{k+1}, t_k)\mathbf{B}_k^{+} \tag{109}$$

where $\alpha_{k+1} \in [\,0 \quad 1\,]$ and represents reduction in knowledge associated with the propagation. This is a classic fading-memory filter [30], where $\alpha_{k+1} = 0$ causes the filter to ignore the old measurements and $\alpha_{k+1} = 1$ makes the filter think there is no process noise and the measurements are combined as if they were all taken simultaneously. Therefore, heuristically speaking, the fading-memory factor α_{k+1} should look something like $\sum \sigma(t_k)_{\phi,i}^2/\sigma(t_{k+1})_{\phi,i}^2$. Note, however, that the filter QUEST algorithm picks a suboptimal value for α_{k+1} to reduce complexity and increase algorithm speed.

Because the Davenport matrix $\mathbf{K}$ is a linear homogeneous function of attitude profile matrix $\mathbf{B}$, Eq. (109) may be written in terms of $\mathbf{K}$ instead of $\mathbf{B}$. This represents the fundamental difference between the filter QUEST and REQUEST algorithm: filter QUEST is a sequentialization on $\mathbf{B}$ and REQUEST is a sequentialization on $\mathbf{K}$. Shuster provides an excellent discussion of the relation between these two algorithms [14], which are shown to be equivalent when the same selection is made for α_{k+1}. Therefore, the conclusions made in the comparison with filter QUEST also hold for REQUEST.

Although not generally true, it is possible to analytically select the forgetfulness factor [α_{k+1} in Eq. (109)] under specific scenarios. Consider a system with the following assumptions:

1) Only three unit vector measurements are available.

2) Each unit vector measurement lies along one of the coordinate axes with a measurement standard deviation of σ_ϕ.

3) The process noise is a zero-mean white noise process with $\mathbf{Q} = q\mathbf{I}_{3\times3}$.

4) Steady-state filter covariances.

5) The attitude covariance matrix is diagonal, $\mathbf{P}_{\theta\theta} = p\mathbf{I}_{3\times3}$.

Shuster [11,12] demonstrates that under these conditions the optimal α_{k+1} may be analytically computed as

$$\alpha_{k+1} = \frac{\sigma_\phi^2/q + 1 - \sqrt{1 + 2\sigma_\phi^2/q}}{\sigma_\phi^2/q} \tag{110}$$

This simple equation for α_{k+1} provides an excellent approximation of the more general optimal fading-memory factor. Note, however, that a typical star tracker may track many more than three stars and all the unit vector observations will be separated by an angle less than the camera field of view (i.e., it is not possible to obtain measurements along each of the coordinate axes). This means that the α_{k+1} given in Eq. (110) is suboptimal in the general sense.

The SOAR filter removes all five of the assumptions made in computing α_{k+1}. As discussed above, the SOAR filter provides the optimal state update under the assumption that state vector error and measurement errors are normally distributed. The SOAR filter is also capable of estimating other states in addition to attitude (recall that one of the major drawbacks to filter QUEST and REQUEST is their inability to also estimate nonattitude states). These advantages come at the expense of additional computations, however.

C. Comparison of SOAR with Optimal REQUEST

The optimal REQUEST algorithm [15] finds the attitude that minimizes the objective function:

$$\text{Min}\, J_{\text{Opt-RQ}} = \text{tr}[E[\Delta\mathbf{K}(\Delta\mathbf{K})^T]] \tag{111}$$

where $\mathbf{K} = \hat{\mathbf{K}} + \Delta\mathbf{K}$. Recognizing that $\Delta\mathbf{K}$ is symmetric, it is straightforward to show that the 2-norm and Frobenius norm of the matrix $\Delta\mathbf{K}$ are given by

$$\|\Delta\mathbf{K}\|_2^2 = \lambda_{\max}(\Delta\mathbf{K}(\Delta\mathbf{K})^T) = \lambda_{\max}^2(\Delta\mathbf{K}) \tag{112}$$

$$\|\Delta\mathbf{K}\|_F^2 = \text{tr}[\Delta\mathbf{K}(\Delta\mathbf{K})^T] \tag{113}$$

This means that optimal REQUEST minimizes the expected value of the square of the Frobenius norm of $\Delta\mathbf{K}$.

Recalling an important result from eigenvalue–eigenvector stability [33], the estimate of the attitude quaternion may be written as

$$\hat{\bar{\mathbf{q}}} = \bar{\mathbf{q}} + \sum_{j=2}^{4}\left[\frac{\mathbf{u}_j^T\Delta\mathbf{K}\bar{\mathbf{q}}}{\lambda_1 - \lambda_j}\right]\mathbf{u}_j + \text{H.O.T.} \tag{114}$$

where $\mathbf{u}_j$ are the eigenvectors and $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$ are the eigenvalues of the true Davenport matrix, $\mathbf{K}$. Also recall that the true quaternion is associated with the largest eigenvalue of $\mathbf{K}$ (i.e., $\mathbf{u}_1 \equiv \bar{\mathbf{q}}$). Because all $\mathbf{u}_j$ are assumed to have unity norm and $\Delta\mathbf{K}$ must be a symmetric matrix, it is straightforward to show (Jordan decomposition) that $\mathbf{u}_j^T\Delta\mathbf{K}\bar{\mathbf{q}} \leq \lambda_{\max}(\Delta\mathbf{K})$. Because it can also be shown that $\text{tr}[\mathbf{P}_{\theta\theta}] = 4\,\text{tr}[E[(\hat{\bar{\mathbf{q}}} - \bar{\mathbf{q}})(\hat{\bar{\mathbf{q}}} - \bar{\mathbf{q}})^T]]$ to first order, then

$$\text{tr}\,[\mathbf{P}_{\theta\theta}] \leq 4E[\lambda_{\max}^2(\Delta\mathbf{K})]\sum_{j=2}^{4}\sum_{l=2}^{4}\left[\frac{1}{(\lambda_1 - \lambda_j)(\lambda_1 - \lambda_l)}\right]\text{tr}[\mathbf{u}_j\mathbf{u}_l^T] \tag{115}$$

Note that the double summation is only a function of the true attitude and therefore may be moved outside of the expected value operator. Additionally, using the 2-norm identity from Eq. (112), the above may be rewritten as

$$\text{tr}\,[\mathbf{P}_{\theta\theta}] \leq 4E[\|\Delta\mathbf{K}\|_2^2]\zeta(\mathbf{K}) \tag{116}$$

where

$$\zeta(\mathbf{K}_k) = \sum_{j=2}^{4}\sum_{l=2}^{4}\left[\frac{1}{(\lambda_1 - \lambda_j)(\lambda_1 - \lambda_l)}\right]\text{tr}[\mathbf{u}_j\mathbf{u}_l^T] \tag{117}$$

Now, recalling that the 2-norm is always less than or equal to the Frobenius norm, $\|\Delta\mathbf{K}\|_2 \leq \|\Delta\mathbf{K}\|_F$, the above inequality may be extended as

$$\text{tr}\,[\mathbf{P}_{\theta\theta}] \leq 4E[\|\Delta\mathbf{K}\|_2^2]\zeta(\mathbf{K}) \leq 4E[\|\Delta\mathbf{K}\|_F^2]\zeta(\mathbf{K}) \tag{118}$$

The SOAR filter, because it provides a maximum likelihood estimate and assumes normally distributed random variables, provides a minimum variance estimate of the attitude, i.e., the SOAR filter minimizes $J_{\text{SOAR}} = \text{tr}[\mathbf{P}_{\theta\theta}]$. Now, recalling that $J_{\text{Opt-RQ}} = E[\|\Delta\mathbf{K}\|_F^2]$, will show that

$$J_{\text{SOAR}} \leq 4J_{\text{Opt-RQ}}\zeta(\mathbf{K}) \tag{119}$$

Therefore, optimal REQUEST and the SOAR filter provide the minimum to two different objective functions, and are related by Eq. (119). Unfortunately, the inequality in Eq. (119) is not as illuminating as one would hope, and it is important to note that this inequality should not be taken to suggest that the SOAR filter's objective function is better. The nature of this inequality makes it difficult to determine which one of these objective functions provides a solution that is closer to the Cramér–Rao bound. Additional work remains to be done before any conclusive statement can be made about the relative optimality of these two filters in attitude-only filtering. The intent of this discussion is simply to show some initial observations regarding the relation between the objective functions of these two different filters. That being said, the SOAR filter has the advantage that it explicitly addresses the objective of providing a maximum likelihood estimate of the spacecraft attitude. Further, because optimal REQUEST is an extension of the REQUEST algorithm [with optimal α_{k+1} chosen to minimize Eq. (111)], the optimal REQUEST algorithm is also unable to estimate nonattitude states. The SOAR filter can estimate both attitude and nonattitude states.

D. Comparison of the SOAR Filter with Extended QUEST

The extended QUEST algorithm is the closest of the existing attitude filtering methods to the SOAR filter. The extended QUEST algorithm is a solution to the following optimization problem [16]:

$$\begin{aligned} J_{\text{Ext-QUEST}} &= \frac{1}{2}\bar{\mathbf{q}}^T\mathbf{K}^m\bar{\mathbf{q}} + \frac{1}{2}[\mathbf{R}_{qq}(\bar{\mathbf{q}} - \bar{\mathbf{q}}^-)]^T[\mathbf{R}_{qq}(\bar{\mathbf{q}} - \bar{\mathbf{q}}^-)] \\ &\quad + \frac{1}{2}[\mathbf{R}_{\beta q}(\bar{\mathbf{q}} - \bar{\mathbf{q}}^-) + \mathbf{R}_{\beta\beta}(\boldsymbol{\beta} - \boldsymbol{\beta}^-)]^T[\mathbf{R}_{\beta q}(\bar{\mathbf{q}} - \bar{\mathbf{q}}^-) \\ &\quad + \mathbf{R}_{\beta\beta}(\boldsymbol{\beta} - \boldsymbol{\beta}^-)] \end{aligned} \tag{120}$$

where $\mathbf{R}_{xx}$ are the square-root information matrices that come from the left QR factorization of the information matrix at each epoch. The extended QUEST algorithm in Eq. (120) has been rewritten using the same nomenclature as in the SOAR filter to facilitate easy comparison. Strong parallels between the SOAR filter and extended QUEST are immediately evident by comparing the SOAR filter objective function in Eq. (63) with the extended QUEST objective function in Eq. (120). These similarities produce similar results for the state vector update. The extended QUEST parameter vector update is given by

$$\hat{\boldsymbol{\beta}}^+ = \hat{\boldsymbol{\beta}}^- - \mathbf{R}_{\beta\beta}^{-1}\mathbf{R}_{\beta q}[\bar{\mathbf{q}} - \hat{\bar{\mathbf{q}}}^-] \tag{121}$$

which is clearly similar to the SOAR filter parameter vector update from Eq. (88). The attitude update is significantly different from in the SOAR filter, however. Extended QUEST finds the optimal a posteriori attitude through the solution to

$$[\mathbf{K}^m + \mathbf{R}_{qq}^T\mathbf{R}_{qq}]\bar{\mathbf{q}} - \mathbf{R}_{qq}^T\mathbf{R}_{qq}\bar{\mathbf{q}}^- = \lambda\bar{\mathbf{q}} \tag{122}$$

Table 1 Summary of noise levels for SOAR filter validations (scenario 1)

Description	Standard deviation
A priori attitude error	$\sigma_\theta = 0.1$ deg
A priori bias error	$\sigma_{\mathbf{b}} = 0.5$ deg/h
Star observation unit vector error	$\sigma_{\phi,i} = 10$ arcsec
Gyro measurement error	$\sigma_\omega = 0.05$ deg/h
Gyro bias Gauss–Markov error	$\sigma = 0.05$ deg/h correlation time of $\tau = 1$ h

The introduction of the constant vector $-\mathbf{R}_{qq}^T \mathbf{R}_{qq} \bar{\mathbf{q}}^-$ makes the solution to this problem more difficult than the simple eigenvector-eigenvalue problem in Eq. (87) for the SOAR filter.

Note that the extended QUEST algorithm described above is a square-root information filter (SRIF) and therefore requires the use of the square-root information matrices that are obtained through *QR* factorization. Although the SOAR filter is presented as a standard information filter and extended QUEST is presented as an SRIF, either filter may be recast in either form. In selecting the appropriate implementation, recall that the computation of the square-root information matrices in an SRIF requires extra computations, but provides increased numerical precision and stability. This is a standard tradeoff associated with SRIFs [30].

Despite strong similarities, there are noteworthy differences between extended QUEST and the SOAR filter. The most significant difference between the SOAR filter and extended QUEST is how they express the attitude error. Extended QUEST treats the attitude quaternion error from an additive standpoint, rather than a multiplicative standpoint. Although the term *additive* is used here, it is important to note that the extended QUEST algorithm explicitly enforces the quaternion unity norm constraint in generating the quaternion update. Therefore, the update here should not be confused with additive updates that do not preserve the unity norm constraint.

The difference between using an additive or multiplicative quaternion update has two significant implications. The first is that the covariance associated with $\mathbf{R}_{qq}$ is related to $\delta\bar{\mathbf{q}}_A = \bar{\mathbf{q}} - \hat{\bar{\mathbf{q}}}^-$, rather than the multiplicative attitude quaternion error described in Eq. (18): $\delta\bar{\mathbf{q}} = \bar{\mathbf{q}} \otimes \hat{\bar{\mathbf{q}}}^-$. This changes the physical significance of the resulting $\mathbf{P}_{qq}$ in extended QUEST.

The second major implication of using the additive quaternion error is related to the optimal update. The additive representation of the attitude quaternion error requires the solution of a more general objective function. When the first differential of Eq. (120) is set to zero, extended QUEST requires the solution to Eq. (122), which is of the following form:

$$(\mathbf{K} - \lambda \mathbf{I}_{4\times4})\bar{\mathbf{q}} + \mathbf{g} = \mathbf{0} \tag{123}$$

Psiaki [16] demonstrates a robust iterative method for finding the desired solution to this problem. Solving Eq. (123), however, requires more computation than simply finding the largest eigenvalue/eigenvector of $\mathbf{K}$ using a method such as QUEST or ESOQ. This means that solving for the optimal attitude in the SOAR filter [as described in Eq. (87)] will be faster than finding the optimal attitude in extended QUEST.

VII. Example: Estimation of Attitude and Gyro Bias

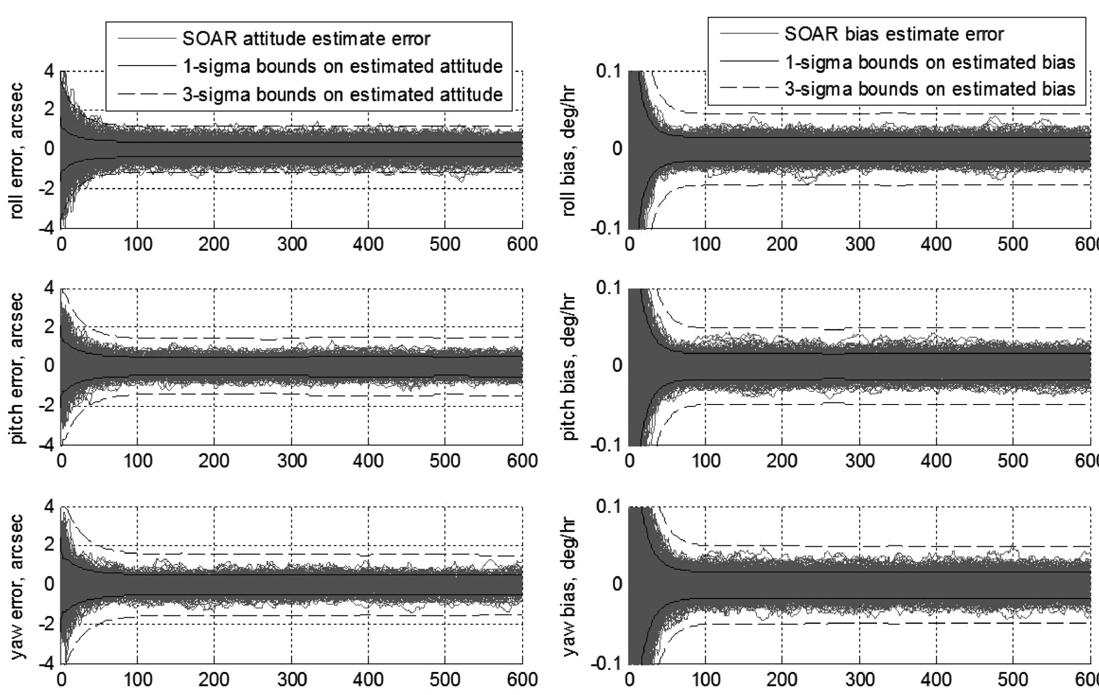
Consider a scenario in which a filter is required to estimate the attitude of a spacecraft and the gyro bias $\mathbf{b}_g$. Rather than have the filter estimate the angular velocity, which would require models for external torques, the gyro measurements of angular velocity are taken to be truth. This is a common practice for spacecraft attitude filters, sometimes referred to as "flying the gyros." In the present analysis, errors associated with gyro scale factor, misalignment, and nonorthogonality are neglected. Therefore, assume that the gyro measurements are given by Farrenkopf's [34] gyro model:

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \mathbf{b}_g + \mathbf{w}_\omega \tag{124}$$

where $\mathbf{b}_g$ is the gyro bias and $\mathbf{w}_\omega$ is zero-mean white noise. Further, assume that the gyro bias behaves as a first-order Gauss–Markov process [30,31]:

$$\dot{\mathbf{b}}_g = -\frac{1}{\tau}\mathbf{b}_g + \mathbf{w}_{\dot{\mathbf{b}}} \tag{125}$$

where τ is the correlation time and $\mathbf{w}_{\dot{\mathbf{b}}}$ is a zero-mean white noise process. It may further be shown that the variance of $\mathbf{w}_{\dot{\mathbf{b}}}$ is $2\sigma^2/\tau$ [30]. Therefore,



a) Attitude estimate error **b) Bias estimate error**

Fig. 2 Monte Carlo performance assessment (1000 runs) of the SOAR filter in scenario 1. Results for the MEKF are very similar.

$$\hat{\boldsymbol{\omega}} = E[\boldsymbol{\omega}] = \tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}}_g \tag{126}$$

$$\dot{\hat{\mathbf{b}}}_g = E[\dot{\mathbf{b}}_g] = -\frac{1}{\tau}\hat{\mathbf{b}}_g \tag{127}$$

Two test cases are considered. The first is a scenario with typical sensor errors and noise. The second scenario is a stressing case that contains very large a priori attitude errors and bias errors. The measurement noise is also increased in the second test scenario.

For the first test scenario, consider a system with performance specifications as described in Table 1. For this scenario, a 1000-run Monte Carlo analysis was performed using the SOAR filter and the results are shown in Fig. 2. A similar analysis was performed using the MEKF and the resulting plots were nearly identical to the results of Fig. 2 and are not shown here, for brevity (the MEKF and SOAR filter plots themselves were visually indistinguishable from each other at the scale shown in Fig. 2). Measurements from two star trackers mounted perpendicular to one another were made available to the filter at a frequency of 1 Hz. As expected for a properly functioning filter, most of the attitude errors and bias errors lie within the $\pm 1\sigma$ bounds and nearly all fall within the $\pm 3\sigma$ bounds.

For the second test scenario, consider a system with performance specifications as described in Table 2. As with the first scenario, the second test scenario also consisted of a 1000-run Monte Carlo analysis and the results are shown in Fig. 3 and 4. This Monte Carlo analysis was repeated twice: once using the SOAR filter and once using the MEKF. To facilitate direct comparison between these two filters, both the SOAR filter and the MEKF were initialized with the same a priori covariance and were subjected to the exact same initial conditions, process noise, observations, and measurement noise for each of the 1000 test cases. Despite being given equivalent a priori information and processing the exact same measurements, significant differences are observed in filter performance and convergence. In both the attitude estimation and the bias estimation, the SOAR filter has lower overshoot (the overshoot is induced by the very poor a priori guess) and faster convergence.

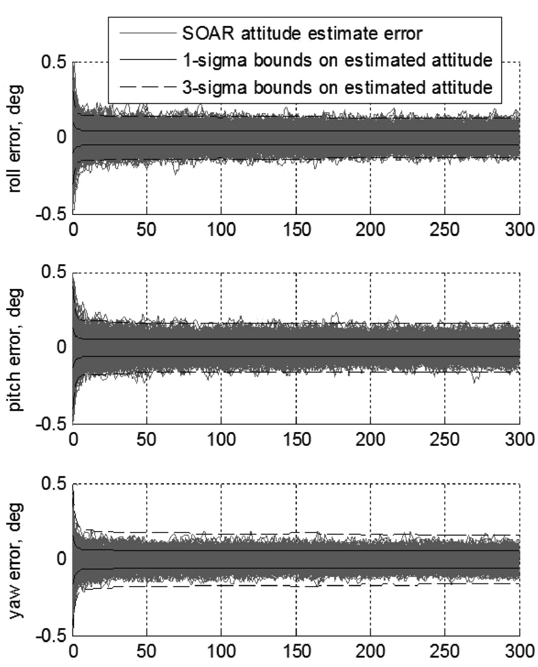
Although all the MEKF cases eventually converged, a large number of these cases remained outside of the covariance bounds even after the MEKF covariance reached steady state. If these covariance bounds were used to estimate filter performance, they would yield overly optimistic results. For this example, the SOAR filter is shown to provide a better reflection of the true covariance throughout the transient and steady-state portions of the simulation.

Table 2 Summary of noise levels for SOAR filter validations (scenario 2)

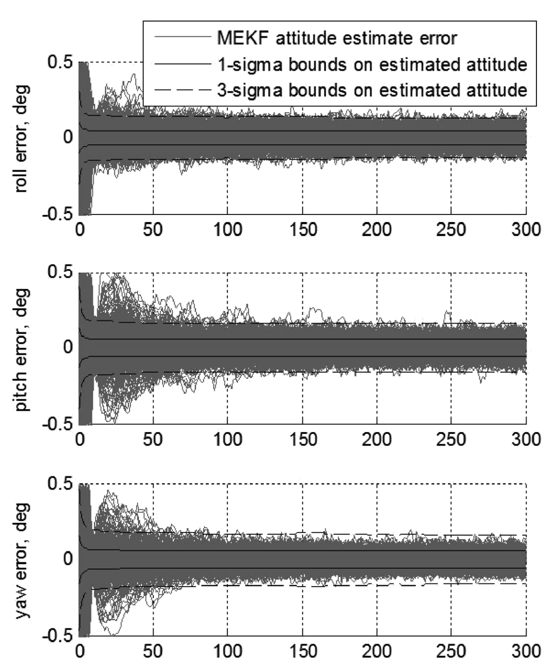
Description	Standard deviation
A priori attitude error	$\sigma_\theta = 40$ deg
A priori bias error	$\sigma_{\mathbf{b}} = 30$ deg/h
Star observation unit vector error	$\sigma_{\phi,i} = 1$ deg
Gyro measurement error	$\sigma_\omega = 0.5$ deg/h
Gyro bias Gauss–Markov error	$\sigma = 0.5$ deg/h correlation time of $\tau = 1$ h

Now consider a single run from this analysis, as is shown in Fig. 5. An interesting observation may be made by also showing the results from filter QUEST on this plot. Note, however, that under these extreme conditions, Filter QUEST is unable to converge. This is because filter QUEST does not estimate gyro bias and the gyro bias is significant in this case. What is interesting here is that the initial behavior of the SOAR filter is similar to that of filter QUEST. Then as the SOAR filter begins to improve the estimate of the gyro bias the SOAR filter solution begins to differ from the filter QUEST solution.

Both the MEKF and the SOAR filter converge for this test case. As was noted before, the SOAR filter and the MEKF approximately minimize the same objective function (under a Gaussian assumption and small angles) and are expected to exhibit similar behavior once the filter converges. The results shown in Fig. 5 support this observation. In some situations (such as roll error and pitch error in Fig. 5), the SOAR filter and MEKF results appear to be nearly identical after both filters converge. In these cases the lines for the two different filter types appear to lie on top of each other. In other situations (such as the yaw error in Fig. 5), a small difference in the attitude estimate produced by the SOAR filter and the MEKF will result in two similar lines that behave nearly identically and have a slowly decreasing offset. This offset eventually disappears and the SOAR filter and MEKF results appear nearly the same. In the case of

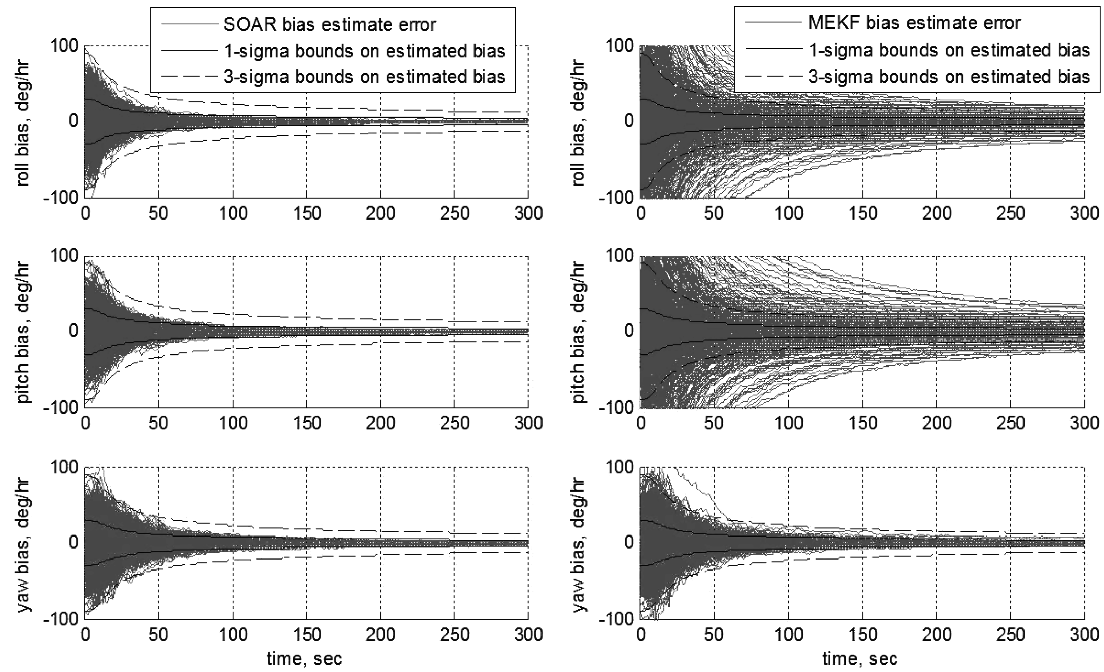


a) Attitude errors with SOAR filter



b) Attitude errors with MEKF

Fig. 3 Monte Carlo performance comparison (1000 runs) of attitude error for the SOAR filter and the MEKF in scenario 2 (stressing case with large errors).



a) Bias errors with SOAR filter

b) Bias errors with MEKF

Fig. 4 Monte Carlo performance comparison (1000 runs) of bias error for the SOAR filter and the MEKF in scenario 2 (stressing case with large errors).

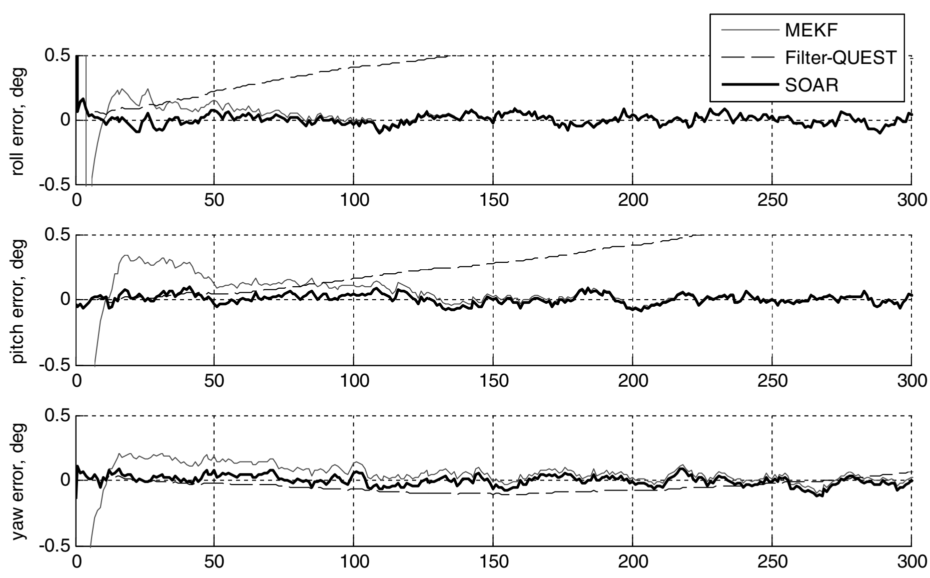


Fig. 5 Performance comparison of the MEKF, filter QUEST, and the SOAR filter in scenario 2 (stressing case with large errors).

the yaw error in Fig. 5, the offset is completely eliminated by about 700 s into the simulation.

VIII. Conclusions

A new nonlinear attitude filter called the sequential optimal attitude recursion (SOAR) filter has been developed. This filter is based on maximum likelihood estimation and is capable of estimating both attitude and nonattitude states. The SOAR filter is shown to provide identical steady-state performance to the MEKF when errors are small, but is frequently capable of providing superior performance when angular errors are large. In its current formulation, however, the SOAR filter is only able to accept two types of measurements: unit vector measurements and complete attitude estimates. The SOAR filter algorithm was compared to numerous other attitude filtering techniques from a theoretical standpoint. Then an example simulation was performed in which the filter was required to estimate attitude and gyro bias. The performance of the SOAR filter was shown to be advantageous relative to previously existing filters in some situations.

Acknowledgments

The authors thank the following individuals for reviewing this manuscript and for providing valuable comments, critiques, and suggestions: Belinda Marchand and Sebastian Muñoz of the University of Texas at Austin; John Crassidis of the University at Buffalo, State University of New York; and Mark Psiaki of Cornell University.

References

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

[2] Keat, J., "Analysis of Least-Squares Attitude Determination Routine DOAOP," Computer Sciences Corp., Rept. CSC/TM-77/6034, Feb. 1977.

[3] Markley, F., and Mortari, D., "Quaternion Attitude Estimation Using Vector Observations," *Journal of the Astronautical Sciences*, Vol. 48, Nos. 2–3, April–Sept. 2000, pp. 359–380.

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

[5] Mortari, D., "ESOQ: A Closed-Form Solution to the Wahba Problem," *Journal of the Astronautical Sciences*, Vol. 45, No. 2, April–June 1997, pp. 195–204.

[6] Mortari, D., "Second Estimator of the Optimal Quaternion," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 5, 2000, pp. 885–888. doi:10.2514/2.4618

[7] Markley, F., "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, Vol. 36, No. 3, July–Sept. 1988, pp. 245–258.

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

[9] Lefferts, E., Markley, F., and Shuster, M., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, Sept.–Oct. 1982, pp. 417–429. doi:10.2514/3.56190

[10] Markley, F., "Attitude Error Representations for Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, March–April 2003, pp. 311–317. doi:10.2514/2.5048

[11] Shuster, M., "A Simple Kalman Filter and Smoother for Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, Jan.–March 1989, pp. 89–106.

[12] Shuster, M., "New Quests for Better Attitudes," *Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, MD, 21–23 May 1991.

[13] Bar-Itzhack, I., "REQUEST: A Recursive QUEST Algorithm for Sequential Attitude Determination," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 5, Sept.–Oct. 1996, pp. 1034–1038. doi:10.2514/3.21742

[14] Shuster, M. D., "Filter QUEST or REQUEST," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 2, March–April 2009, pp. 643–645. doi:10.2514/1.40423

[15] Choukroun, D., Bar-Itzhack, I., and Oshman, Y., "Optimal-REQUEST Algorithm for Attitude Determination," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 3, May–June 2004, pp. 418–425. doi:10.2514/1.10337

[16] Psiaki, M., "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 2, March–April 2000, pp. 206–214. doi:10.2514/2.4540

[17] Markley, F., "Attitude Determination and Parameter Estimation Using Vector Observations: Theory," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, Jan.–March 1989, pp. 41–58.

[18] Markley, F., "Attitude Determination and Parameter Estimation Using Vector Observations: Application," *Journal of the Astronautical Sciences*, Vol. 39, No. 3, July–Sept. 1991, pp. 367–381.

[19] Crassidis, J., Markley, F., and Cheng, Y., "Survey of Nonlinear Attitude Estimation Methods," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 1, 2007, pp. 12–28. doi:10.2514/1.22452

[20] Mortari, D., and Majji, M., "Multiplicative Measurement Model and Single-Point Attitude Estimation," *AAS F. Landis Markley Astronautics Symposium*, Cambridge, MD, 29 June–2 July 2008.

[21] Zanetti, R., "A Multiplicative Residual Approach to Attitude Kalman Filtering with Unit-Vector Measurements," *AAS Flight Mechanics Conference*, San Diego, CA, 14–17 Feb. 2010.

[22] Cheng, Y., Crassidis, J., and Markley, F., "Attitude Estimation for Large Field-of-View Sensors," *Journal of the Astronautical Sciences*, Vol. 54, Nos. 3–4, July–Dec. 2006, pp. 433–448.

[23] Markley, F., "Parameterization of the Attitude," *Spacecraft Attitude Determination and Control*, edited by J. Wertz, D. Reidel, Dordrecht, The Netherlands, 1985, pp. 410–420.

[24] Lerner, G., "Three-Axis Attitude Determination," *Spacecraft Attitude Determination and Control*, edited by J. Wertz, D. Reidel, Dordrecht, The Netherlands, 1985, pp. 420–428.

[25] Hull, D., *Optimal Control Theory for Applications*, Springer, New York, 2003, pp. 52–53.

[26] Mortari, D., "n-Dimensional Cross Product and Its Application to the Matrix Eigenanalysis," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 3, 1997, pp. 509–515. doi:10.2514/3.60598

[27] Sorenson, H., *Parameter Estimation: Principles and Problems*, Marcel Dekker, New York, 1980, pp. 99–100.

[28] Morrison, D., *Multivariate Statistical Methods*, 4th ed., Brooks/Cole, Belmont, CA, 2005, pp. 99–103.

[29] Shuster, M., "Constraint in Attitude Estimation Part 1: Constrained Estimation," *Journal of the Astronautical Sciences*, Vol. 51, No. 1, Jan.–March 2003, pp. 51–74.

[30] Gelb, A., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974, pp. 182–190, 285–288, 306–307.

[31] Brown, R., and Hwang, P., *Introduction to Random Signals and Applied Kalman Filtering: With MATLAB Exercised and Solutions*, 3rd ed., Wiley, New York, 1997, pp. 94–96, 200–202, 343–347.

[32] Crassidis, J., and Junkins, J., *Optimal Estimation of Dynamical Systems*, CRC Press, Boca Raton, FL, 2004, pp. 285–292.

[33] Atkinson, K., *An Introduction to Numerical Analysis*, 2nd ed., Wiley, New York, 1989, pp. 600–601.

[34] Farrenkopf, R., "Analytic Steady-State Accuracy Solutions for Two Common Spacecraft Attitude Estimators," *Journal of Guidance and Control*, Vol. 1, No. 4, 1978, pp. 282–284. doi:10.2514/3.55779