EXTENDED QUEST ATTITUDE DETERMINATION FILTERING

Mark L. Psiaki
Cornell University

ABSTRACT

The quaternion estimation (QUEST) batch attitude determination algorithm has been extended to work in a general Kalman-filter framework. This has been done in order to allow the inclusion of a complicated dynamics model and to allow the estimation of additional quantities beyond the attitude quaternion. The QUEST algorithm, which works with vector attitude observations, serves as a starting point because it is able to work with a poor (or no) first guess of the attitude. It is able to do this because its nonlinear estimation problem can be solved exactly by solving an eigenvalue/eigenvector problem. This paper's extended version of QUEST uses square-root information filtering techniques and linearization of the dynamics to handle all of the non-QUEST parts of the estimation problem. The remaining QUEST-type part of the problem can be solved by a technique that is an extension of the original QUEST algorithm's eigenvalue/eigenvector solution. The paper shows that two previously-proposed iterative QUEST techniques are special cases of the present algorithm. It also demonstrates the new algorithm's performance on an attitude determination problem that uses star-tracker and rate-gyro measurements to estimate the attitude time history and the rate-gyro biases. The new algorithm is able to converge from initial attitude errors of 180^o and initial rate-gyro bias errors as large as $2{,}400^o$/hour.

INTRODUCTION

Most space missions require knowledge of the spacecraft's attitude. This knowledge is normally derived from on-board sensor data. Possible measured quantities include the Sun direction vector, the Earth's magnetic field vector, the Earth nadir vector, the Earth-limb crossing time of a horizon scanner bore sight, the direction vectors to bright stars, and the differential carrier phase of Global Positioning System signals [1,2]. Some of these quantities contain 2-axis attitude information: a unit direction vector measured in spacecraft coordinates and known in inertial coordinates. Other quantities contain only 1 axis worth of information: the cosine of an angle between a spacecraft referenced vector and an inertially-referenced vector.

In order to derive a spacecraft's full 3-axis attitude, 2 or more attitude measurements must be processed together. A variety of methods exist for processing attitude data. These include geometric-based methods [1], extended Kalman filters [3-8], and a special algorithm known as the quaternion estimation (QUEST) algorithm [9,10]. Geometric-based methods can operate on any type of attitude data, but they do not easily make use of redundant data or complex dynamic models. Kalman filters are excellent at handling multiple redundant sensor signals and at incorporating dynamic models and data that has been measured at different times. Unfortunately, an extended Kalman filter can exhibit sensitivity to the initial attitude guess because it relies on linearizations of the spacecraft's nonlinear measurement and dynamics models. In some situations this sensitivity can cause an extended Kalman filter to diverge [6].

The basic QUEST algorithm solves Wahba's problem [11]. Given a set of known unit direction vectors in inertial coordinates, r_i for $i = 1, \ldots, m$, their measured values in spacecraft coordinates, b_i for $i = 1, \ldots, m$, and their per-axis direction uncertainties (in radians), s_i for $i = 1, \ldots, m$, the problem is to

find: $\quad A(q) \qquad$ (1a)

to minimize: $$J_{QUEST}\{A(q)\} = \frac{1}{2}\sum_{i=1}^{m}\frac{1}{s_i^2}\{b_i - A(q)r_i\}^T\{b_i - A(q)r_i\} \qquad (1b)$$

subject to: $\quad q^T q = 1 \qquad$ (1c)

where q is the attitude quaternion for the transformation from inertial coordinates to spacecraft coordinates and $A(q)$ is the direction cosines matrix for that same transformation. The formula for $A(q)$ can be found on p. 414 of Ref. 1.

The QUEST algorithm has advantages in comparison to standard extended Kalman filters. One great advantage is that it can be solved exactly by solving an eigenvalue problem [9]. It can never diverge because this solution procedure does not

Presented at the Flight Mechanics Symposium, NASA Goddard Space Flight Center, Greenbelt, MD, May 18, 1999.

depend on having a first guess. Another advantage is that it explicitly and optimally preserves the attitude quaternion's normalization.

The QUEST algorithm also has disadvantages in comparison to extended Kalman filters. One disadvantage is that it can deal only with vector-type measurements, not with cosine-type measurements. This limits its use to missions where the attitude measurements are all vector-type measurements. A more significant disadvantage of the QUEST algorithm is that it can deal only with very simple dynamic models. The attitude rate time history must be input to QUEST in order for it to use data that has been measured at different times [10]. If this attitude rate time history is derived from rate-gyro measurements, then they cannot have significant biases, which constitutes a severe limitation.

A related disadvantage is the QUEST algorithm's inability to estimate anything other than the attitude quaternion. In its original form, it cannot be used to estimate quantities such as sensor misalignments, rate-gyro biases, or other typical filter states. Therefore, QUEST cannot be used as part of a general attitude determination filter. A linearized version of the QUEST measurement equation can be used in a generalized filter [5], but linearization makes any such algorithm prone to diverge if the initial attitude uncertainty is large.

It would be a great advantage if the QUEST algorithm could be extended to handle an arbitrary dynamic model and the estimation of states other than the attitude quaternion. With such extensions, the QUEST algorithm could be applied using Euler's equations to estimate the attitude rates or using realistic rate gyros to measure them. If Euler's equations were used, then the attitude rates would be estimated as part of the filter state. If rate gyros were used, then the filter's state could include estimates of the rate gyros' biases.

References 12 and 13 document an attempt to extend QUEST to include estimation of other parameters. That approach is based on solving the following extended problem:

find: q and x (2a)

$$\text{to minimize:}\quad J(q,x) = \frac{1}{2}\sum_{i=1}^{m}\frac{1}{s_i^2}\{b_i(x) - A(q)r_i(x)\}^T\{b_i(x) - A(q)r_i(x)\} + \frac{1}{2}(x - x^0)^T W^0 (x - x^0) \tag{2b}$$

$$\text{subject to:}\quad q^T q = 1 \tag{2c}$$

where x is a vector of additional parameters, x^0 is it's *a priori* value, and W^0 is a symmetric positive-semidefinite weighting matrix. A problem with dynamics and rate-gyro measurements can be posed in this form [13]. The rate-gyro biases are estimated as part of the x vector.

Reference 12 develops a solution algorithm for problem (2a)-(2c). It works with guesses of the optimal x and solves exactly for the corresponding optimal q by using the QUEST procedure. An outer loop improves the x guesses by numerical iteration. The algorithm is a batch algorithm, and it showed no significant advantages in comparison to a standard batch algorithm when compared on a test problem [13].

This paper has two goals. One is to extend the QUEST algorithm to include arbitrary dynamics and additional estimated states while retaining QUEST's measurement-error cost function and its explicit constraint of the quaternion norm. The other goal is to develop an iterative QUEST filter rather than a batch filter, one that functions as much as possible like an extended Kalman filter while retaining the good features of the QUEST algorithm.

Achievement of these goals will constitute an advance over the work of Ref. 12. That paper's algorithm cannot handle arbitrary dynamics, and it requires numerical iteration to converge to an estimate of the auxiliary x vector. The new algorithm uses extended-Kalman-filter-type stage-to-stage iterations to achieve convergence to its x estimate. This type of iteration is normally much faster than batch-filter iteration because this type computes problem function gradients only once per stage.

This paper also represents an advance in the area of preserving quaternion normalization within the context of an extended Kalman filter. References 4, 5, and 7 develop special techniques to preserve quaternion normalization within a linearizing extended Kalman filter. They do not explicitly enforce the quaternion normalization constraint during the attitude update. Rather, they use ways that implicitly enforce the constraint [5], or they develop ways to re-normalize the quaternion after the update [4,7]. The present paper explicitly enforces the quaternion normalization in an optimal manner.

The extended QUEST algorithm is presented and analyzed in the four main sections of this paper. The second section reviews the QUEST algorithm and its associated quadratically-constrained quadratic optimization procedure. It then presents and solves an extended quadratically-constrained quadratic problem that is compatible with the extended QUEST algorithm. The third section presents the extended QUEST filtering problem statement and the algorithm that solves the problem. The fourth section shows how to compute the estimation error covariance of the filter. The last main section presents test-case results that are based on data from a simulated truth model.

REVIEW AND EXTENSION OF QUEST SOLUTION MATHEMATICS

Original QUEST Solution

QUEST's efficient solution of Wahba's problem hinges on the fact that the cost function in eq. (1b) can be written as a quadratic form in q:

$$J_{QUEST}\{A(q)\} \;=\; \sum_{i=1}^{m} \frac{1}{s_i^2} \;+\; \frac{1}{2} q^T H_{meas} q \qquad (3)$$

where the symmetric Hessian matrix in eq. (3) is [9,10,12]:

$$H_{meas} \;=\; \sum_{i=1}^{m} \frac{1}{s_i^2} \begin{bmatrix} \{ I(b_i^T r_i) - r_i b_i^T - b_i r_i^T \} & -(b_i \times r_i) \\ -(b_i \times r_i)^T & -b_i^T r_i \end{bmatrix} \qquad (4)$$

The minimization of $J\{A(q)\}$ in eq. (3) subject to the quaternion normalization constraint, $q^T q = 1$, constitutes a quadratically-constrained quadratic program. If one adjoins the quaternion normalization constraint to the cost function using the Lagrange multiplier $l/2$, then differentiation of the resulting Lagrangian function with respect to q leads to the following optimality necessary condition:

$$(H_{meas} + lI)q \;=\; 0 \qquad\qquad \text{or} \quad H_{meas}\, q \;=\; -l\, q \qquad (5)$$

From the right-hand version of eq. (5), it is plain that q is a normalized eigenvector of H_{meas} and that $-l$ is the corresponding eigenvalue. The optimal solution to Wahba's problem is achieved when q corresponds to the $-l$ value that is the smallest (the most negative) eigenvalue of H_{meas}.

A Generalized Quadratic Program

The extended QUEST filter needs to be able to solve a slightly more general quadratically-constrained quadratic program:

find: $\quad q$ (6a)

to minimize: $\quad J(q) \;=\; \frac{1}{2} q^T H q + g^T q \;+ constant$ (6b)

subject to: $\quad q^T q \;=\; 1$ (6c)

In this formulation, H is the cost function's Hessian matrix, and g is the cost function's gradient vector at $q = 0$.

There are two differences between this problem and the quaternion optimization form of Wahba's problem. Both of them arise from the inclusion of *a priori* information at the given sample instant. The first difference is that there is a linear cost term, $g^T q$. The second difference is that the Hessian matrix, H, is no longer the H_{meas} matrix given in eq. (4). Instead, H will be a combination of H_{meas} and an *a priori* term.

Solution of the Generalized Quadratic Program

Problem (6a)-(6c) can be solved by forming a Lagrangian and deriving optimality necessary conditions [14]. In this case, eq. (5) generalizes to become:

$$(H + lI)q + g \;\; = \; 0 \tag{7}$$

This equation can be solved for q, $q = \; -(H + lI)^{-1} g$, and the result can be substituted into constraint (6c) to yield a scalar equation in the scalar unknown l:

$$g^T (H + lI)^{-2} g \;\; = \; 1 \tag{8}$$

If one multiplies both sides of eq. (8) by the square of the determinant of $(H + lI)$, then the resulting equation is an 8th-order polynomial in l, and the optimal l can be determined by solving that polynomial. The global minimum of problem (6a)-(6c) occurs at the l that is the largest (most positive) real solution of eq. (8) [14]. This l value is the only real solution to eq. (8) that is greater than or equal to the negative of the minimum eigenvalue of H, which guarantees that the Hessian of the Lagrangian function, $H+lI$, will at least be positive-semidefinite.

An efficient solution procedure for eq. (8) makes use of an eigenvalue decomposition of H:

$$H \; = \; V\begin{bmatrix} -l_1 & 0 & 0 & 0 \\ 0 & -l_2 & 0 & 0 \\ 0 & 0 & -l_3 & 0 \\ 0 & 0 & 0 & -l_4 \end{bmatrix} V^T \qquad \text{and} \begin{bmatrix} g_{z1} \\ g_{z2} \\ g_{z3} \\ g_{z4} \end{bmatrix} \; = \; V^T g \tag{9}$$

where $V = [v_1, v_2, v_3, v_4]$ is a matrix of orthogonal eigenvectors, $-l_1 \;^3 -l_2 \;^3 -l_3 \;^3 -l_4$ are the four eigenvalues, and g_{z1}, g_{z2}, g_{z3}, and g_{z4} are the components of a transformed gradient vector. Using the V transformation and the notation in eq. (9), eq. (8) can be rewritten in the following form:

$$f(l) \; = \; \frac{g_{z1}^2}{(\,l - l_1\,)^2} \; + \; \frac{g_{z2}^2}{(\,l - l_2\,)^2} \; + \; \frac{g_{z3}^2}{(\,l - l_3\,)^2} \; + \; \frac{g_{z4}^2}{(\,l - l_4\,)^2} \; - \; 1 \; = \; 0 \tag{10}$$

Figure 1 shows a typical plot of $f(l)$ vs. l. The negatives of the eigenvalues of H are marked on the bottom of the plot as l_1, l_2, l_3, and l_4, and the optimal solution to eq. (8) is marked as l_{opt}. Notice that it occurs at the highest value of l for which $f(l) = 0$. Notice, also, that there is an infinite peak in the plot at each of the l_i values.

There is an efficient solution procedure that exploits the form of the curve in Fig. 1. It takes advantage of the facts that $l_{opt} \;^3\; l_4$ and that $f(l)$ is monotonically decreasing for $l \;^3\; l_4$. These characteristics make it possible for any efficient numerical scalar equation solver, such as the guarded secant method, to determine l_{opt} in very few iterations.

The solution of the original QUEST problem is a special case of this solution technique. As g_{z4} approaches 0, the $f(l)$ spike at $l = l_4$ becomes infinitely narrow, and l_{opt} will approach l_4 if g_{z1}, g_{z2}, and g_{z3} are sufficiently small. All of the g_{zi} values are zero in the original QUEST algorithm; so, $l_{opt} = l_4$ in this case. When $l_{opt} = l_4$, the matrix $(H + lI)$ is singular, but eq. (7) still has a solution. In fact, it has multiple solutions. An optimal solution is determined by inverting the nonsingular part of $(H + lI)$ in eq. (7) to solve for the part of q that is a linear combination of the eigenvectors v_1, v_2, and v_3. The solution is completed by adding the term $a\, v_4$ to q and by selecting a to be large enough to satisfy the normality constraint in eq. (6c). There will be positive and negative values of a that satisfy the quaternion normalization, and both solutions will be global minima to problem (6a)-(6c) [14]. In the case of QUEST, the two q solutions, $\pm v_4$, are equivalent estimates of the attitude because $A(q) = A(-q)$.

EXTENDED QUEST FILTER ALGORITHM

Filtering Problem Statement

The extended QUEST filtering problem statement is defined for a single stage of a sampled data or discrete-time system. As will be shown, its solution leads to a natural method for iteration when dealing with multiple-stage systems. The problem is stated as a least-squares optimization problem, in keeping with the original Wahba formulation:

find: $q_{(k)}$ and $x_{(k)}$ $\{$and $q_{(k-1)}$, $x_{(k-1)}$, and, $w_{(k-1)}\}$ (11a)

$$\text{to minimize: } J = \frac{1}{2}\sum_{i=1}^{m_{(k)}} \frac{1}{s_{i(k)}^{2}} \left\{ b_{i(k)} - A[q_{(k)}]r_{i(k)} \right\}^T \left\{ b_{i(k)} - A[q_{(k)}]r_{i(k)} \right\}$$

$$+ \frac{1}{2}\left\{ R_{ww(k-1)} w_{(k-1)} \right\}^T \left\{ R_{ww(k-1)} w_{(k-1)} \right\} + \frac{1}{2}\left\{ R_{qq(k-1)}[q_{(k-1)} - \hat{q}_{(k-1)}] \right\}^T \left\{ R_{qq(k-1)}[q_{(k-1)} - \hat{q}_{(k-1)}] \right\}$$

$$+ \frac{1}{2}\left\{ R_{xq(k-1)}[q_{(k-1)} - \hat{q}_{(k-1)}] + R_{xx(k-1)}[x_{(k-1)} - \hat{x}_{(k-1)}] \right\}^T \left\{ R_{xq(k-1)}[q_{(k-1)} - \hat{q}_{(k-1)}] + R_{xx(k-1)}[x_{(k-1)} - \hat{x}_{(k-1)}] \right\}$$

$$+ \textit{constant} \tag{11b}$$

subject to:

$$q_{(k)} = F\{ t_{(k)}, t_{(k-1)}; q_{(k-1)}, x_{(k-1)}, w_{(k-1)} \} q_{(k-1)} \tag{11c}$$

$$x_{(k)} = f_x\{ t_{(k)}, t_{(k-1)}; q_{(k-1)}, x_{(k-1)}, w_{(k-1)} \} \tag{11d}$$

$$q_{(k)}^T q_{(k)} = 1 \tag{11e}$$

The quantities in the above problem statement are defined as follows: q is the attitude quaternion, x is the vector of auxiliary filter states, and w is the process noise vector. The subscript $[\ ]_{(k)}$ refers to sample instant k, which occurs at time $t_{(k)}$, and the subscript $[\ ]_{(k-1)}$ refers to sample instant $k-1$, which occurs at time $t_{(k-1)}$ $(< t_{(k)})$. Just as in the usual QUEST cost function, the measured vectors $b_{i(k)}$ for $i = 1,\dots,m_{(k)}$ are the attitude reference unit direction vectors as measured in spacecraft coordinates at sample time $t_{(k)}$, the unit vectors $r_{i(k)}$ for $i = 1,\dots,m_{(k)}$ are the known inertial directions of the measured vectors, and the $s_{i(k)}$ standard deviations are the per-axis accuracies of the $b_{i(k)}$ measurements. The vectors $\hat{q}_{(k-1)}$ and $\hat{x}_{(k-1)}$ are the *a posteriori* (or best) estimates of q and x at sample time $t_{(k-1)}$. The matrices $R_{ww(k-1)}$, $R_{qq(k-1)}$, $R_{xq(k-1)}$, and $R_{xx(k-1)}$ are weights that penalize the differences between $q_{(k-1)}$, $x_{(k-1)}$, and, $w_{(k-1)}$ and their *a posteriori* estimates at sample time $t_{(k-1)}$.

Equations (11c) and (11d) constitute the filter's dynamic model. The *4´4* matrix $F\{ t_{(k)}, t_{(k-1)}; q_{(k-1)}, x_{(k-1)}, w_{(k-1)} \}$ is the orthogonal state transition matrix from time $t_{(k-1)}$ to time $t_{(k)}$ that is associated with the quaternion's kinematic differential equation:

$$\dot{q} = \frac{1}{2}\begin{bmatrix} 0 & w_3(t) & -w_2(t) & w_1(t) \\ -w_3(t) & 0 & w_1(t) & w_2(t) \\ w_2(t) & -w_1(t) & 0 & w_3(t) \\ -w_1(t) & -w_2(t) & -w_3(t) & 0 \end{bmatrix} q \tag{12}$$

In this equation $[w_1(t); w_2(t); w_3(t)] = w\{t; t_{(k-1)}, q_{(k-1)}, x_{(k-1)}, w_{(k-1)}\}$ is the attitude rate vector during the time interval $t_{(k-1)}$ to $t_{(k)}$. As shown in this formula, $w(t)$ may depend on the quaternion, the auxiliary state vector, or the process noise vector at time $t_{(k-1)}$. The specific form of this dependence will be dictated by the specific dynamic model that is used in the filter.

The remainder of the dynamic model consists of the discrete-time auxiliary state transition function $f_x\{ t_{(k)}, t_{(k-1)}; q_{(k-1)}, x_{(k-1)}, w_{(k-1)} \}$. This is a vector function whose result has the same dimension as the auxiliary state vector x. This discrete-time function may be the result of numerical integration of auxiliary dynamic differential equations from time $t_{(k-1)}$ to time $t_{(k)}$, or it may be directly defined in discrete-time. The former situation holds if the x vector contains attitude rate estimates that get propagated via numerical integration of Euler's equations for an attitude dynamics model of the spacecraft. The latter situation may hold for a spacecraft that has rate gyros whose biases are estimated as part of x. If the auxiliary state vector is propagated between samples by numerical integration of a differential equation, then it will usually be best to integrate that equation and eq. (12) simultaneously.

It may be useful to pose a filter problem that has no measurements associated with some of the problem stages. In other words, $m_{(k)} = 0$ would hold for some (but not all) values of k. This might be needed because of the way in which process noise enters this model. It is modeled as a discrete-time process, $w_{(k-1)}$, $w_{(k)}$, $w_{(k+1)}$, … One might want to use this discrete-time process noise to approximate the effects of a continuous-time white process noise. If the time period between actual attitude measurements is too large, then one needs to add extra "pseudo measurement" times between the actual attitude measurements. These extra "pseudo samples" break up the interval between measurements into multiple intervals over which different $w_{(k)}$ act. This modeling trick prevents the effective continuous-time process noise from losing its whiteness by having too large of a correlation time.

The problem in eqs. (11a)-(11e) is closely related to a combination of the square-root information filtering update and propagation problems of Ref. 15. Square-root information filters are normally developed by using least-squares estimation techniques, but they also admit a statistical interpretation. The statistical interpretation of the present filter is as follows: $R_{ww(k-1)}$ is the square root of the *a priori* information matrix for the random $w_{(k-1)}$ vector, and $[R_{qq(k-1)}, 0; R_{xq(k-1)}, R_{xx(k-1)}]$ is the square root of the *a posteriori* information matrix for the state estimate $[\hat{q}_{(k-1)}; \hat{x}_{(k-1)}]$.

The above problem form only admits vector attitude measurements, but it can easily be extended to include general attitude measurements. The attitude measurements appear in the first least-squares cost term on the right-hand side of eq. (11b). If other measurements were available, such as the cosine of the angle between a known inertial reference vector and a known spacecraft instrument vector, then an additional cost term would need to be added to the above problem statement to penalize the difference between the measured value of the cosine and the filter's best estimate of it. In that case, the filtering algorithm that will be presented below would need to linearize the resulting measurement equation. It would deal with such measurements just like any standard extended Kalman filter deals with such measurements. In this situation there would be no benefit from using the current filter beyond that of optimal maintenance of quaternion normalization, but there would also be no harm in using it.

New Filtering Algorithm

The new filtering algorithm is an iterative procedure that is made up of two phases per sample period. The algorithm that is developed here is an extended square-root information filtering algorithm [15], which forms its estimates by using least-squares-type procedures. Its first phase is a state and square-root information matrix propagation step. This phase starts with the *a posteriori* estimates at stage $k-1$ and dynamically propagates them in order to compute the *a priori* estimates at stage k. This phase is carried out as in an extended Kalman filter, which means that the dynamics get linearized about stage $k-1$'s *a posteriori* state estimate in order to propagate the information matrices' square roots. The second phase of the process is the measurement update. This is the process of combining the *a priori* information at stage k with the vector attitude measurements at that stage in order to produce the best (or *a posteriori*) estimates of $q_{(k)}$ and $x_{(k)}$ given all of the available data up to that stage. This latter phase of the filtering process is where this paper makes its contributions.

Dynamic Propagation Phase

Although not new to this paper, the state propagation phase is presented here for purposes of completeness and in order to define notation for use later in the paper. The goal of propagation is to eliminate $q_{(k-1)}$, $x_{(k-1)}$, and $w_{(k-1)}$ from problem (11a)-(11e). The propagation procedure uses the constraints in eqs. (11c) and (11d) to eliminate $q_{(k-1)}$ and $x_{(k-1)}$. It next eliminates $w_{(k-1)}$ by partially optimizing the resulting cost function with respect to that quantity. This yields a cost function that depends only on $q_{(k)}$ and $x_{(k)}$. This cost is only an approximation of the original cost function because of the linearization of the dynamics about stage $k-1$'s *a posteriori* state estimate.

The propagation procedure first determines $\tilde{q}_{(k)}$ and $\tilde{x}_{(k)}$, the *a priori* estimates of $q_{(k)}$ and $x_{(k)}$. It does this by nonlinear propagation of eqs. (11c) and (11d):

$$\tilde{q}_{(k)} = F\{t_{(k)}, t_{(k-1)}; \hat{q}_{(k-1)}, \hat{x}_{(k-1)}, 0\}\,\hat{q}_{(k-1)} \tag{13a}$$

$$\tilde{x}_{(k)} = f_x\{t_{(k)}, t_{(k-1)}; \hat{q}_{(k-1)}, \hat{x}_{(k-1)}, 0\} \tag{13b}$$

Note that $w_{(k-1)} = 0$ is used in this propagation because zero is its *a priori* expected value.

The filter next develops a linearized dynamic model. This model takes the form:

$$\begin{bmatrix} Dq_{(k)} \\ Dx_{(k)} \end{bmatrix} = \begin{bmatrix} F_{qq(k-1)} & F_{qx(k-1)} \\ F_{xq(k-1)} & F_{xx(k-1)} \end{bmatrix} \begin{bmatrix} Dq_{(k-1)} \\ Dx_{(k-1)} \end{bmatrix} + \begin{bmatrix} G_{q(k-1)} \\ G_{x(k-1)} \end{bmatrix} w_{(k-1)} \tag{14}$$

where the perturbations are defined to be $Dq_{(k)} = q_{(k)} - \tilde{q}_{(k)}$, $Dx_{(k)} = x_{(k)} - \tilde{x}_{(k)}$, $Dq_{(k-1)} = q_{(k-1)} - \hat{q}_{(k-1)}$, and $Dx_{(k-1)} = x_{(k-1)} - \hat{x}_{(k-1)}$. The state-transition and process noise effectiveness matrices in eq. (14) are Jacobians of the dynamics equations, (11c) and (11d):

$$F_{qq(k-1)} = F + \left[\frac{\partial F}{\partial q}\right]\hat{q}_{(k-1)}, \qquad F_{qx(k-1)} = \left[\frac{\partial F}{\partial x}\right]\hat{q}_{(k-1)}, \qquad G_{q(k-1)} = \left[\frac{\partial F}{\partial w}\right]\hat{q}_{(k-1)} \tag{15a}$$

$$F_{xq(k-1)} = \frac{\partial f_x}{\partial q}, \qquad F_{xx(k-1)} = \frac{\partial f_x}{\partial x}, \qquad G_{x(k-1)} = \frac{\partial f_x}{\partial w} \tag{15b}$$

with all of the partial derivatives evaluated at the point $[q_{(k-1)};x_{(k-1)};w_{(k-1)}] = [\,\hat{q}_{(k-1)};\hat{x}_{(k-1)};0]$.

The final operations of the propagation step are to form a large information matrix and to left QR-factorize it:

$$Q \begin{bmatrix} \tilde{R}_{qq(k)} & 0 & 0 \\ \tilde{R}_{xq(k)} & \tilde{R}_{xx(k)} & 0 \\ \tilde{R}_{wq(k)} & \tilde{R}_{wx(k)} & \tilde{R}_{ww(k-1)} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} R_{qq(k-1)} & 0 \\ R_{xq(k-1)} & R_{xx(k-1)} \end{bmatrix} \begin{bmatrix} F_{qq(k-1)} & F_{qx(k-1)} \\ F_{xq(k-1)} & F_{xx(k-1)} \end{bmatrix}^{-1} \begin{bmatrix} I & 0 & -G_{q(k-1)} \\ 0 & I & -G_{x(k-1)} \end{bmatrix} \\ \begin{bmatrix} 0 & 0 & R_{ww(k-1)} \end{bmatrix} \end{bmatrix} \tag{16}$$

where Q is an orthogonal matrix. The matrices $\tilde{R}_{qq(k)}$, $\tilde{R}_{xx(k)}$, and $\tilde{R}_{ww(k-1)}$ are square matrices, and $\tilde{R}_{xx(k)}$ and $\tilde{R}_{ww(k-1)}$ are both nonsingular. This orthogonal transformation can be carried out using techniques that are standard to square-root information filtering [15].

The result of the propagation step is a modified form of the cost function in eq. (11b):

$$\begin{aligned} J = {} & \tfrac{1}{2}\sum_{i=1}^{m_{(k)}} \frac{1}{s^2_{i(k)}} \left\{b_{i(k)} - A[q_{(k)}]r_{i(k)}\right\}^T \left\{b_{i(k)} - A[q_{(k)}]r_{i(k)}\right\} + \tfrac{1}{2}\left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\}^T \left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\} \\ & + \tfrac{1}{2}\left\{\tilde{R}_{xq(k)}[q_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \tilde{x}_{(k)}]\right\}^T \left\{\tilde{R}_{xq(k)}[q_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \tilde{x}_{(k)}]\right\} \end{aligned} \tag{17}$$

This modified cost function is an approximation of the cost in eq. (11b) if the dynamic constraints in eqs. (11c) and (11d) are nonlinear, but it is exact if they are linear. Note that eq. (17) assumes that $w_{(k-1)}$ is set to its optimal value:

$$\{w_{(k-1)}\}_{opt} = -\tilde{R}^{-1}_{ww(k-1)}\left\{\tilde{R}_{wq(k)}[q_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{wx(k)}[x_{(k)} - \tilde{x}_{(k)}]\right\}. \tag{18}$$

<u>Measurement Update Phase</u>

The measurement update procedure solves a quadratically-constrained quadratic program. In order to form the quadratic program, it uses eqs. (3) and (4) to express the QUEST-type squared measurement-error cost terms as a quadratic form in $q_{(k)}$. Given that the resulting measurement error Hessian matrix is $H_{meas(k)}$, the measurement update problem becomes:

find: $q_{(k)}$ and $x_{(k)}$ (19a)

to minimize:

$$\begin{aligned} J = {} & \tfrac{1}{2} q^T_{(k)} H_{meas(k)} q_{(k)} + \tfrac{1}{2}\left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\}^T \left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\} \\ & + \tfrac{1}{2}\left\{\tilde{R}_{xq(k)}[q_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \tilde{x}_{(k)}]\right\}^T \left\{\tilde{R}_{xq(k)}[q_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \tilde{x}_{(k)}]\right\} + constant \end{aligned} \tag{19b}$$

subject to: $q^T_{(k)} q_{(k)} = 1$ (19c)

Optimization problem (19a)-(19c) can be solved in two steps, one that finds the optimum $x_{(k)}$ as a function of $q_{(k)}$ and the other that optimizes $q_{(k)}$. The $x_{(k)}$ auxiliary state vector does not enter constraint (19c). Therefore, it can be optimized by setting $\partial J/\partial x_{(k)}$ equal to zero and solving the resulting linear equation for $x_{(k)}$. The optimal value is

$$\{x_{(k)}\}_{opt} = \tilde{x}_{(k)} - \tilde{R}_{xx(k)}^{-1}\tilde{R}_{xq(k)}[q_{(k)} - \tilde{q}_{(k)}] \tag{20}$$

Substitution of the optimal $x_{(k)}$ from eq. (20) into the eq.-(19b) cost function yields the following least-squares optimization problem for determining the measurement update:

find: $q_{(k)}$ (21a)

to minimize: $J = \frac{1}{2} q_{(k)}^T H_{meas(k)} q_{(k)} + \frac{1}{2}\left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\}^T\left\{\tilde{R}_{qq(k)}[q_{(k)} - \tilde{q}_{(k)}]\right\} + constant$ (21b)

subject to: $q_{(k)}^T q_{(k)} = 1$ (21c)

This problem is in the same general form as the quadratically-constrained quadratic program in eqs. (6a)-(6c). This equivalence is evident if one defines: $H = H_{meas(k)} + \tilde{R}_{qq(k)}^T\tilde{R}_{qq(k)}$ and $g = -\tilde{R}_{qq(k)}^T\tilde{R}_{qq(k)}\tilde{q}_{(k)}$.

The measurement update is completed by solving problem (6a)-(6c) with the H matrix and g vector given above. The solution to this problem is the *a posteriori* estimate of the attitude quaternion at sample k, $\hat{q}_{(k)}$. This value of $q_{(k)}$ is next substituted into eq. (20) to compute the *a posteriori* auxiliary state vector estimate at sample k, $\hat{x}_{(k)} = \tilde{x}_{(k)} - \tilde{R}_{xx(k)}^{-1}\tilde{R}_{xq(k)}[\hat{q}_{(k)} - \tilde{q}_{(k)}]$.

In order for the algorithm to be recursive, the measurement update procedure must re-express the cost function in eq. (19b) in terms like those of eq. (11b) that constitute the *a posteriori* cost function at sample k-1 – the 3rd and 4th terms on the right-hand side of eq. (11b). This cost function must be a quadratic form in $[q_{(k)} - \hat{q}_{(k)}]$ and $[x_{(k)} - \hat{x}_{(k)}]$.

It is possible to derive a cost function in the required form that is equivalent to the cost in eq. (19b) on the manifold $q_{(k)}^T q_{(k)} = 1$:

$$\begin{aligned} J = \; & \tfrac{1}{2}[q_{(k)} - \hat{q}_{(k)}]^T\left\{H_{meas(k)} + \tilde{R}_{qq(k)}^T\tilde{R}_{qq(k)} + l_{(k)}I\right\}[q_{(k)} - \hat{q}_{(k)}] \\ & + \tfrac{1}{2}\left\{\tilde{R}_{xq(k)}[q_{(k)} - \hat{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \hat{x}_{(k)}]\right\}^T\left\{\tilde{R}_{xq(k)}[q_{(k)} - \hat{q}_{(k)}] + \tilde{R}_{xx(k)}[x_{(k)} - \hat{x}_{(k)}]\right\} + constant \end{aligned} \tag{22}$$

Note that $l_{(k)}$ in this equation is two times the optimal Lagrange multiplier for the quaternion normalization constraint. Equivalence between this cost function and the eq.-(19b) cost function on the manifold is sufficient for the purpose of recursive filtering because all optimizations at successive sample instants are effectively constrained to this manifold, at least to first order in the perturbations $[q_{(k)} - \hat{q}_{(k)}]$ and $[x_{(k)} - \hat{x}_{(k)}]$.

The cost function in eq. (22) can be derived by making use of the optimality necessary conditions for the measurement update,

$$\tilde{R}_{xx(k)}^T\left\{\tilde{R}_{xq(k)}[\hat{q}_{(k)} - \tilde{q}_{(k)}] + \tilde{R}_{xx(k)}[\hat{x}_{(k)} - \tilde{x}_{(k)}]\right\} = 0 \tag{23a}$$

$$\left\{H_{meas(k)} + \tilde{R}_{qq(k)}^T\tilde{R}_{qq(k)} + l_{(k)}I\right\}\hat{q}_{(k)} - \tilde{R}_{qq(k)}^T\tilde{R}_{qq(k)}\tilde{q}_{(k)} = 0 \tag{23b}$$

and the following condition

$$\tfrac{1}{2}[q_{(k)} - \hat{q}_{(k)}]^T[q_{(k)} - \hat{q}_{(k)}] = -\hat{q}_{(k)}^T q_{(k)} + 1 \tag{24}$$

which can be derived from the quaternion normalization constraint. The detailed derivation of eq. (22) is straightforward, but it has been omitted for the sake of brevity.

One last step completes the preparations for recursive application of the algorithm at an incremented value of k. This step is to express the eq.-(22) cost as a sum of squares:

$$J = + \tfrac{1}{2}\Big\{R_{qq(k)}[q_{(k)} - \hat{q}_{(k)}]\Big\}^T\Big\{R_{qq(k)}[q_{(k)} - \hat{q}_{(k)}]\Big\}$$
$$+ \tfrac{1}{2}\Big\{R_{xq(k)}[q_{(k)} - \hat{q}_{(k)}] + R_{xx(k)}[x_{(k)} - \hat{x}_{(k)}]\Big\}^T\Big\{R_{xq(k)}[q_{(k)} - \hat{q}_{(k)}] + R_{xx(k)}[x_{(k)} - \hat{x}_{(k)}]\Big\} + constant \tag{25}$$

where $R_{qq(k)}$ is a matrix square root: $R^T_{qq(k)}R_{qq(k)} = \{H_{meas(k)} + \tilde{R}^T_{qq(k)}\tilde{R}_{qq(k)} + l_{(k)}I\}$ and where $R_{xq(k)} = \tilde{R}_{xq(k)}$ and $R_{xx(k)} = \tilde{R}_{xx(k)}$. The matrix square root $R_{qq(k)}$ is guaranteed to exist because the matrix $\{H_{meas(k)} + \tilde{R}^T_{qq(k)}\tilde{R}_{qq(k)} + l_{(k)}I\}$ is at least positive-semidefinite at the global minimum of problem (21a)-(21c), and it may be positive definite. This matrix square root can be computed using the eigenvalue/eigenvector decomposition of $\{H_{meas(k)} + \tilde{R}^T_{qq(k)}\tilde{R}_{qq(k)}\}$ which is part of the solution procedure for a quadratically-constrained quadratic program that is presented in the second section of this paper.

Comparison with other QUEST Algorithms.

The present algorithm is a generalization of the QUEST algorithms presented in Refs 9 and 10. It reduces exactly to the previous algorithms under appropriate problem modeling assumptions. First of all, if no dynamics are assumed in the problem and if there is no auxiliary state vector to estimate, then only the first term on the right-hand side of eq. (11b) remains in the cost function, and constraints (11c) and (11d) disappear. This yields a problem that is exactly Wahba's original problem [11]. The measurement update part of the filter will calculate a $q_{(k)}$ estimate that is the appropriate eigenvector of the cost function's Hessian matrix. This will happen because $g = 0$ will hold in the general quadratically-constrained quadratic program, and its solution algorithm will recognize this as a special case in which it must make use of eigenvectors.

Alternate modeling assumptions yield the recursive QUEST algorithms of Ref. 10. That paper presents two recursive algorithms, one for use with perfect rate-gyro measurements and one for use with noisy rate gyros whose noise is approximately modeled as white noise. In both cases these filters can be re-derived by using problem models that have quaternion dynamics as in eq. (11c), but that have no auxiliary state vector.

In the perfect rate-gyro case, one can show that the cost in eq. (19b) is equivalent to

$$J = \tfrac{1}{2}q^T_{(k)}H_{meas(k)}q_{(k)} + \tfrac{1}{2}[F^T_{qq(k-1)}q_{(k)} - \hat{q}_{(k-1)}]^T\Big\{H_{(k-1)} + l_{(k-1)}I\Big\}[F^T_{qq(k-1)}q_{(k)} - \hat{q}_{(k-1)}] + constant \tag{26}$$

where $H_{(k-1)}$ is the filter cost function's cumulative Hessian matrix at sample k-1. Equation (26) holds because $F_{qq(k-1)}$ equals the orthogonal F matrix of eq. (11c) in this situation and because there is no $w_{(k-1)}$ process noise. The optimality necessary condition at sample k-1, $\{H_{(k-1)} + l_{(k-1)}I\}\hat{q}_{(k-1)} = 0$, can be used to reduce eq. (26) to the following form

$$J = \tfrac{1}{2}q^T_{(k)}\Big\{H_{meas(k)} + F_{qq(k-1)}H_{(k-1)}F^T_{qq(k-1)}\Big\}q_{(k)} + \tfrac{1}{2}\,l_{(k-1)} + constant \tag{27}$$

Note that the bracketed expression in this equation constitutes $H_{(k)}$.

This cost form proves that the present algorithm is equivalent to the first recursive algorithm of Ref. 10 under the stated modeling assumptions. Except for a scale factor, the result in eq. (27) is equivalent to the cost that would be calculated by the algorithm of Ref. 10. The present filter estimates $q_{(k)}$ by minimizing the cost in eq. (26) subject to the normalization constraint. Reference 10's first algorithm minimizes the cost in eq. (27) and, therefore, calculates the same quaternion estimate as does the present algorithm.

The second recursive QUEST algorithm of Ref. 10 can be reproduced by the present filter with slightly different modeling assumptions. Reference 10's second filter has a forgetting factor, $r_{(k-1)}$. It is a number between 0 and 1 and is used to de-weight the cumulative measurement error cost terms up to sample time k-1 before adding them to the new measurement error cost terms at sample time k.

This same forgetting factor effect can be reproduced in the present filter by adding an appropriately modeled process noise, $w_{(k-1)}$. In this model, one assumes rate-gyro measurements that are numerically integrated to determine $F_{qq(k-1)}$, which again equals the orthogonal F matrix of eq. (11c). The discrete-time, white-noise rate-gyro measurement error is a 3-dimensional process noise vector that can produce rotation errors in 3 orthogonal directions. The linearized influence of $w_{(k-1)}$ on the dynamics model is characterized by the 4×3 matrix $G_{q(k-1)}$. The 3 columns of this matrix are orthogonal to each other and to $F\,\hat{q}_{(k-1)}$. If the model also assumes that the process disturbance noise information matrix is

$$R^T_{ww(k-1)}R_{ww(k-1)} \;=\; \left\{\frac{r_{(k-1)}}{1-r_{(k-1)}}\right\}G^T_{q(k-1)}F_{qq(k-1)}\left\{H_{(k-1)} + l_{(k-1)}I\right\}F^T_{qq(k-1)}G_{q(k-1)} \tag{28}$$

then the cost function in eq. (19b) can be reduced to the form

$$J \;=\; \tfrac{1}{2}q^T_{(k)}\left\{H_{meas(k)} + r_{(k-1)}F_{qq(k-1)}H_{(k-1)}F^T_{qq(k-1)}\right\}q_{(k)} + constant \tag{29}$$

Except for a scale factor, this is the same sample-k cost as is used in the 2nd filter of Ref. 10, which proves that this implementation of the present filter is equivalent to it.

THE ESTIMATION ERROR COVARIANCE

There is a direct connection between the square-root information filter that has been developed and the covariance of the optimal estimate. It is well known that there is such a connection for standard square-root information filters [15]. Three issues must be dealt with in order to generalize the standard results to the present case. One is system nonlinearity, another is the statistical model of the QUEST measurement errors, and the third is the quaternion normalization constraint.

The issue of nonlinearity will be dealt with in the usual way for extended Kalman filters. It will be assumed that accurate calculations can be made using linearizations around *a priori* and *a posteriori* estimates. This assumption will hold so long as the measurement noise and process noise are not "large" compared to the nonlinear terms according to some sensible definition of "large."

The measurement errors can be modeled statistically by means of the following probability density function for the measurements $b_{1(k)}, \ldots, b_{m(k)}$ conditioned on the quaternion, $q_{(k)}$:

$$p[b_{1(k)}, \ldots, b_{m(k)} \,|\, q_{(k)}] \;=\; C\,exp\{-J_{QUEST(k)}[q_{(k)}]\} \tag{30}$$

where C is a constant. This function defines probability density on the manifold $q^T_{(k)}q_{(k)} = 1$. This probability density function is sensible. For each measured b vector this function gives an error probability density function of $C\,exp\{-2sin^2(q/2)/s^2\}$ where q is the angle between the measured b vector and its true direction and where s is the angular error's standard deviation. This density function approaches a Gaussian for small s.

The constraint that the estimate lie on the manifold $q^T_{(k)}q_{(k)} = 1$ is handled in the following way. First, one recognizes that this constraint causes the covariance matrix to be singular, with its null space being the normal to the constraint manifold. Expressed in plain terminology, the estimator knows the quaternion length exactly; so, the variance of the quaternion length is zero. The remainder of the covariance calculations are carried out in the manifold's local tangent space. Although not valid for large uncertainties, this approach is consistent with the linearizing assumptions that are used to deal with all other problem nonlinearities for purposes of calculating covariances.

The rest of the covariance calculation proceeds in a manner analogous to standard square-root information filter theory. One assumes that the *a priori* process noise covariance matrix is $\{R^T_{ww(k-1)}R_{ww(k-1)}\}^{-1}$. The *a posteriori* estimation error covariance at sample k is then the inverse of the projection of the matrix $[R_{qq(k)}, 0; R_{xq(k)}, R_{xx(k)}]^T[R_{qq(k)}, 0; R_{xq(k)}, R_{xx(k)}]$ onto the subspace that is tangent to the quaternion normalization constraint. This covariance is calculated as follows. First, one uses left QR factorization to determine 3 quaternion vectors that are mutually orthogonal to each other and to $\hat{q}_{(k)}$:

$$[\hat{q}_{(k)}, Q_2] \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = Q_{ns} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \hat{q}_{(k)} \tag{31}$$

where Q_{ns} is an orthogonal matrix. The 4×3 matrix Q_2, which forms the last 3 columns of Q_{ns}, is used to perform the projections that are needed in order to calculate the covariance matrix:

$$\begin{bmatrix} \hat{P}_{qq(k)} & \hat{P}_{qx(k)} \\ \hat{P}_{qx(k)}^T & \hat{P}_{xx(k)} \end{bmatrix} = \begin{bmatrix} Q_2 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} Q_2^T (R_{qq(k)}^T R_{qq(k)} + R_{xq(k)}^T R_{xq(k)}) Q_2 & Q_2^T R_{xq(k)}^T R_{xx(k)} \\ R_{xx(k)}^T R_{xq(k)} Q_2 & R_{xx(k)}^T R_{xx(k)} \end{bmatrix}^{-1} \begin{bmatrix} Q_2^T & 0 \\ 0 & I \end{bmatrix} \tag{32}$$

It is straightforward to show that this covariance matrix is singular, having its one zero eigenvalue along the eigenvector direction $[\, \hat{q}_{(k)};0]$.

It can be proved rigorously that the filter's *a posteriori* estimates $\hat{q}_{(k)}$ and $\hat{x}_{(k)}$ are the expected values of $q_{(k)}$ and $x_{(k)}$ conditioned on the measurements up through sample *k* and on the *a priori* statistics. It can also be proved that the formula in eq. (32) gives the correct covariance for the errors in these optimal estimates. The proofs make use of the linearization assumptions; so, they are valid only for "small" uncertainties. The proofs work with various conditional probability density functions that define probabilities on manifolds of the form $q^T q = 1$. These probability density functions all take the form $p(q,x) = C\, exp\{- J(q,x)\}$, where $J(q,x)$ is one of the least-squares cost functions defined above. Although straightforward, the proofs are lengthy. They have been omitted for the sake of brevity.

SIMULATION TESTS OF THE ALGORITHM

The algorithm has been tested using simulated data. There are several reasons for simulation testing. One is to check for any unforeseen difficulties with implementation of the algorithm. Another is to determine whether the algorithm indeed has better convergence properties than a standard extended Kalman filter. Yet a third reason for testing is to check out the practicality of the small-angle assumptions in the covariance analysis.

A baseline extended Kalman filter has also been tested in order to determine whether the new filter has an improved ability to converge. The baseline filter is almost the same as the above filter except that it linearizes the measurement error equations about the *a priori* quaternion estimate before it calculates the measurement error. This leads to a measurement error cost function of the following form:

$$J_{meas}(q) = \frac{1}{2} \sum_{i=1}^{m} \frac{1}{s_i^2} \Big\{ b_i - [A(\tilde{q}) + \frac{\partial A}{\partial q}(q - \tilde{q})] r_i \Big\}^T \Big\{ b_i - [A(\tilde{q}) + \frac{\partial A}{\partial q}(q - \tilde{q})] r_i \Big\} \tag{33}$$

This is exactly what a standard extended Kalman filter does with measurement errors if it is implemented as a square-root information filter. This baseline filter uses the quadratically-constrained measurement update, which optimally enforces quaternion normalization. Therefore, it is slightly more sophisticated than the extended Kalman filters of Refs. 4, 5, and 7.

Filter Design

A relatively simple filtering case has been tested. It assumes the availability of noisy rate-gyro data and star-tracker data. The rate gyro is assumed to have biases. The filter's estimation vector is $[q;w_{bias}]$, where w_{bias} is the estimated rate-gyro bias vector. It constitutes this filter's auxiliary state vector, *x*. The dynamic model of the filter consists of a model for the angular velocity vector between measurement samples and a model for the rate-gyro bias dynamics. The angular velocity model that gets used in eq. (12) is:

$$w\{t;\, t_{(k\text{-}1)},\ \hat{q}_{(k\text{-}1)}, \hat{w}_{bias(k\text{-}1)}, w_{(k\text{-}1)}\} = w_{rg(k\text{-}1)} + \left(\frac{t - t_{(k-1)}}{t_{(k)} - t_{(k-1)}}\right)[w_{rg(k)} - w_{rg(k\text{-}1)}] - \hat{w}_{bias(k\text{-}1)} - w_{a(k\text{-}1)} \tag{34}$$

where $w_{rg(k)}$ is the rate-gyro measurement at sample k and $w_{a(k\text{-}1)}$ constitutes the first 3 elements of the 6×1 process noise vector $w_{(k\text{-}1)}$. The dynamic model of the rate-gyro bias vector is a random walk:

$$w_{bias(k)} = w_{bias(k\text{-}1)} + w_{b(k\text{-}1)} \tag{35}$$

where $w_{b(k\text{-}1)}$ constitutes the last 3 elements of the $w_{(k\text{-}1)}$ process noise vector; i.e., $w_{(k\text{-}1)} = [\, w_{a(k\text{-}1)} ; w_{b(k\text{-}1)} \,]$.

Truth Model

This filter has been tested with data from a simulated truth model. The simulated truth model is that of a rigid-body spacecraft in a low Earth orbit. The truth model simulates Euler's equations and the quaternion kinematics. It includes gravity-gradient torques and white-noise disturbance torques. One scenario tests a spin-stabilized spacecraft that undergoes nutations. It has a spin period of *50* sec and a nutation period of *504* sec. The other scenario tests a nadir-pointing gravity-gradient stabilized spacecraft that librates at frequencies on the order of the orbital frequency.

The truth model includes a model of the star tracker and of the rate gyro. The star tracker is assumed to have a limited field of view; it has only a 5^o radius. For the spin-stabilized spacecraft, the center of the star-tracker's field of view points approximately perpendicular to the nominal spin vector. For the nadir-pointing spacecraft it points towards the nominal zenith direction. Similar to what was done in Ref. 13, star-tracker measurements have been simulated by randomly generating a direction vector in the star tracker's field of view at each measurement sample. That direction and the spacecraft's true attitude have been used to generate the "known" inertial direction vector for that attitude measurement, $r_{i(k)}$. The $b_{i(k)}$ measured vector has been calculated by taking the original randomly generated direction vector and adding a random direction error component that has a Gaussian distribution with a standard deviation of *10* arc sec per axis. In the case of the spin stabilized spacecraft, the star tracker makes one measurement every 15 seconds, which is about 3 times per spin period. For the nadir-pointing spacecraft, the star tracker measures the direction to one star once every 58.5 seconds, or 100 times per orbit.

The measurement model for the rate gyro includes errors due to white noise and errors due to a bias that can drift as a random-walk. Two different intensities have been used for the white-noise component of the rate gyro error: 0.02^o/hour$^{1/2}$ and 0.10^o/hour$^{1/2}$. The intensity of the white noise that drives the bias drift has been set at (0.10^o/hour)/hour$^{1/2}$ for all cases in this paper.

It is important to choose small enough sample rates for the rate gyros. The sample rates used in this study are once every *0.625* sec for the spin-stabilized spacecraft and once every *2.92* sec for the nadir-pointing spacecraft. Lower sampling rates can cause systematic estimation errors due to the truncation error that is inherent in the filter's angular rate model, i.e., in eq. (34).

There are sample times with no measurement update. This happens because of differences between the star-tracker sampling rates and the rate-gyro sampling rates. In the spinning spacecraft case there is one star-tracker measurement for every *24* rate-gyro measurements. In this case, $t_{(k)} - t_{(k\text{-}1)}$ is fixed at *0.625* sec, but $m_{(k)}$ varies. It is *0* for *23* out of *24* samples, and then it is *1* for the 24^{th} sample. The nadir-pointing spacecraft has a $t_{(k)} - t_{(k\text{-}1)}$ of *2.92* sec, and $m_{(k)}$ switches from *0* to *1* every 20^{th} sample.

Results

This filter shows very good performance under nominal conditions and when subjected to large initial attitude uncertainty. As an example, Fig. 2 shows attitude error results for a filtering case that used simulated data for the spinning spacecraft. This case starts with a moderate initial attitude error; the error is 11^o in magnitude and directed about an axis that is 43^o away from the nominal spin axis. Its initial rate-gyro bias error is small, with a magnitude of 1.4^o/hour, and its rate-gyro error's white-noise intensity is 0.02^o/hour$^{1/2}$. Figure 2 shows the total attitude error – the total rotational error between the *a posteriori* attitude estimate and the true attitude. It also shows the filter's predicted *a posteriori* standard deviation for this quantity. It is clear from this plot that the filter does a good job of attitude determination and that its covariance calculations are consistent with the actual errors.

A number of cases have been run in order to test the new filter's ability to converge from large initial attitude errors. In each of these cases, the baseline extended Kalman filter has been used to filter the same data. This provides a point of comparison that allows one to determine whether the new filter can converge in situations where the baseline filter cannot. Figures 3 and 4 present results for one of these cases, a case that considers the nadir-pointing spacecraft. Both filters were given the same large initial errors in the attitude and in the rate-gyro bias estimates. The initial attitude error was 180° about the spacecraft's roll axis, and the initial rate-gyro bias error was 100°/hour about the pitch axis. To make matters even worse, the filter was given erroneous standard deviations for its *a priori* per-axis attitude errors and rate-gyro bias errors, 0.1° and 1°/hour, respectively. The rate-gyro's white-noise error component had an intensity of 0.10°/hour$^{1/2}$ for this case.

The baseline extended Kalman filter fails to converge from this poor first guess, but the extended QUEST attitude filter does very well. The extended QUEST filter converges to an attitude error of less than 3° in its first 500 sec of filtering, and its rate-gyro bias error gets reduced to a magnitude of under 3°/hour after one orbit of filtering.

Even with substantial decreases in the initial errors, the standard extended Kalman filter still has problems. If the initial attitude and rate-gyro bias errors are reduced by *25%* from those used in Figs. 3 and 4, then the standard filter still fails to converge. With a *33.3%* reduction – to an initial attitude error of 120° – the standard filter finally converges, but its attitude error is still 3.8° after one orbit. When the extended QUEST filter is used for this same 120°-case, it reduces the error to less than 3.8° right at the outset and to about 0.1° in less than half an orbit.

In almost all large-initial-error cases considered, the extended QUEST filter displayed better performance than the extended Kalman filter. In many cases the extended Kalman filter failed to converge when the extended QUEST filter succeeded. There was one case where the extended Kalman filter converged and the new filter failed to converge, but this was a spurious case in which the extended Kalman filter converged only by accident. Even when the extended Kalman filter did converge, it almost always took much more time to achieve a good estimate than did the new extended QUEST filter.

Another difference between the two filters is in their sensitivities to tuning when initial errors are large. The extended Kalman filter can be made to diverge or converge by changing the initial *a priori* state error covariance or by changing the rate gyro's white-noise error intensity. In at least one case, the relationship between convergence and the covariance tuning proved to be counter-intuitive. The extended QUEST filter, on the other hand, exhibits insensitivity to changes in covariance tuning.

It must be noted that a poor initial guess can have detrimental effects on the new extended QUEST attitude filter, especially if the initial filter covariances are not set properly. This can be seen if one looks at the filter's predicted estimation error standard deviations for the case that produced Figs. 3 and 4. The initial attitude error is *1,800* times larger than the initial standard deviation for this quantity. The filter does not get the actual attitude error to be less than *10* times the filter's *a posteriori* standard deviation until after *4,500* sec of filtering. The rate-gyro bias behaves the same way: for the whole first orbit, the actual rate-gyro errors remain more than an order of magnitude larger than the filter's *a posteriori* standard deviation. If the measurement and process-noise covariances in the filter were close to the real system's actual values, then this problem would go away after a long time. In the interim, before the covariances settled down to their correct values, this discrepancy would cause the actual errors to be larger than they would have been with better tuning.

Even though the new filter has an increased ability to converge from initial attitude errors, it still can fail to converge if the error in its initial estimate of the auxiliary state vector is too large. With 180° initial attitude errors for the nadir-pointing spacecraft, as in Figs. 3 and 4, the filter successfully converged from an initial rate-gyro bias error magnitude of $2{,}400^\circ$/hour, but an initial error of $4{,}000^\circ$/hour caused divergence. At a $3{,}200^\circ$/hour initial error, the filter did not converge after one orbit, and it was unclear whether it would ever converge. Similar behavior has been found for the spinning spacecraft, where divergence has been observed for initial rate-gyro bias errors above $12{,}000^\circ$/hour.

This demonstrates that the filter may be unable to converge from large initial errors in its auxiliary state vector, x. Any failure to converge is due to the linearization assumption that was made in eq. (14). Probably the only way to ensure that a filter cannot diverge is to do batch filtering with numerical iteration, as in Ref. 12.

The new filter's convergence properties are probably sufficiently robust for almost any mission. Even though the filter can diverge for a wrong initial guess of x, it appears to have a very large domain of convergence in the examples that have been considered. In the nadir-pointing case, convergence has been achieved for an initial rate-gyro bias error that is more than 10 times as large as the orbital rate. In the spinning spacecraft case, the filter successfully converged even when the initial rate-gyro bias error was 48% of the spacecraft's spin rate.

The extended QUEST attitude determination filter is capable of working with a system model that propagates its estimate of the attitude rate vector using Euler's equations. None of the examples included in this section use such a model because of limits to the scope of this effort. If Euler's equations are used, then filter convergence may depend on not having too large of an initial error in q, contrary to what has been found in the above examples. This could happen because of q's effect on gravity-gradient torques or on other terms in Euler's equations. Although the filter might be more prone to diverge, it still might out-perform a standard extended Kalman filter in this regard. Of course, more work is needed in order to investigate this issue.

CONCLUSIONS

A new spacecraft attitude determination filter has been developed. It operates on vector attitude data. Its goal is to incorporate QUEST-type measurement updates into an extended Kalman filter framework in hopes of improving the filter's convergence robustness in the face of large initial attitude errors and nonlinear effects. The extended filter's state vector uses a quaternion attitude parameterization and can include other elements such as angular rates or rate-gyro biases. The filter uses standard square-root information filtering techniques wherever possible. The only exception is that the quaternion part of the measurement update involves the solution of a quadratically-constrained quadratic program, as in the original QUEST algorithm. The quadratic cost function gets modified in this case to include the effects of *a priori* information and the effects of the other state vector elements.

The new algorithm has proved successful at increasing the range of initial attitude uncertainties from which the filter can converge to the true attitude. In an example that involved star-tracker and rate-gyro measurements with rate-gyro bias estimation, the extended QUEST algorithm was able to converge from simultaneous initial errors of 180^o in attitude and $2{,}400^o$/hour in rate-gyro bias. Larger initial rate-gyro bias errors can cause divergence, but for the cases considered, the size of the initial attitude error has no effect on the filter's ability to converge. Such properties will be important to spacecraft missions that require an increased degree of autonomy and, therefore, an increased domain of convergence for the attitude filter.

REFERENCES

1. Wertz, J.R. ed., *Spacecraft Attitude Determination and Control*, D. Reidel Pub. Co., (Boston, 1978).

2. Cohen, C.E., "Attitude Determination," in *Global Positioning System: Theory and Applications, Vol. II*, Parkinson, B.W. and Spilker, J.J. Jr., eds., American Institute of Aeronautics and Astronautics, (Washington, 1996), pp. 519-538.

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

4. Bar-Itzhack, I.Y. and Oshman, Y., "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-21, No. 1, 1985, pp. 128-136.

5. Shuster, M.D., "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *Journal of the Astronautical Sciences*, Vol. 38, No. 3, 1990, pp. 377-393.

6. Psiaki, M.L., Martel, F., and Pal, P.K., "Three-Axis Attitude Determination via Kalman Filtering of Magnetometer Data," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 3, 1990, pp. 506-514.

7. Bar-Itzhack, I.Y., Deutschmann, J., and Markley, F.L., "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," *Proceedings of the AIAA Guidance, Navigation, and Control Conf.*, New Orleans, LA, Aug. 1991, pp. 908-916.

8. Psiaki, M.L., Theiler, J., Bloch, J., Ryan, S., Dill, R.W., and Warner, R.E. “ALEXIS Spacecraft Attitude Reconstruction with Thermal/Flexible Motions Due to Launch Damage”, *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, Sept.-Oct. 1997, pp. 1033-1041.

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

10. Bar-Itzhack, I.Y., "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.

11. Wahba, G., "A Least Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, p. 409.

12. Markley, F.L., "Attitude Determination and Parameter Estimation Using Vector Observations: Theory," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, 1989, pp. 41-58.

13. Markley, F.L., "Attitude Determination and Parameter Estimation Using Vector Observations: Application," *Journal of the Astronautical Sciences*, Vol. 39, No. 3, 1991, pp. 367-381.

14. Sorensen, D.C., "Newton's Method with a Model Trust Region Modification," *SIAM Journal on Numerical Analysis*, Vol. 19, No. 2, 1982, pp. 409-426.

15. Bierman, G.J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, (New York, 1977), pp. 57-67, 69-76, and 115-122.

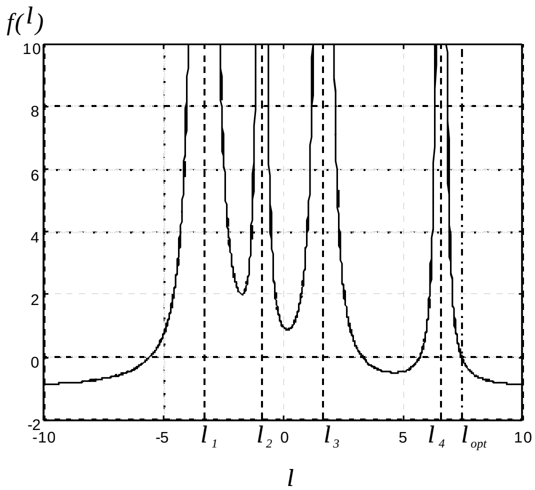

Fig. 1. Constraint function dependence on the Lagrange multiplier in the solution of the quadratically-constrained quadratic program.

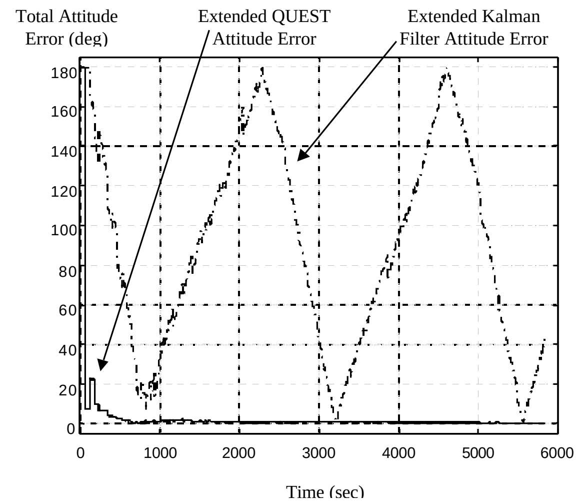

Fig. 3. Comparison of total attitude error time histories for 2 filters that start with a poor a priori state estimate.

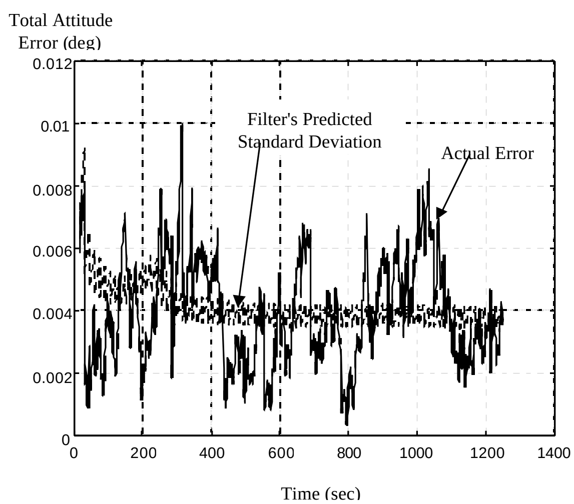

Fig. 2. Total attitude error time history and its predicted standard deviation for the extended QUEST filter operating on data from the spinning spacecraft.

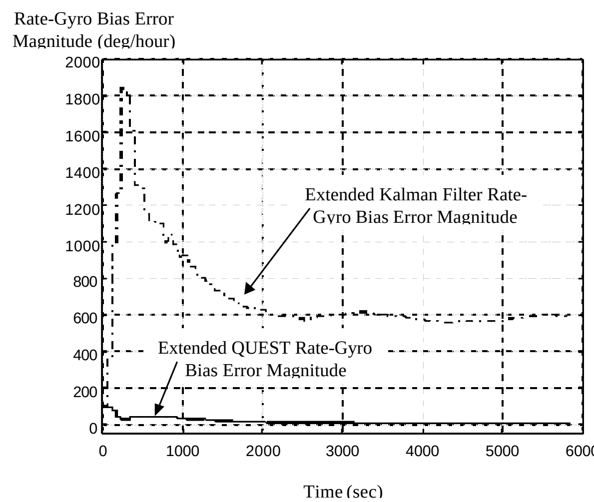

Fig. 4. Comparison of rate-gyro bias error magnitude time histories for 2 filters that start with a poor a priori state estimate.