JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 27, No. 2, March–April 2004

Global Magnetometer-Based Spacecraft Attitude and Rate Estimation

Mark L. Psiaki*
Cornell University, Ithaca, New York 14853-7501

A globally self-initializing attitude determination filter has been developed to estimate three-axis attitude and attitude rate from magnetometer data only. This filter can serve as the main attitude determination system for a low-weight, low-budget, low-accuracy mission or as a backup system for other missions. The filter uses a new attitude representation that is based on the minimum quaternion to align the body-axis and reference magnetic field vectors. It uses implicit quaternion dynamics and implicit Euler dynamics to propagate its attitude and attitude-rate estimates. It is based on a generalization of the iterated extended Kalman-filter concept. Its global initialization strategy employs a number of filters that cover a two-dimensional box in which the principal initial condition uncertainties lie. Hypothesis testing is used to eliminate all but one of the filters after the best one has reached steady state. Flight data from two missions have been used to test the new system. Convergence without the need for initial attitude or rate knowledge has been demonstrated in all cases, and per-axis 3-σ accuracies on the order 5 deg in attitude and 0.03 deg/s in rate have been achieved.

I. Introduction

THE goals of the present work are to develop and test a three-axis attitude and rate determination system that can function autonomously using only vector magnetometer data. The design strategy for this system uses extended-Kalman-filtering (EKF) principles in order to achieve the best possible accuracy. Autonomy requires that the method work reliably without a first guess of the attitude or the rate. This calls for a global initialization scheme that avoids the possibility of divergence that exists for EKFs.

This system can serve as the primary attitude determination subsystem for micro- and nano-satellites because a miniature, low-power magnetometer is the only required sensor. The target 3-σ attitude accuracy is on the order of several degrees, which will suffice for many such missions. This system can also be used as a backup attitude and rate determination system for larger satellites. For example, the Hubble Space Telescope (HST) experienced rate-gyro failures in 1999 that rendered its primary attitude determination system nonfunctional. A number of unsuccessful attempts have been made to use HST data from this time period in traditional attitude filters. This paper's new method has produced the only known successful attitude solution for this HST data. Thus, it could have provided three-axis attitude and rate determination capabilities that would have enabled HST to avoid a series of control anomalies, which resulted in large attitude gyrations.

A number of previous efforts have addressed the problem of magnetometer-based three-axis attitude and rate determination. Psiaki et al.[1] developed a magnetometer-based extended Kalman filter system for attitude and rate estimation for a gravity-gradient-stabilized spacecraft. It used Euler's equation to propagate rates, but it required a reasonable initialization in order to converge. Hashmall et al.,[2] Hashmall and Sedlak,[3] Sedlak and Hashmall,[4] and Sedlak[5] developed batch filters and Kalman filters that use rate gyros to measure rates and that use magnetometers to estimate attitude and rate-gyro biases. These filters achieve very high attitude accuracies for a magnetometer-based system, on the order of arc minutes. This is possible because the accurate rate gyros allow the estimators to average out magnetometer and magnetic field model errors over long periods of time. The approach of Challa et al.[6] and Challa et al.[7] estimates three-axis attitude and rate using only magnetometer data. The estimation system uses an EKF that includes Euler dynamics for rate propagation, and it achieves accuracies on the order of several degrees. Initialization is sometimes handled by a deterministic batch algorithm. The authors report that their filter converges without good a priori attitude or rate knowledge, but they give no guarantee that their initialization techniques will work reliably in all circumstances.

The present work makes three main contributions. First, it develops a new three-parameter attitude representation that isolates the difficult part of the estimation problem. The difficult quantities to estimate are the attitude and rate about the magnetic field vector. The new representation includes a large uncertain rotation about the measured field and small rotations perpendicular to it. It is based on the minimum rotation to align two vectors[8] and on a small rotations model of magnetometer measurement errors.[9]

This paper's second contribution is a specially designed iterated extended Kalman filter. This filter tailors its linearizations and iterations in order to deal effectively and efficiently with large uncertainties in attitude and rate about the measured magnetic field. These large uncertainties are resolved by finding the time histories of the rotation and rotation rate about the field vector that best fit the Euler attitude dynamics equation. Two optional additions to the filter are 1) to include inertia matrix estimation in order to improve the accuracy of the Euler equation and 2) to incorporate incomplete rate-gyro data if available. The second option avoids the limitations of Refs. 2–5, whose filters can function only if full three-axis rate-gyro data are available.

The third contribution is a new global initialization strategy. It uses the new attitude representation to confine the large components of the initial uncertainty to a two-dimensional box embedded within the filter's state space. It uses multiple Kalman filters and statistical hypothesis testing to determine the part of that box which contains the true initial state. Unlike the nonlinear attitude filters of Refs. 10 and 11, that seek only to expand their regions of convergence, the present technique effectively guarantees that the initialization produces filter convergence.

The remainder of this paper is divided into four main sections plus conclusions. Section II presents the system model that gets used in the filter, including the new attitude representation. Section III defines the nonlinear filtering problem and explains the special iterated extended Kalman filter that approximately solves it. This section also describes a related nonlinear smoother. Section IV develops the global initialization strategy. Section V presents test results for the new filter that use simulated data and flight data. Section VI gives the conclusions.

Received 6 March 2003; revision received 3 July 2003; accepted for publication 27 July 2003.

*Associate Professor, Sibley School of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

II. Attitude and Dynamics Model

A. New Three-Parameter Attitude Representation

The new attitude representation presumes the availability of measured and reference magnetic field unit direction vectors. Suppose that the measured field direction vector in spacecraft coordinates is $\hat{\boldsymbol{b}}_{\rm sc}$ and that the corresponding reference vector in inertial coordinates is $\hat{\boldsymbol{b}}_{\rm in}$. The former is measured by the magnetometer, and the latter is calculated using a model and the known spacecraft position. Consider the following attitude quaternion:

$$q_{\min}(\hat{\boldsymbol{b}}_{\rm sc}, \hat{\boldsymbol{b}}_{\rm in}) = \begin{cases} \dfrac{1}{\sqrt{2\left(1+\hat{\boldsymbol{b}}_{\rm sc}^T\hat{\boldsymbol{b}}_{\rm in}\right)}} \begin{bmatrix} \hat{\boldsymbol{b}}_{\rm sc}\times\hat{\boldsymbol{b}}_{\rm in} \\ \left(1+\hat{\boldsymbol{b}}_{\rm sc}^T\hat{\boldsymbol{b}}_{\rm in}\right) \end{bmatrix} & \text{if} \quad -1 < \hat{\boldsymbol{b}}_{\rm sc}^T\hat{\boldsymbol{b}}_{\rm in} \\ \begin{bmatrix} \hat{\boldsymbol{v}} \\ 0 \end{bmatrix} & \text{if} \quad -1 = \hat{\boldsymbol{b}}_{\rm sc}^T\hat{\boldsymbol{b}}_{\rm in} \end{cases} \tag{1}$$

where the vector $\hat{\boldsymbol{v}}$ in the lower condition is any unit direction vector that is perpendicular to $\hat{\boldsymbol{b}}_{\rm sc}$. This spacecraft attitude quaternion is the minimum rotation that satisfies the measurement constraint $\hat{\boldsymbol{b}}_{\rm sc} = A(q)\hat{\boldsymbol{b}}_{\rm in}$ (Ref. 8), where $A(q)$ is the associated direction cosines matrix.[12]

The new spacecraft attitude representation defines additional rotations that get concatenated with $q_{\min}(\hat{b}_{\rm sc}, \hat{b}_{\rm in})$. Suppose that $\hat{\boldsymbol{v}}_1$ and $\hat{\boldsymbol{v}}_2$ are two unit direction vectors chosen so that $[\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \hat{\boldsymbol{b}}_{\rm sc}]$ forms a right-handed orthonormal triad. Then any attitude can be represented by

$$q(\alpha_1, \alpha_2, \theta) = \frac{1}{\sqrt{1+\alpha_1^2+\alpha_2^2}} \begin{bmatrix} (\alpha_1\hat{\boldsymbol{v}}_1 + \alpha_2\hat{\boldsymbol{v}}_2) \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \hat{\boldsymbol{b}}_{\rm sc}\sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix} \otimes q_{\min}(\hat{\boldsymbol{b}}_{\rm sc}, \hat{\boldsymbol{b}}_{\rm in}) \tag{2}$$

where the symbol $\otimes$ denotes quaternion multiplication so that $A(q_a \otimes q_b) = A(q_a)A(q_b)$. The quantities α_1, α_2, and θ parameterize the attitude. The angle θ is a rotation about the measured magnetic field vector. The two parameters α_1 and α_2 are the nonzero elements of the Gibbs vector for a rotation defined in the $[\hat{\boldsymbol{v}}_1, \hat{\boldsymbol{v}}_2, \hat{\boldsymbol{b}}_{\rm sc}]$ coordinate system and constrained to be perpendicular to $\hat{\boldsymbol{b}}_{\rm sc}$ (Ref. 12). In general, $\hat{\boldsymbol{v}}_1$ and $\hat{\boldsymbol{v}}_2$ will vary with time.

The values of α_1 and α_2 will be small compared to one and will approximately equal the magnetometer attitude measurement error half angles, as in Ref. 9. Their smallness ensures that this representation will avoid its singularity, which occurs for any 180-deg rotation perpendicular to $\hat{\boldsymbol{b}}_{\rm sc}$. The smallness of α_1 and α_2 allows the filter equations to be linearized with respect to these quantities without a significant loss of accuracy.

This representation isolates the least accurate component of the attitude estimate in a single parameter θ. The main challenge of magnetometer-based attitude estimation is to determine this angle and its rate based on the resultant fit to Euler's equation. Isolation of the large uncertainties to θ and $\dot{\theta}$ allows an iterated nonlinear Kalman filter to reserve its expensive, complicated nonlinear estimation techniques for application only in a small subspace of the state space.

B. Filter State Vector and Dynamics Model

The extended Kalman filter's nominal state vector contains the three-axis attitude rate and the three attitude parameters

$$\boldsymbol{x} = \left[\boldsymbol{\omega}^T, \alpha_1, \alpha_2, \theta\right]^T \tag{3}$$

where $\boldsymbol{x}$ is the six-dimensional state vector and $\boldsymbol{\omega}$ is the three-dimensional spacecraft angular rate with respect to inertial coordinates but expressed in body axes.

The state vector dynamics are modeled using implicit trapezoidal integration of Euler's equation and the attitude kinematics. Suppose the k subscript refers to quantities at magnetometer sample time t_k and that the $k+1$ subscript refers to quantities at sample time t_{k+1}. Then the trapezoidally integrated Euler's equation in inertial coordinates takes the form:

$$\begin{aligned} A^T[\alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}]\Big\{ & I_m\boldsymbol{\omega}_{k+1} + \boldsymbol{h}_{k+1} - \tfrac{1}{2}(t_{k+1}-t_k) \\ & \times [\boldsymbol{n}_{\rm gg}(\alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}; \boldsymbol{r}_{{\rm sck}+1}, I_m) + \boldsymbol{m}_{k+1}\times\boldsymbol{b}_{{\rm sck}+1}]\Big\} \\ = A^T[\alpha_{1k}, \alpha_{2k}, \theta_k]\Big\{ & I_m\boldsymbol{\omega}_k + \boldsymbol{h}_k + \tfrac{1}{2}(t_{k+1}-t_k) \\ & \times [\boldsymbol{n}_{\rm gg}(\alpha_{1k}, \alpha_{2k}, \theta_k; \boldsymbol{r}_{\rm sck}, I_m) + \boldsymbol{m}_k\times\boldsymbol{b}_{\rm sck}]\Big\} + \Delta\boldsymbol{h}_{\rm dk} \end{aligned} \tag{4}$$

The notation $A[\alpha_1, \alpha_2, \theta]$ is shorthand for the direction cosines matrix $A[q(\alpha_1, \alpha_2, \theta)]$, which transforms from inertial coordinates to spacecraft coordinates. The 3×3 matrix I_m is the spacecraft moment-of-inertia matrix in body-fixed coordinates. The vector $\boldsymbol{h}$ is the net reaction wheel angular momentum with respect to the spacecraft body expressed in body coordinates. The vector function $\boldsymbol{n}_{\rm gg}(\alpha_1, \alpha_2, \theta; \boldsymbol{r}_{\rm sc}, I_m)$ is the gravity-gradient torque given in body coordinates, $\boldsymbol{r}_{\rm sc}$ is the spacecraft position relative to the Earth expressed in inertial coordinates, $\boldsymbol{m}$ is the dipole moment vector of the magnetic torque rods in body coordinates, and $\boldsymbol{b}_{\rm sc}$ is the dimensional magnetic field vector in body coordinates. The vector $\Delta\boldsymbol{h}_{\rm dk}$ is the process noise expressed as an angular momentum impulse. It is modeled as a Gaussian, zero-mean discrete-time white noise sequence. If one neglects $\Delta\boldsymbol{h}_{\rm dk}$, then each side of Eq. (4) represents an approximation of the inertially referenced angular momentum vector at time $(t_k + t_{k+1})/2$.

The attitude kinematics model uses quaternion multiplication. Given a quaternion q, suppose that $\tilde{q}$ is the inverse quaternion, so that $A(\tilde{q}) = A^T(q)$. One computes $\tilde{q}$ by negating the first three elements of q. The quaternion for the spacecraft rotation that occurs during the time interval from t_k to t_{k+1} is

$$\begin{aligned} & q_{\rm rot}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \\ & \quad = q(\alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \otimes \tilde{q}(\alpha_{1k}, \alpha_{2k}, \theta_k) \\ & \quad = \begin{bmatrix} \boldsymbol{q}_{\rm rotv}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \\ q_{\rm rot4}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \end{bmatrix} \end{aligned} \tag{5}$$

where the vector $\boldsymbol{q}_{\rm rotv}$ consists of the first three elements of $q_{\rm rot}$ and $q_{\rm rot4}$ is the fourth element. If the spacecraft rotates less than 180 deg during this interval, then the total rotation angle is

$$\begin{aligned} & \psi(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \\ & \quad = 2\,\text{atan2}\left[\|\boldsymbol{q}_{\rm rotv}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})\|,\right. \\ & \quad \left. |\boldsymbol{q}_{\rm rot4}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})|\right] \end{aligned} \tag{6}$$

and this rotation takes place about the unit direction vector

$$\begin{aligned} & \boldsymbol{e}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \\ & \quad = \frac{\boldsymbol{q}_{\rm rotv}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})}{\|\boldsymbol{q}_{\rm rotv}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})\|} \\ & \quad \times \text{sign}[q_{\rm rot4}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})] \end{aligned} \tag{7}$$

where atan2[,] is the standard two-argument arctangent function and where $\|\ \|$ stands for the Euclidean norm (this usage holds throughout the paper). The quantities defined in Eqs. (6) and (7) can be used to develop the following implicit trapezoidal integration of the kinematics model:

$$\begin{aligned} \frac{\boldsymbol{\omega}_{k+1} + \boldsymbol{\omega}_k}{2} = & \left[\frac{\psi(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1})}{t_{k+1}-t_k}\right] \\ & \times \boldsymbol{e}(\alpha_{1k}, \alpha_{2k}, \theta_k, \alpha_{1k+1}, \alpha_{2k+1}, \theta_{k+1}) \end{aligned} \tag{8}$$

Note that Eq. (8) is well defined even if $\|\boldsymbol{q}_{\rm rotv}\| = 0$. Although the denominator term on the right-hand side of Eq. (7) goes to zero, the rotation angle in Eq. (6) also goes to zero, and the right-hand side of Eq. (8) becomes zero.

The validity of the implicit dynamics models in Eqs. (4) and (8) depends on having a short magnetometer sample interval $t_{k+1} - t_k$. If $t_{k+1} - t_k$ is too large, then the trapezoidal integration approximations will break down. Also, the spacecraft might violate the assumption that it rotates less than 180 deg between samples. The models will be valid if the magnetometer sampling rate is significantly faster than the underlying spacecraft dynamics, which is usually the case. If $t_{k+1} - t_k$ is too large, then it might be possible to alter Eqs. (4) and (8) to use a higher-order integration formula, but such a development is beyond the scope of this paper.

C. Auxiliary Filter States and Their Dynamics Models

The filter state vector can be augmented with auxiliary states, such as estimates of moment-of-inertia matrix parameters, as in Ref. 9, or rate-gyro bias estimates if rate-gyro data are available. The general state vector takes the form

$$\boldsymbol{x} = \left[\boldsymbol{\omega}^T, \alpha_1, \alpha_2, \boldsymbol{x}_{\text{aux}}^T, \theta\right]^T \tag{9}$$

This form keeps θ as the last element in order to facilitate calculations that use square-root information-filtering (SRIF) techniques.

If $\boldsymbol{x}_{\text{aux}}$ includes moment-of-inertia matrix parameters in its first six elements, then the body-axis moment-of-inertia matrix is defined as

$$I_m = A_1(x_{\text{aux4}})A_2(x_{\text{aux5}})A_3(x_{\text{aux6}}) \times \begin{bmatrix} \dfrac{x_{\text{aux2}}^2 + x_{\text{aux3}}^2}{12} & 0 & 0 \\ 0 & \dfrac{x_{\text{aux1}}^2 + x_{\text{aux3}}^2}{12} & 0 \\ 0 & 0 & \dfrac{x_{\text{aux1}}^2 + x_{\text{aux2}}^2}{12} \end{bmatrix} \times A_3^T(x_{\text{aux6}})A_2^T(x_{\text{aux5}})A_1^T(x_{\text{aux4}}) \tag{10}$$

where $A_i(x)$ is the 3×3 direction cosines matrix for a rotation about the ith axis. This model represents the inertia of a uniform rectangular box. The first three elements of $\boldsymbol{x}_{\text{aux}}$ are the lengths of its sides multiplied by the square root of its mass, and the 4th–6th elements of $\boldsymbol{x}_{\text{aux}}$ give its orientation with respect to body coordinates. This parameterization guarantees the physical realizability of all inertia matrix estimates.

The dynamic model of the inertia matrix parameters presumes that they are constants:

$$(x_{\text{auxi}})_{k+1} = (x_{\text{auxi}})_k \qquad \text{for} \qquad i = 1, \ldots, 6 \tag{11}$$

If rate-gyro biases are included in $\boldsymbol{x}_{\text{aux}}$, then they can be modeled as random walks:

$$(x_{\text{auxi}})_{k+1} = (x_{\text{auxi}})_k + w_{\text{rgik}} \qquad \text{for} \qquad i > 6 \tag{12}$$

The present study models these as constants by setting the process noise w_{rgik} to zero. This is reasonable because the data spans considered in this paper are relatively short.

D. Measurement Models

The basic measurement for this system is the magnetic field unit direction vector in spacecraft coordinates $\hat{\boldsymbol{b}}_{\text{sc}}$. The use of this vector in Eq. (2) to define the attitude representation gives rise to the following magnetometer measurement model:

$$0 = \begin{bmatrix} \alpha_{1k} \\ \alpha_{2k} \end{bmatrix} + \boldsymbol{\nu}_{\text{magk}} \tag{13}$$

where the 0 on the left-hand side is the measurement. The 2×1 vector $\boldsymbol{\nu}_{\text{magk}}$ is the measurement noise. It is modeled as a Gaussian, zero-mean discrete-time white-noise sequence. The measurement noise in Eq. (13) includes field model uncertainty in the vector $\hat{\boldsymbol{b}}_{\text{in}}$ and errors in the magnetometer's measured $\hat{\boldsymbol{b}}_{\text{sc}}$ vector as a result of random sensor noise, miscalibration, and cross talk from other spacecraft instruments.

Magnetic field measurement errors can be reduced by calibration of errors as a result of magnetometer misalignment, scale-factor uncertainty, and cross talk and by improved magnetometer design and placement. Some of the calibrations could be incorporated into the estimator that has been developed, but this approach has not been adopted here. Instead, the flight magnetometer data used in Section V has been precalibrated using an attitude-independent calibration algorithm. The algorithm compares the measured magnitude of the magnetic field to the modeled magnitude in order to estimate magnetometer relative misalignments, magnetometer scale factors, and cross talk from magnetic torque rods.

If n_{rg} rate gyros are present, then their measurement model takes the form

$$\boldsymbol{y}_{\text{rgk}} = H_{\text{rg}}\boldsymbol{\omega}_k + \begin{bmatrix} x_{\text{aux7}} \\ \vdots \\ x_{\text{aux}(6+\text{nrg})} \end{bmatrix} + \boldsymbol{\nu}_{\text{rgk}} \tag{14}$$

where $\boldsymbol{y}_{\text{rgk}}$ is the $n_{\text{rg}} \times 1$ vector of rate-gyro outputs in rad/s units, H_{rg} is the $n_{\text{rg}} \times 3$ rate-gyro alignment matrix, and $x_{\text{aux7}}, \ldots, x_{\text{aux}(6+\text{nrg})}$ are the rate-gyro bias states. The $n_{\text{rg}} \times 1$ vector $\boldsymbol{\nu}_{\text{rgk}}$ is the rate-gyro measurement noise, which is modeled as a discrete-time white-noise sequence with a mean of zero and a Gaussian distribution.

This method of including rate-gyro data differs from the usual technique that substitutes the rate-gyro outputs directly into the attitude kinematics after bias estimates have been removed. The use of Euler's equation in conjunction with Eq. (14) allows the filter to operate with fewer than three rate-gyro channels, which is the principal advantage of this new method of incorporating rate-gyro data. If accurate three-axis rate measurements are available, then their small error standard deviation will cause the filter, in effect, to use those measurements in its kinematic model and ignore Euler's equation.

III. Iterated Extended Kalman Filter and Nonlinear Smoother Algorithms

The iterated extended Kalman filter approximately minimizes the following weighted least-squares cost function in order to perform its state propagation from time t_k to time t_{k+1} and its measurement update at time t_{k+1}:

$$J = \frac{1}{2}\|R_{\text{xxk}}[\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k]\|^2 + \frac{1}{2}\|\bar{R}_{\text{hhk}}\Delta\boldsymbol{h}_{\text{dk}}\|^2 + \frac{1}{2}\left\|\frac{2}{\sigma_{\text{mag}}}\begin{bmatrix}\alpha_{1k+1}\\ \alpha_{2k+1}\end{bmatrix}\right\|^2 + \frac{1}{2}\left\|\frac{1}{\sigma_{\text{rg}}}\left(H_{\text{rg}}\boldsymbol{\omega}_{k+1} + \begin{bmatrix} x_{\text{aux7}} \\ \vdots \\ x_{\text{aux}(6+n_{\text{rg}})}\end{bmatrix}_{k+1} - \boldsymbol{y}_{\text{rgk}+1}\right)\right\|^2 \tag{15}$$

where $\hat{\boldsymbol{x}}_k$ is the a posteriori state estimate at time t_k, R_{xxk} is the square root of the inverse of the a posteriori state estimation error covariance matrix at time t_k (otherwise known as the square-root information matrix), $\bar{R}_{\text{hhk}}$ is the a priori square-root information matrix for the process-noise $\Delta\boldsymbol{h}_{\text{dk}}$, σ_{mag} is the magnetometer's direction measurement error standard deviation, and σ_{rg} is the rate-gyro measurement error standard deviation (excluding bias effects).

This cost function is the negative natural logarithm of the state vector's a posteriori conditional probability density function, and the corresponding nonlinear least-squares estimator is a maximum a posteriori estimator. SRIF techniques like those found in Ref. 13 are used to minimize J and to represent the state estimate and its estimation error covariance.

The least-squares cost in Eq. (15) gets minimized by varying the process noise vector $\Delta\boldsymbol{h}_{\text{dk}}$ and the states at sample times t_k and t_{k+1} subject to the constraints that they (approximately) obey the dynamic relationships in Eqs. (4), (8), (11), and (12). The optimization uses an inner procedure that optimizes everything except θ_{k+1} and an outer loop that optimizes this angle. The inner procedure works with a fixed guess of θ_{k+1} and performs the following calculations.

A. Inner Optimization Procedure for Fixed θ_{k+1}

1) Assume the a priori values $\alpha_{1k+1} = \alpha_{2k+1} = 0$, $\boldsymbol{x}_{\text{auxk}+1} = \hat{\boldsymbol{x}}_{\text{auxk}}$, and $\Delta\boldsymbol{h}_{\text{dk}} = 0$, and solve for the a priori $\boldsymbol{\omega}_{k+1}$ from Eq. (4). Use $\boldsymbol{x}_k = \hat{\boldsymbol{x}}_k$ in this calculation.

2) Linearize dynamics equations (4) and (8) about the a posteriori and a priori estimates from step 1. These linearized dynamics equations along with dynamics equations (11) and (12) can be solved for x_k to yield an equation of the form

$$\boldsymbol{x}_k = \Phi_{k+1}(\theta_{k+1})\boldsymbol{x}_{\text{sk}+1} + \Gamma_k(\theta_{k+1})\Delta\boldsymbol{h}_{\text{dk}} + \boldsymbol{\phi}_{k+1}(\theta_{k+1}) \tag{16}$$

where the solution determines the matrix functions $\Phi_{k+1}(\theta_{k+1})$ and $\Gamma_k(\theta_{k+1})$ and the vector function $\boldsymbol{\phi}_{k+1}(\theta_{k+1})$. Their dependence on θ_{k+1} is explicitly noted for use in calculations needed by the outer-loop θ_{k+1} optimization. Note that $\boldsymbol{x}_{\text{sk}+1} = [\boldsymbol{\omega}_{k+1}^T, \alpha_{1k+1}, \alpha_{2k+1}, \boldsymbol{x}_{\text{auxk}+1}^T]^T$, that is, $\boldsymbol{x}_{\text{sk}+1}$ is $\boldsymbol{x}_{k+1}$ without θ_{k+1}.

3) Use Eq. (16) to eliminate $\boldsymbol{x}_k$ from the cost function in Eq. (15). The resulting cost is half the sum of the square errors in the overdetermined system of equations

$$H_{\text{totk}}(\theta_{k+1}) \begin{bmatrix} \Delta\boldsymbol{h}_{\text{dk}} \\ \boldsymbol{x}_{\text{sk}+1} \end{bmatrix} = \boldsymbol{z}_{\text{totk}}(\theta_{k+1}) \tag{17}$$

where the matrix $H_{\text{totk}}(\theta_{k+1})$ and the vector $\boldsymbol{z}_{\text{totk}}(\theta_{k+1})$ are defined as

$$H_{\text{totk}}(\theta_{k+1}) = \begin{bmatrix} \bar{R}_{\text{hhk}} & 0 \\ R_{\text{xxk}}\Gamma_{k+1}(\theta_{k+1}) & R_{\text{xxk}}\Phi_{k+1}(\theta_{k+1}) \\ 0 & \left[0,\ \dfrac{2}{\sigma_{\text{mag}}}I,\ 0,\ 0\right] \\ 0 & \dfrac{1}{\sigma_{\text{rg}}}[H, 0, 0, I] \end{bmatrix} \tag{18a}$$

$$\boldsymbol{z}_{\text{totk}}(\theta_{k+1}) = \begin{bmatrix} 0 \\ R_{\text{xxk}}\{\hat{\boldsymbol{x}}_k - \boldsymbol{\phi}_{k+1}(\theta_{k+1})\} \\ 0 \\ \dfrac{1}{\sigma_{\text{rg}}}\boldsymbol{y}_{\text{rgk}+1} \end{bmatrix} \tag{18b}$$

Note that the last $6 + n_{\text{rg}}$ columns of $H_{\text{totk}}(\theta_{k+1})$ correspond to the inertia matrix states and the rate-gyro bias states. These drop out if the auxiliary states are not included. Similarly, the last n_{rg} rows of $H_{\text{totk}}(\theta_{k+1})$ and $\boldsymbol{z}_{\text{totk}}(\theta_{k+1})$ drop out if rate-gyro measurements are not available.

4) Left QR factorize $H_{\text{totk}}(\theta_{k+1})$:

$$Q_{\text{totk}}(\theta_{k+1}) \begin{bmatrix} R_{\text{hhk}}(\theta_{k+1}) & R_{\text{hsk}+1}(\theta_{k+1}) \\ 0 & R_{\text{ssk}+1}(\theta_{k+1}) \\ 0 & 0 \end{bmatrix} = H_{\text{totk}}(\theta_{k+1}) \tag{19}$$

and use the result to transform Eq. (17) into the form

$$\begin{bmatrix} R_{\text{hhk}}(\theta_{k+1}) & R_{\text{hsk}+1}(\theta_{k+1}) \\ 0 & R_{\text{ssk}+1}(\theta_{k+1}) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{h}_{\text{dk}} \\ \boldsymbol{x}_{\text{sk}+1} \end{bmatrix} = Q_{\text{totk}}^T(\theta_{k+1})\boldsymbol{z}_{\text{totk}}(\theta_{k+1})$$

$$= \begin{bmatrix} \boldsymbol{z}_{\text{hk}}(\theta_{k+1}) \\ \boldsymbol{z}_{\text{sk}+1}(\theta_{k+1}) \\ \boldsymbol{z}_{\text{rk}+1}(\theta_{k+1}) \end{bmatrix} \tag{20}$$

where Q_{totk} is an orthogonal matrix, R_{hhk} and $R_{\text{ssk}+1}$ are square, nonsingular, uppertriangular matrices; $R_{\text{hsk}+1}$ is a matrix; and $\boldsymbol{z}_{\text{hk}}$, $\boldsymbol{z}_{\text{sk}+1}$, and $\boldsymbol{z}_{\text{rk}+1}$ are vectors.

5) Solve Eq. (20) in a least-squares sense to determine $\boldsymbol{x}_{\text{sk}+1}$ and $\Delta\boldsymbol{h}_{\text{dk}}$:

$$\boldsymbol{x}_{\text{sk}+1}(\theta_{k+1}) = [R_{\text{ssk}+1}(\theta_{k+1})]^{-1}\boldsymbol{z}_{\text{sk}+1}(\theta_{k+1}) \tag{21a}$$

$$\Delta\boldsymbol{h}_{\text{dk}}(\theta_{k+1}) = [R_{\text{hhk}}(\theta_{k+1})]^{-1}[\boldsymbol{z}_{\text{hk}}(\theta_{k+1}) - R_{\text{hsk}+1}(\theta_{k+1})\boldsymbol{x}_{\text{sk}+1}(\theta_{k+1})] \tag{21b}$$

B. Least-Squares Cost and the Information Matrix

The results of the inner optimization can be used to calculate several important quantities. One of these quantities is the residual cost:

$$J_r(\theta_{k+1}) = \tfrac{1}{2}\boldsymbol{z}_{\text{rk}+1}^T(\theta_{k+1})\boldsymbol{z}_{\text{rk}+1}(\theta_{k+1}) \tag{22}$$

which uses the residual error vector $\boldsymbol{z}_{\text{rk}+1}$ from Eq. (20).

The transformed sensitivity of the least-squares equations to changes in θ_{k+1} is

$$\begin{bmatrix} \boldsymbol{r}_{h\theta k+1} \\ \boldsymbol{r}_{s\theta k+1} \\ \boldsymbol{r}_{r\theta k+1} \end{bmatrix} = Q_{\text{totk}}^T(\theta_{k+1}) \left\{ \frac{\mathrm{d}H_{\text{totk}}}{\mathrm{d}\theta_{k+1}} \begin{bmatrix} \Delta\boldsymbol{h}_{\text{dk}} \\ \boldsymbol{x}_{\text{sk}+1} \end{bmatrix} - \frac{\mathrm{d}\boldsymbol{z}_{\text{totk}}}{\mathrm{d}\theta_{k+1}} \right\} \tag{23}$$

where the derivatives with respect to θ_{k+1} are evaluated at its current guessed value and where the values of $\Delta\boldsymbol{h}_{\text{dk}}$ and $\boldsymbol{x}_{\text{sk}+1}$ come from Eqs. (21a) and (21b).

The vectors $\boldsymbol{r}_{s\theta k+1}$ and $\boldsymbol{r}_{r\theta k+1}$ from Eq. (23) can be used to complete the a posterior square-root information matrix. If θ_{k+1} is the optimal estimate that globally minimizes $J_r(\theta_{k+1})$, then the a posteriori square-root information matrix for the state estimation error at time t_{k+1} is

$$R_{\text{xxk}+1} = \begin{bmatrix} R_{\text{ssk}+1} & \boldsymbol{r}_{s\theta k+1} \\ 0 & \|\boldsymbol{r}_{r\theta k+1}\| \end{bmatrix} \tag{24}$$

This information matrix is needed in order to iterate the EKF calculation by redefining the Eq. (15) cost for the sample interval that begins at time t_{k+1}.

The vector $\boldsymbol{z}_{\text{rk}+1}$ from Eq. (20) and the vector $\boldsymbol{r}_{r\theta k+1}$ from Eq. (23) can be used to calculate the derivative of $J_r(\theta_{k+1})$ with respect to θ_{k+1}:

$$\frac{\mathrm{d}J_r}{\mathrm{d}\theta_{k+1}} = -\boldsymbol{z}_{\text{rk}+1}^T\boldsymbol{r}_{r\theta k+1} \tag{25}$$

This derivative is useful in the outer-loop optimization of $J_r(\theta_{k+1})$.

C. Outer-Loop Optimization to Complete the Propagation and Measurement Update

An outer-loop minimization of $J_r(\theta_{k+1})$ with respect to θ_{k+1} completes the combined state-propagation/measurement-update procedure. A global one-dimensional minimization procedure gets used in order to derive the best possible estimate of θ_{k+1}. The global minimization starts with M_{k+1} equally spaced guesses of θ_{k+1} that span the range from 0 to 2π. $\theta_{k+1} = 0, 2\pi/M_{k+1}, 4\pi/M_{k+1}, \ldots, 2(M_{k+1}-1)\pi/M_{k+1}$. The cost $J_r(\theta_{k+1})$ is calculated at each of these grid values. The lowest cost is chosen as a candidate for the global minimum, and the two neighboring θ_{k+1} points are used to define a bounded region in which the minimum is known to lie. The chosen value of M_{k+1} must be sufficiently large to enable correct identification of an interval on which the global minimum lies. The value $M_{k+1} = 5$ has been found to suffice in all cases considered in this paper.

The global minimization next performs a guarded search for the minimum by looking in the interval that contains the lowest-cost grid point. This procedure is similar to one that is defined in Ref. 14. It works with a current best point and the endpoints of an uncertainty interval that contains the best point. It uses $\mathrm{d}J_r/\mathrm{d}\theta_{k+1}$ at the best point to determine whether the minimum lies above or below that point, and it computes a new trial point on the appropriate side of the best point. Under nominal conditions the new point is the point where $\mathrm{d}J_r/\mathrm{d}\theta_{k+1} = 0$ as predicted by a linear fit of $\mathrm{d}J_r/\mathrm{d}\theta_{k+1}$ using values at the current best point and at the endpoint of the interval that has the lowest cost. The value of $J_r(\theta_{k+1})$ at the new trial point is used to refine the uncertainty interval. If it is lower than the value at the old best point, then the trial point becomes the new best point, and the old best point becomes an endpoint of the refined uncertainty interval. Otherwise, the old best point remains the best point, and the trial point becomes a new endpoint of the refined

interval. Binomial guarding techniques use a trial point equal to the average of the interval endpoint and the best point if the original trial point is nearer to the interval's endpoint or if the finite difference estimate of $d^2J_r/d\theta_{k+1}^2$ is negative. The guarded search algorithm terminates when the slope at the best point is zero or when the uncertainty interval is smaller than some prespecified tolerance.

The a posteriori state estimate at time t_{k+1} is defined by the optimal value of θ_{k+1} and the corresponding value of $\boldsymbol{x}_{sk+1}$ from Eq. (21a):

$$\hat{\boldsymbol{x}}_{k+1} = \begin{bmatrix} (\boldsymbol{x}_{sk+1})_{opt} \\ (\theta_{k+1})_{opt} \end{bmatrix} \tag{26}$$

This estimate allows the EKF to iterate using a redefined Eq. (15) cost for the sample interval that starts at time t_{k+1}.

D. Iteration of the Inner Procedure

An alternate version of the time-propagation/measurement-update algorithm employs an iteration of the inner optimization procedure. This iteration enables the algorithm to improve its approximation of the global minimum of the nonlinear least-squares estimation problem. It relinearizes the left-hand side of dynamics Eq. (4) about the $\boldsymbol{x}_{sk+1}$ estimate from Eq. (21a). Kinematics equation (8) gets relinearized about this $\boldsymbol{x}_{sk+1}$ and about a new estimate of x_k that gets determined by substituting $\boldsymbol{x}_{sk+1}$ and $\Delta\boldsymbol{h}_{dk}$ from Eqs. (21a) and (21b) into Eq. (16). Next, the relinearized versions of Eqs. (4) and (8) get used to derive an updated version of Eq. (16). Finally, the result gets used in a second pass through steps 3–5 of the inner procedure. This procedure is like that of an iterated extended Kalman filter. Standard iterated EKFs relinearize only the measurement equations, but this procedure also partially relinearizes the dynamics.

This inner iteration complicates the process of determining the derivatives of $H_{totk}(\theta_{k+1})$ and $\boldsymbol{z}_{totk}(\theta_{k+1})$ with respect to θ_{k+1} for use in Eq. (23). The difficulty arises because the θ_{k+1} derivatives of Eq. (16)'s iterated values of $\Phi_{k+1}(\theta_{k+1})$, $\Gamma_k(\theta_{k+1})$, and $\phi_{k+1}(\theta_{k+1})$ must account for the dependence on θ_{k+1} of the estimates about which Eqs. (4) and (8) get relinearized. The derivatives of these estimates with respect to θ_{k+1} can be calculated by differentiating the first iterates of Eqs. (21a) and (21b). This latter calculation requires the θ_{k+1} derivatives of $R_{hhk}(\theta_{k+1})$, $R_{hsk+1}(\theta_{k+1})$, $R_{sskk+1}(\theta_{k+1})$, $\boldsymbol{z}_{sk+1}(\theta_{k+1})$, and $\boldsymbol{z}_{hk}(\theta_{k+1})$. The derivative of $Q_{totk}(\theta_{k+1})$ is required in order to calculate the derivatives of $\boldsymbol{z}_{sk+1}(\theta_{k+1})$ and $\boldsymbol{z}_{hk}(\theta_{k+1})$. The derivatives of $R_{hhk}(\theta_{k+1})$, $R_{hsk+1}(\theta_{k+1})$, $R_{sskk+1}(\theta_{k+1})$, and $Q_{totk}(\theta_{k+1})$ can be determined by differentiating the QR factorization in Eq. (19) and the relationship $Q_{totk}^T(\theta_{k+1})Q_{totk}(\theta_{k+1}) = I$. The details of these calculations have been omitted for the sake of brevity.

E. Rapid Approximate θ_{k+1} Optimization

The computational cost of the outer-loop optimization can be reduced by a factor of about seven or greater if one replaces its global numerical optimization with an EKF-like approximation. Such an approach will be useful for implementation on a flight computer of limited capability.

The simplified outer loop requires a reasonable first guess of θ_{k+1}. Calculation of this guess begins by forming an a priori attitude quaternion estimate at time t_{k+1} using Euler integration of the quaternion kinematics equation:

$$\bar{q}_{k+1} = \begin{bmatrix} \hat{\boldsymbol{\omega}}_k \sin\{\|\hat{\boldsymbol{\omega}}_k\|(t_{k+1}-t_k)/2\}/\|\hat{\boldsymbol{\omega}}_k\| \\ \cos\{\|\hat{\boldsymbol{\omega}}_k\|(t_{k+1}-t_k)/2\} \end{bmatrix} \otimes \hat{q}_k \tag{27}$$

where $\hat{\boldsymbol{\omega}}_k$ and $\hat{q}_k$ are the a posteriori attitude rate and quaternion estimates at time t_k, with $\hat{q}_k$ computed from Eq. (2). The quaternion $\bar{q}_{k+1}$ can be used to define an error quaternion as a function of θ_{k+1}:

$$q_{err}(\theta_{k+1}) = \cos(\theta_{k+1}/2)\, q_{ck+1} + \sin(\theta_{k+1}/2)\, q_{sk+1} \tag{28}$$

where

$$q_{ck+1} = q_{min}(\hat{\boldsymbol{b}}_{sck+1}, \hat{\boldsymbol{b}}_{ink+1}) \otimes \bar{\bar{q}}_{k+1}$$

$$\text{and} \quad q_{sk+1} = \begin{bmatrix} \hat{\boldsymbol{b}}_{sck+1} \\ 0 \end{bmatrix} \otimes q_{ck+1} \tag{29}$$

The a priori θ_{k+1} estimate is the value that minimizes the rotation error in Eq. (28):

$$\bar{\theta}_{k+1} = 2\,\text{atan2}[(q_{sk+1})_4, (q_{ck+1})_4] \tag{30}$$

The EKF-like outer-loop optimization starts with $\bar{\theta}_{k+1}$ from Eq. (30) and executes the inner optimization once. Next, the procedure calculates $\boldsymbol{r}_{s\theta k+1}$ and $\boldsymbol{r}_{r\theta k+1}$ using Eq. (23). These, in turn, are used to calculate the following estimation increments:

$$\Delta\theta_{k+1} = \frac{\boldsymbol{z}_{rk+1}^T \boldsymbol{r}_{r\theta k+1}}{\boldsymbol{r}_{r\theta k+1}^T \boldsymbol{r}_{r\theta k+1}} \tag{31a}$$

$$\Delta x_{sk+1} = -[R_{sskk+1}(\theta_{k+1})]^{-1} \boldsymbol{r}_{s\theta k+1} \Delta\theta_{k+1} \tag{31b}$$

These increments are then used to form the a posteriori state estimate

$$\hat{\boldsymbol{x}}_{k+1} = \begin{bmatrix} \{(\boldsymbol{x}_{sk+1})_{opt} + \Delta\boldsymbol{x}_{sk+1}\} \\ \{\bar{\theta}_{k+1} + \Delta\theta_{k+1}\} \end{bmatrix} \tag{32}$$

where $(\boldsymbol{x}_{sk+1})_{opt}$ is the computed value from Eq. (21a). The square-root information matrix is still computed according to Eq. (24). Finally, the optimal least-squares cost from Eq. (22) gets estimated using the following formula:

$$J_r(\bar{\theta}_{k+1} + \Delta\theta_{k+1}) \cong \tfrac{1}{2}\boldsymbol{z}_{rk+1}^T(\bar{\theta}_{k+1})\boldsymbol{z}_{rk+1}(\bar{\theta}_{k+1}) - \tfrac{1}{2}\Delta\theta_{k+1}^2 \boldsymbol{r}_{r\theta k+1}^T \boldsymbol{r}_{r\theta k+1} \tag{33}$$

This cost estimate is needed by the global initialization procedure.

F. Nonlinear Batch Attitude and Rate Smoother

A nonlinear fixed-interval smoother has also been developed for this estimation problem. It minimizes the cumulative cost function

$$J = \frac{1}{2}\|\bar{R}_{xx0}[\boldsymbol{x}_0 - \bar{\boldsymbol{x}}_0]\|^2 + \frac{1}{2}\sum_{k=0}^{N-1}\|\bar{R}_{hhk}\Delta\boldsymbol{h}_{dk}\|^2 + \frac{1}{2}\sum_{k=0}^{N}\left\{\left\|\frac{2}{\sigma_{mag}}\begin{bmatrix}\alpha_{1k}\\ \alpha_{2k}\end{bmatrix}\right\|^2 + \left\|\frac{1}{\sigma_{rg}}\left(H_{rg}\boldsymbol{\omega}_k + \begin{bmatrix} x_{aux7} \\ \vdots \\ x_{aux(6+n_{rg})}\end{bmatrix}_k - \boldsymbol{y}_{rgk}\right)\right\|^2\right\} \tag{34}$$

subject to the dynamics constraints in Eqs. (4), (8), (11), and (12) for $k = 0, \ldots, N-1$. The quantities $\bar{x}_0$ and $\bar{R}_{xx0}$ in the first term of this cost function are the a priori state estimate and its estimation error square-root information matrix at time t_0.

The smoother uses an iterative Gauss–Newton algorithm.[14] Starting with estimates of $\boldsymbol{\omega}_0, \alpha_{10}, \alpha_{20}, \theta_0, \ldots, \alpha_{1N}, \alpha_{2N}, \theta_N$, and $\boldsymbol{x}_{aux}$, it iterates the kinematic model in Eq. (8) to compute its $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N$ estimates, and it uses the Euler dynamics model in Eq. (4) to compute its $\Delta\boldsymbol{h}_{d0}, \ldots, \Delta\boldsymbol{h}_{dN-1}$ estimates. Increments to the estimates are determined by solving a linear smoothing problem that uses linearizations of Eqs. (4) and (8) about the current estimates. The linear smoothing problem gets solved using techniques like those of Ref. 13. The resulting increments get added to $\boldsymbol{\omega}_0$, α_{10}, α_{20}, $\theta_0, \ldots, \alpha_{1N}$, α_{2N}, θ_N, and $\boldsymbol{x}_{aux}$ after being scaled by an optimization step length, and new estimates of $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N$ and $\Delta\boldsymbol{h}_{d0}, \ldots, \Delta\boldsymbol{h}_{dN-1}$ get computed from the constraints. A step length of one is used if it produces a decrease of the Eq. (34) cost; otherwise, the step length gets halved in an iterative process until the cost decreases. The overall procedure computes a sequence of refined guesses using a sequence of

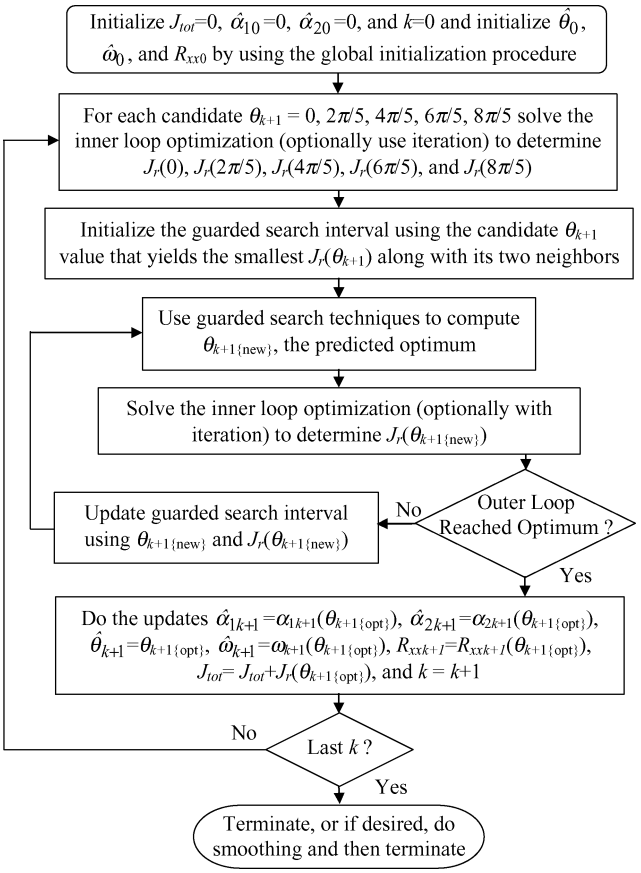

Fig. 1 Flowchart of nominal estimation algorithm.

increments from linearized problems, and it terminates when the optimal increments are very small.

The nonlinear smoother's first guess comes from a nonlinear filter run. The smoother modifies its first guess of $\boldsymbol{\omega}_0$ by optimizing the Eq. (34) cost only with respect to $\boldsymbol{\omega}_0, \ldots, \boldsymbol{\omega}_N$ and $\Delta\boldsymbol{h}_{d0}, \ldots, \Delta\boldsymbol{h}_{dN-1}$ subject to the dynamics constraints, which are linear in these quantities.

G. Summary of Algorithm

The interactions between the principal algorithm components are summarized in the flowchart that appears in Fig. 1. The second block of the flow chart assumes that the outer-loop optimization starts with $M_{k+1}=5$ initial θ_{k+1} grid points. Several blocks in the chart would change slightly if inertia matrix parameters or rate-gyro biases were being estimated. If the rapid θ_{k+1} optimization of Section III.E were used, then the 2nd–5th blocks would be replaced by the procedure of Section III.E, and the first decision branch would be removed along with the block immediately to the left of it.

IV. Statistical Method to Generate Initial Estimates that Ensure EKF Convergence

Extended Kalman filters can diverge if they start with poor initial guesses.[10] This can happen if the filter makes large corrections to its state estimates that invalidate its linearizations. The present section develops a strategy to determine good first guesses of the attitude and rate state estimates. These good first guesses avert the possibility of filter divergence.

A. Two-Dimensional Space of Large Initial Uncertainties

Suppose that θ_0 and θ_1, the rotations about the magnetic field vector at sample times t_0 and t_1, somehow were known. It would be reasonable to guess that the magnetometer measurement errors were zero at these sample times, that is, that $\alpha_{10}=\alpha_{20}=\alpha_{11}=\alpha_{21}=0$. This θ knowledge and these α guesses would constitute full three-axis attitude estimates at t_0 and t_1. The implicit kinematic model in Eq. (8) could be used to derive an initial rate estimate from the attitude estimates under the assumption that $\omega_0=\omega_1$, which would complete the attitude and rate initialization. If inertia parameters, rate-gyro biases, or both were also to be estimated, then good initial estimates of these quantities would have to be available, which is usually the case. The covariance initialization would follow from the uncertainties in the initial attitude parameters.

The problem of finding a good initialization boils down to a problem of finding reasonable estimates of θ_0 and θ_1. Each of these must lie in the range from 0 to 2π. Therefore, the global initialization problem seeks a reasonable initial point in a $2\pi \times 2\pi$ rectangle in (θ_0, θ_1) space.

B. Statistical Discrimination Between Points on a Grid of Possible Initial Conditions

The initialization strategy starts multiple filters from a set of grid points in the (θ_0, θ_1) rectangle and uses a statistical hypothesis test to determine which grid point gives the best initialization. The optimal probabilistic test statistic is the sum of the optimal residual cost values from Eq. (22) [or from Eq. (33) if the simplified outer θ_{k+1} optimization gets used]:

$$J^m_{\text{tot}} = \sum_{k=1}^{N} J^m_r\left([\theta_{k+1}]^m_{\text{opt}}\right) \tag{35}$$

The m superscript gives the index of the filter run that starts from the mth grid point, $[\theta_{k+1}]^m_{\text{opt}}$ is the filter's a posteriori estimate of θ_{k+1}, and N is the number of samples.

The initialization runs M different filters from M different initial grid points for N samples and then decides which is the best. The initialization that produces the lowest value of J^m_{tot} is chosen as the correct one, and the others are discarded. N must be set large enough to make θ_{k+1} sufficiently observable. Observability dictates that the spacecraft must move far enough along its orbit to cause the magnetic field to change its inertial direction by about 45 deg or more.

C. Design of the Two-Dimensional Grid

Certain details about how the grid is formed and about how the square-root information matrix gets initialized can have a significant impact on the efficiency of this method. A good approach is to use a regular θ_0 grid and to develop a unique variable-sized θ_1 grid for each different θ_0 point. A θ_0 grid of size M_0 includes the points 0, $2\pi/M_0$, $4\pi/M_0$, $\ldots$, $2(M_0-1)\pi/M_0$.

The variable-sized θ_1 grids cluster more points around values that yield the smallest initial attitude rate estimates. This strategy makes the reasonable assumption that the magnetometer sampling rate is much faster than twice the spin rate. For a given θ_0 the true θ_1 will not take on a value that maximizes the attitude rate because the maximum rate is nearly half the sampling rate and, therefore, inaccurate. The calculation of θ_1 grid values makes use of a function $\boldsymbol{\omega}_0(\theta_0, \theta_1)$ that is defined by using Eq. (8) at $k=0$ to compute $\boldsymbol{\omega}_0$ under the assumptions $\boldsymbol{\omega}_0=\boldsymbol{\omega}_1$ and $\alpha_{10}=\alpha_{20}=\alpha_{11}=\alpha_{21}=0$. An example of the dependence of $\|\boldsymbol{\omega}_0(\theta_0, \theta_1)\|$ on θ_1 is shown in Fig. 2. Note how the maximum rate magnitude corresponds closely to half the sampling frequency: $\pi/(t_1-t_0)=0.251$ rad/s because $t_1-t_0=12.5$ s. The θ_1 that minimizes $\|\boldsymbol{\omega}_0(\theta_0, \theta_1)\|$ can be computed by using Eq. (30) with inputs from Eq. (29) evaluated under the assumption that $\bar{q}_1=q_0(\alpha_1=0, \alpha_2=0, \theta_0)$. Call this minimizing value $\theta_{1\,\min}$. An additional function that contributes to the θ_1 grid calculation is a target θ_1 standard deviation about each grid point. It is computed as a function of θ_1 according to the following formula:

$$\sigma_{\theta 1}(\theta_1) = \min\left(\left\{\frac{\max\left[f\|\boldsymbol{\omega}_0(\theta_0, \theta_1)\|, \mathrm{d}\Omega_{\min}\right]}{\|\partial\boldsymbol{\omega}_0(\theta_0, \theta_1)/\partial\theta_1\|}\right\}, \sigma_{\theta 1\max}\right) \tag{36}$$

The idea behind Eq. (36) is to use $2\sigma_{\theta 1}(\theta_1)$ as the distance between neighboring θ_1 grid points and to pick this value so that the norm of the change in $\boldsymbol{\omega}_0(\theta_0, \theta_1)$ from one grid point to the next is nominally $2f\|\boldsymbol{\omega}_0(\theta_0, \theta_1)\|$, where $2f$ is a small fractional change. The max[ ]

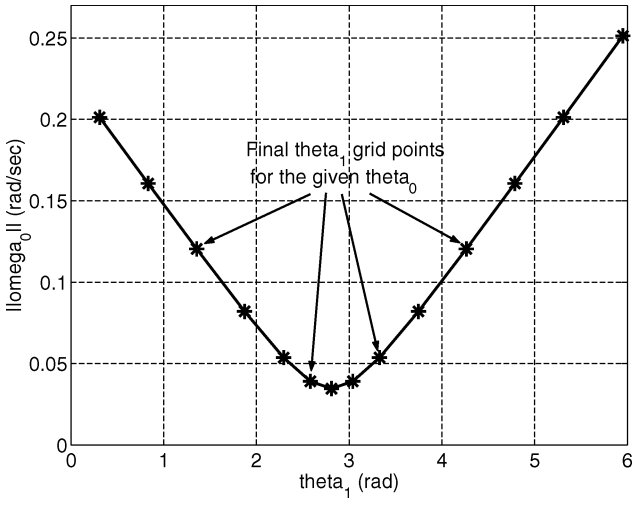

Fig. 2 $\|\boldsymbol{\omega}_0(\theta_0, \theta_1)\|$ vs θ_1 for a given θ_0 under the assumption of zero magnetometer measurement error.

function in Eq. (36) forces the norm of the change in $\boldsymbol{\omega}_0(\theta_0, \theta_1)$ between θ_1 grid points to be at least $2d\Omega_{\min}$. The min( ) function bounds the maximum distance between grid points to be below $2\sigma_{\theta 1\max}$ in cases where $\|\partial\boldsymbol{\omega}_0(\theta_0, \theta_1)/\partial\theta_1\|$ is small. Proper selection of the grid spacing parameters f, $d\Omega_{\min}$, and $\sigma_{\theta 1\max}$ is an art. Typical ranges of these values are $f = 0.20$ to 0.25, $d\Omega_{\min} = 0.001\pi/(t_1 - t_0)$ to $0.01\pi/(t_1 - t_0)$ rad/s, and $\sigma_{\theta 1\max} = 0.26$ to 0.35 rad (15 to 20 deg).

Using the function in Eq. (36), the θ_1 grid is defined implicitly by the relations

$$\theta_1^0 = \theta_{1\,\min} \tag{37a}$$

$$\theta_1^{m-1} = \theta_1^m - \sigma_{\theta 1}\left(\theta_1^m\right) - \sigma_{\theta 1}\left(\theta_1^{m-1}\right) \quad \text{for} \quad m = 0, -1, -2, \ldots, (m_{\min} + 1) \tag{37b}$$

$$\theta_1^{m+1} = \theta_1^m + \sigma_{\theta 1}\left(\theta_1^m\right) + \sigma_{\theta 1}\left(\theta_1^{m+1}\right) \quad \text{for} \quad m = 0, 1, 2, \ldots, (m_{\max} - 1) \tag{37c}$$

$$\theta_{1\,\min} - \pi - \sigma_{\theta 1}\left(\theta_1^{m_{\min}}\right) < \theta_1^{m_{\min}} \le \theta_{1\,\min} - \pi + \sigma_{\theta 1}\left(\theta_1^{m_{\min}}\right) \tag{37d}$$

$$\theta_{1\,\min} + \pi - \sigma_{\theta 1}\left(\theta_1^{m_{\max}}\right) \le \theta_1^{m_{\max}} < \theta_{1\,\min} + \pi + \sigma_{\theta 1}\left(\theta_1^{m_{\max}}\right) \tag{37e}$$

These relations can be used to generate the grid points by the following procedure: Eq. (37b) can be iterated backward from the known θ_1^0 for decreasing values of the superscript m until the Eq. (37d) termination criterion for $\theta_1^{m_{\min}}$ has been achieved. This termination effectively defines the value of the minimum grid index $m_{\min} < 0$. Similarly, Eq. (37c) can be iterated forward from θ_1^0 for increasing values of m until the Eq. (37e) termination criterion for $\theta_1^{m_{\max}}$ has been achieved, which defines the maximum grid index $m_{\max} > 0$. Equations (37b) and (37c) are implicit; θ_1^{m-1} appears on both sides of the former, and θ_1^{m+1} appears on both sides of the latter. Both of these implicit equations can be solved by a suitable iterative numerical procedure. An example procedure uses the preceding guess of θ_1^{m-1} in the right-hand side of Eq. (37b) to compute an improved guess. The initial guess of θ_1^{m-1} is set equal to θ_1^m. A similar iteration can be developed for Eq. (37c). One need not compute a very precise solution to either equation. These calculations are performed only at the initialization of the filter. Therefore, it is acceptable to spend a modest amount of computation time in order to develop each θ_1 grid.

Figure 2 shows a typical grid of θ_1 points that have been generated using Eqs. (37a–37e). The parameters that have been used to generate this grid are $f = 0.25$, $d\Omega_{\min} = 0.0021$ rad/s, $\sigma_{\theta 1\,\max} = 0.2618$ rad (15 deg), and $t_1 - t_0 = 12.5$ s.

The initialization algorithm typically uses between 144 and 360 grid points. An equal number of filters must be run for the first N samples in order to determine which grid point yields the most reasonable initialization. Typically there are $M_0 = 12$ grid points in the θ_0 direction. This yields a 30-deg spacing between θ_0 grid points, and the best initialization will be within 15 deg of the true initial θ_0. An angular deviation of this magnitude is well within the convergence region of most attitude estimation EKFs. Typical θ_1 grids employ anywhere from 12 to 30 points. This number varies for different θ_0 grid points because of the nature of the calculations defined in Eqs. (37a–37e). It might be that these numbers could be cut back by a factor of 2 to 4 by using coarser grid spacings, but this issue has not been fully investigated.

D. State and Information Matrix Initialization at Grid Points

At each grid point (θ_0, θ_1) the filter performs the following calculations in order to initialize its state estimate and its square-root information matrix. First, it uses the grid values of (θ_0, θ_1) and the values $\alpha_{10} = \alpha_{20} = \alpha_{11} = \alpha_{21} = 0$ and $\Delta\boldsymbol{h}_{d0} = 0$ in Eqs. (4) and (8) to set up and solve two linear equations for the two angular rate vectors $\boldsymbol{\omega}_0$ and $\boldsymbol{\omega}_1$. It uses the $\boldsymbol{\omega}_0$ solution and the θ_0 grid value to define $\hat{\boldsymbol{x}}_0 = [\boldsymbol{\omega}_0^T, 0, 0, \hat{\boldsymbol{x}}_{\text{aux}0}^T, \theta_0]^T$. This formula assumes that reasonable initial estimates are available for all inertia-parameter and rate-gyro-bias states in $\hat{\boldsymbol{x}}_{\text{aux}0}$.

The square-root information matrix gets initialized to reflect the initial magnetometer measurement errors and the uncertainties associated with the (θ_0, θ_1) grid spacings. The (θ_0, θ_1) uncertainties create weightings that penalize filter runs that, in effect, adjust their corresponding smoothed estimates of θ_0 and θ_1 to be far away from their grid points. This ensures that the filter with the lowest Eq. (35) test statistic will employ reasonable linear approximations because of the proximity of its actual estimates to the values about which Eqs. (4) and (8) get linearized. The θ_1 standard deviation has already been defined in Eq. (36). The θ_0 standard deviation equals half the θ_0 grid spacing:

$$\sigma_{\theta 0} = \pi/M_0 \tag{38}$$

Using these uncertainties, the initial square-root information matrix becomes

$$R_{\text{xx}0} = \begin{bmatrix} \dfrac{(\partial\boldsymbol{\omega}_0/\partial\theta_1)^T}{\sigma_{\theta 1}\|\partial\boldsymbol{\omega}_0/\partial\theta_1\|^2} & 0 & 0 & 0 & 0 \\ 0 & \dfrac{1}{\sigma_{\text{mag}}} & 0 & 0 & 0 \\ 0 & 0 & \dfrac{1}{\sigma_{\text{mag}}} & 0 & 0 \\ 0 & 0 & 0 & R_{\text{aux}0} & 0 \\ 0 & 0 & 0 & 0 & \dfrac{1}{\sigma_{\theta 0}} \end{bmatrix} \tag{39}$$

where the vector function $\boldsymbol{\omega}_0(\theta_0, \theta_1)$ is the one that has already been defined prior to Eq. (36). Its partial derivative gets evaluated at the (θ_0, θ_1) grid point. The upper-left-hand block of $R_{\text{xx}0}$ models the θ_1 uncertainty as though it were caused solely by uncertainty in ω_0 along the direction that most significantly affects θ_1. The total θ_1 uncertainty will be larger because of the added effect of the uncertainty in θ_0. This method of backpropagating θ_1 uncertainty into $R_{\text{xx}0}$ is not fully consistent with the idea of defining uncertainties in terms of θ grid spacings. It has been used because this indirect penalization of the initial estimate's distance from a particular θ_1 grid point seems to enhance performance. The matrix $R_{\text{aux}0}$ is the initial square-root information matrix for $\boldsymbol{x}_{\text{aux}}$ and is assumed to be given.

This $R_{\text{xx}0}$ definition has two fewer rows than columns because there is only one row corresponding to the attitude rate vector. This is nonstandard for an SRIF, but it poses no difficulties for the method. The availability of $\hat{\boldsymbol{x}}_0$ eliminates the need to invert $R_{\text{xx}0}$. $R_{\text{xx}1}$ becomes square and nonsingular after the first propagation and measurement update operations.

E. Reduction to a One-Dimensional Grid Using Rate-Gyro Data

Rate-gyro data, even if available for only one or two axes, usually reduce the space of large initial condition uncertainties to be one dimensional. In this case a grid is needed only in the θ_0 direction, and the global initialization strategy can be implemented with as few as 12 parallel filters. The only requirement is the following: The sensitive axis of at least one rate gyro must have a component that is parallel to $\hat{\boldsymbol{b}}_{\text{sc0}}$. At a given θ_0 grid point, the initialization starts by finding the θ_1 value that optimizes the cost function

$$J_{\text{init}}(\theta_1) = \tfrac{1}{2}\left\| \tfrac{1}{2}(\boldsymbol{y}_{\text{rg0}} + \boldsymbol{y}_{\text{rg1}}) - H_{\text{rg}}\boldsymbol{\omega}_0(\theta_0, \theta_1) \right\|^2 \tag{40}$$

$J_{\text{init}}(\theta_1)$ can be optimized over the interval 0 to 2π using a global numerical procedure. This procedure adds little computational expense to the estimator because these optimizations are performed only at the initial time and only for about 12 different θ_0 grid points. One uses $\boldsymbol{\omega}_0(\theta_0, \theta_{1\text{opt}})$ to initialize the rate estimate in the null space of H_{rg}, and one uses $\boldsymbol{y}_{\text{rg0}}$ to initialize the remaining components of $\boldsymbol{\omega}_0$. The first row of R_{xx0} in Eq. (39) gets modified to use $H_{\text{rg}}/\sigma_{\text{rg}}$ in its upper-left-hand block, and the rate-gyro bias block of that row becomes I/σ_{rg}.

This initialization strategy can be used with a magnetometer-based attitude determination EKF that eliminates Euler's equations by using three-axis rate-gyro data directly in the kinematics model, as in Refs. 4 and 5. Depending on the choice of filter implementation, one might need to modify the Eq. (35) hypothesis test statistic. In a standard Kalman-filter formulation the correct statistic is the sum of the squares of the normalized filter innovations.[15]

V. Results

This new magnetometer-based attitude determination algorithm has been tested using data from actual spacecraft and from a truth-model simulation. Flight data have been used from the Far-Ultraviolet Spectroscopic Explorer mission (FUSE) and from the HST mission. The FUSE results demonstrate the filter's performance using real flight data when there is a known accurate attitude from star sensors for use as a "truth" benchmark. The HST data show how the filter performs in a backup mode situation. This data come from a period when HST had only two working rate gyros, and so traditional gyro-driven kinematics models could not be used during this period. The truth model simulation is useful for testing the algorithm's ability to estimate inertia parameters because this is the only case in which truth values for these quantities are available. Note that all of the filter runs reported below use the two-iteration estimation procedure for $\boldsymbol{x}_{\text{sk}+1}$ that is described in Section III.D.

A. FUSE Mission Results

The FUSE spacecraft was launched in June 1999 to study the far-ultraviolet spectrum of the universe. It carries a three-axis magnetometer, a three-axis star tracker, and a three-axis rate-gyro system. The latter two sensors have been used to provide truth attitude and rate values. Their respective accuracies are on the order of arc seconds and tenths of a degree per hour.

Calibrated magnetometer data from 16 February 2002 have been filtered using the algorithm described in this paper. At that time the FUSE spacecraft was in an orbit with an apogee of 759 km, a perigee of 749 km, and an inclination of 25 deg. The magnetometer was sampled once every 2 s. The filter used a spacecraft inertia matrix, a magnetic torquer dipole moment time history, and a reaction wheel angular momentum time history that were based on actuator telemetry and on in-flight calibrations of moments of inertia, alignments, and scale factors. These calibrations had been performed independent of the new attitude estimation algorithm using data from the star tracker and the rate gyros. The calibrated moment of inertia matrix was $I_m = [3524, 3, 24; 3, 3303, -10; 24, -10, 642]\ \text{kg}\cdot\text{m}^2$.

The filter has been run for this case using the following properties. It has estimated only attitude and rate. The tuning parameters that have been used are $\bar{R}_{\text{hhk}} = I/\sqrt{[q_{\text{hct}}(t_{k+1} - t_k)]}$, where $q_{\text{hct}} = 8 \times 10^{-6}\ \text{N}^2\cdot\text{m}^2\cdot\text{s}$ and $\sigma_{\text{mag}} = 0.0062$ rad (0.35 deg). The chosen value of q_{hct} reflects the expected level of torque uncertainty in Eq. (4) for this mission. This represents a rough estimate, but the filter is insensitive to the value of q_{hct}; a value that is 10 times smaller increases the peak attitude errors by only 25%. The grid of possible initial conditions has been set up with $M_0 = 12$ grid points for the θ_0 grid, and the parameter values $f = 0.2$, $d\Omega_{\text{min}} = 0.0017$ rad/s, and $\sigma_{\theta 1\,\text{max}} = 0.3491$ rad have been used to set up the θ_1 grid. The total number of (θ_0, θ_1) initial condition grid points that have been used is 360, which corresponds to an average of 30 θ_1 grid points for each value of θ_0. The corresponding 360 filters have been run for the first 719 samples (0.4 h) after which only the best filter has been run. During this time, the Earth's magnetic field direction rotates with respect to inertial space by 52 deg, which makes the full three-axis attitude strongly observable.

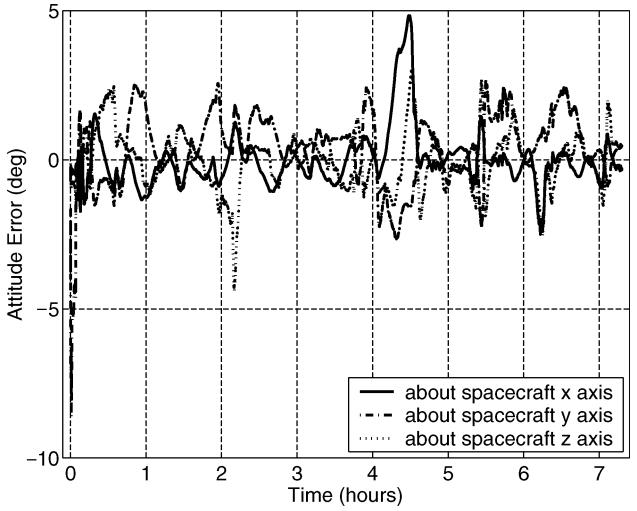

Fig. 3 Attitude error time histories for the basic filter operating on FUSE data.

The initialization procedure works well in this case. The hypothesis test statistic has a distinct global minimum near the true initial attitude. All of the initial conditions that produce costs within 20% of the optimum have θ_0 and θ_1 values within 34 deg of the truth. Any of these first guesses could have been used successfully, which implies that the initialization procedure is insensitive to small changes in the input data.

The attitude error time histories for this case are shown in Fig. 3. The maximum per-axis error for this 7.2 h run is 8.6 deg, and this maximum occurs during the filter's initial transient. In steady state the filter has a maximum per-axis error of 4.8 deg and an rms per-axis error of 1 deg. The maximum steady-state per-axis rate error is 0.015 deg/s, and the rms steady-state per-axis error is 0.0024 deg/s. The spacecraft was maneuvering during this data interval, and its maximum maneuver angular rate was 0.2057 deg/s. The rms per-axis rate error of the new filter is only 1.2% of the maximum actual rate, which is a reasonable resolution for a system that uses such a crude sensor.

A second filtering case has been run using this same data set. The differences for this case are that it uses incomplete rate-gyro information and that it estimates inertia matrix corrections and rate-gyro biases. The incomplete rate-gyro information uses only the x- and y-axis gyro outputs, and $\sigma_{\text{rg}} = 1.17 \times 10^{-6}$ rad/s (6.7×10^{-5} deg/s). The a priori inertia parameters have been chosen to yield the nominal I_m from the first FUSE case. The 1-σ a priori uncertainties of the first three inertia parameters equal 7% of their nominal values, and the principal axes uncertainties are 14 deg per axis. The two rate-gyro bias states has been initialized to $0 \pm 2.3 \times 10^{-5}$ rad/s (0.0013 deg/s) 1-σ. The rate-gyro data have enabled the statistical initialization to use a one-dimensional grid in θ_0 space. Twenty-four parallel initialization filters with a 15-deg grid spacing have been run for the first 0.4 h of data. Only the best filter has been run after that.

This second case demonstrates the filter's ability to take advantage of incomplete rate-gyro information. The attitude and rate errors decrease dramatically about the two axes for which rate-gyro data

are available. The peak steady-state attitude errors decrease from 4.8 to 0.5 deg about the spacecraft's x axis and from 2.7 to 0.6 deg about the spacecraft's y axis. Similarly, the peak steady-state attitude rate error about the x axis decreases from 0.006 to 0.0002 deg/s, and the peak rate error about the y axis decreases from 0.006 to 0.0003 deg/s. The peak steady-state z-axis attitude error decreases slightly for this filter run, from 3.0 to 2.9 deg, but the peak steady-state z-axis attitude rate error increases somewhat, from 0.015 to 0.024 deg/s. The steady-state rms errors show similar trends, except that the z-axis rms rate error also decreases slightly when using partial rate-gyro information. These results demonstrate that the use of incomplete rate-gyro information in the filter can provide significant accuracy improvements for the axes about which data are available and that rms errors can be decreased even on an axis for which rate-gyro data are not available.

B. HST Mission Results

Another case has used data from the Hubble Space Telescope that was recorded on 16 November 1999. The HST spacecraft was operating in a contingency mode because only two axes worth of rate-gyro data were available. The star trackers were not operating, and the only attitude data came from a magnetometer and a coarse sun sensor. Magnetometer data have been calibrated using an attitude independent calibration, and they have been filtered using this paper's techniques.

Two different filter cases have been run. One has used the incomplete rate-gyro data along with the magnetometer data, and it has estimated inertia parameters and rate-gyro biases along with attitude and attitude rate. The other filter has used only the magnetometer data and has estimated only the attitude and the attitude rate.

The input parameters for the filter that uses rate-gyro data are as follows: The tuning parameters are $\bar{R}_{\text{hhk}} = I/\sqrt{[q_{\text{hctk}}(t_{k+1} - t_k)]}$, where q_{hctk} ranges from a nominal value of 0.01 $\text{N}^2 \cdot \text{m}^2 \cdot \text{s}$ up to 0.02 $\text{N}^2 \cdot \text{m}^2 \cdot \text{s}$ during times of large reaction wheel or magnetic torque rod activity, $\sigma_{\text{mag}} = 0.035$ rad (2 deg), and $\sigma_{\text{rg}} = 1.7 \times 10^{-5}$ rad/s (0.0010 deg/s). The chosen q_{hctk} values reflect the expected level of torque uncertainty in Eq. (4) for the mission. Telemetry accuracy limitations cause this uncertainty to be larger when the reaction wheels or the magnetic torque rods are active. The a priori inertia parameters yield $I_m = \text{diag}(31000; 76000; 76000)$ $\text{kg} \cdot \text{m}^2$. The first three inertia parameters' a priori 1-σ uncertainties are set to 7% of their nominal values, and the principal axes uncertainties are set to 14 deg per axis. The initial rate-gyro bias estimates are set to zero, and their a priori 1-σ uncertainties are set to 9×10^{-6} rad/s (0.0005 deg/s). A one-dimensional θ_0 grid with a spacing of 15 deg has been used in the statistical initialization to generate 24 parallel filters. These filters have been run for 2.2 h of data using a sample period of 30 s. The most likely θ_0 grid point from these 24 runs has been used to initialize a new filter that has run for 8 h using a sample period of 3 s.

The filter that uses only magnetometer data and that estimates only attitude and rate has been run with slightly different inputs. The same values have been used for the tuning parameters q_{hctk} and σ_{mag}. The best inertia matrix estimate from the other HST filter run has been used in this filter: $I_m = [30517, -1105, -506; -1105, 77776, -276; -506, -276, 76743]$ $\text{kg} \cdot \text{m}^2$. The statistical initialization has used a θ_0 grid spacing of 30 deg, and the parameter values $f = 0.2$, $d\Omega_{\min} = 0.0017$ rad/s, and $\sigma_{\theta 1 \max} = 0.3491$ rad have been used to generate each θ_1 grid. These parameters yield a total of 264 initialization grid points. The measurement sample period for this filter is 12 s. All 264 filters have been run in parallel for 0.7 h worth of data, and then the best filter has been selected and run for the remainder of a 4.6-h data interval.

It is difficult to evaluate the performance of these two HST filter runs because the truth attitude and rate are not known. One means of checking their accuracy is the coarse sun sensor. During orbit day, a filter's attitude estimate and the known inertial sun direction can be used to get an estimate of the sun direction vector in spacecraft coordinates. The angle between this vector and the coarse sun-sensor's output can be calculated as a metric of filter accuracy. Figure 4 shows this angle for the filter that estimates attitude, attitude rate, inertia matrix parameters, and rate-gyro biases using the magnetometer outputs and the two axes worth of rate-gyro data. Figure 5 plots the sun error angle for the filter that uses only magnetometer data and that estimates only attitude and rate. These figures plot the sun error vs the angular distance between the sun and the center of the sun sensor's field of view (FOV). The coarse sun sensor's accuracy is known to be about 1.5 deg near the center of its FOV, but the calibration errors can increase to about 10 deg when the sun is 30 deg from the center of the FOV.[16] Thus, the coarse sun sensor can be used as a check on the filter only when the sun is near the center of its FOV, which corresponds to the left-most data on Figs. 4 and 5. Note how these two figures have their lowest average sun errors when the sun is near the center of its field of view. The average curve plotted in these two figures represents an approximate calibration of the coarse sun sensor.

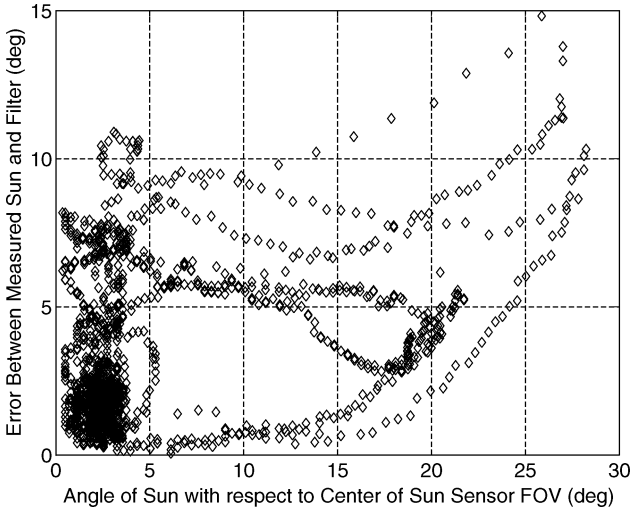

Fig. 4 Comparison between HST coarse sun sensor outputs and the attitude estimate of the filter that uses magnetometer and rate-gyro data.

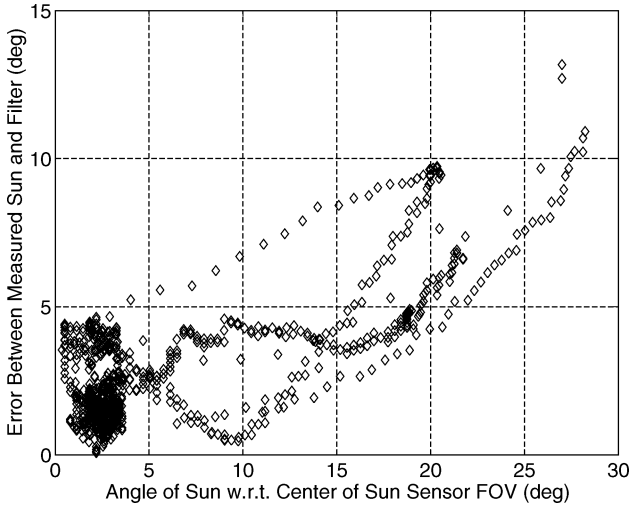

Fig. 5 Comparison between HST coarse sun sensor outputs and the attitude estimate of the filter that uses only magnetometer data.

The relative accuracy of the two filters seems counterintuitive. Figure 4 indicates that the filter which uses rate-gyro information has an accuracy of only 10 deg because the left-most points on the figure range this high. The left-most points on Fig. 5 imply that the filter which uses only magnetometer data has a superior accuracy, that is, better than 5 deg. One normally expects the addition of data, such as the rate-gyro data, to improve the accuracy, not to degrade it. The reason for this discrepancy is not completely understood. It might stem from suboptimal tuning of the filter that uses rate gyros, or it might be the result of an uncalibrated misalignment between the

rate gyros and the coarse sun sensor. Another possible explanation is that the filter run, which uses rate gyros, also estimates additional quantities, rate-gyro biases, and inertia matrix parameters. An increase in the number of estimated quantities tends to degrade the observability of a system.

The gyro-less filter has also been checked by comparing its attitude rate estimates to the available rate-gyro data. The maximum discrepancy is 0.0276 deg/s per axis, and the rms difference is 0.0057 deg/s per axis. The maximum discrepancy occurs during an attitude gyration that produced rate-gyro outputs as high as 0.2367 deg/s. Thus, the gyro-less filter's attitude rate estimates are accurate to within 12% during large maneuvers.

The nonlinear smoother has also been run for these two cases, and the results are as expected. Each smoother produces better agreement with the coarse sun-sensor data then its corresponding filter. Each smoother is about twice as accurate as the corresponding filter.

Another metric of whether these solutions are reasonable is the unmodeled disturbance torque magnitude $\|\Delta \boldsymbol{h}_{\mathrm{dk}}/(t_{k+1}-t_k)\|$ as estimated by the smoother. The peak value of this norm for the gyro-less smoother case is 0.023 N · m, and the mean value of this magnitude is 0.008 N · m. The peak and mean magnitudes of the magnetic control torque are 0.249 and 0.106 N · m, respectively, and the gravity-gradient torque's respective peak and mean magnitudes are 0.088 and 0.064 N · m. The relative smallness of the unmodeled torques is convincing evidence of the accuracy of this system's attitude and rate estimates and of the acceptability of using the Euler attitude dynamics model in the estimation algorithm.

C. Results Using a Truth-Model Simulation

A truth model simulation has been used to test two aspects of the filter that could not be tested using the available flight data. The simulation has been used to test the ability of the filter to estimate the moment-of-inertia parameters of Eq. (10) and its ability to converge when the spacecraft is tumbling. The simulation includes realistic models of the disturbance torques caused by atmospheric drag, solar-radiation pressure, and the Earth's gravity gradient. It also includes magnetometer errors such as scale factor, misalignment, bias, and thermal noise.

Consider the results for an example that involves a microsatellite in a circular orbit with an altitude of 823 km and an inclination of 82 deg. The initial rotation rate is 2 deg/s, and the truth moment-of-inertia matrix is $I_m = [1.6135, -0.0151, -0.0703; -0.0151, 1.5497, -0.0435; -0.0703, -0.0435, 1.5571]$ kg · m^2. The filter's initial estimate of the moment-of-inertia parameters yields the matrix $I_m = \mathrm{diag}(1.5715, 1.5715, 1.5715)$ kg · m^2. The spacecraft has no momentum wheels, and no magnetic control torque is applied, which makes one of the six inertia parameters unobservable. The filter can estimate the two independent ratios of the principal inertias and the principal axis directions, but it must rely on its a priori estimate of the absolute scale of the inertia matrix. This is a difficult problem because the natural attitude dynamics are almost neutrally stable, and unmodeled disturbance torques tend to dominate the dynamics.

The other parameters of the filter are as follows: The sampling period is 10 s. $\bar{R}_{\mathrm{hhk}} = I/\sqrt{[q_{\mathrm{hct}}(t_{k+1}-t_k)]}$, where $q_{\mathrm{hct}} = 9 \times 10^{-15}$ N^2 · m^2 · s, and $\sigma_{\mathrm{mag}} = 0.0087$ rad (0.5 deg). The first two inertia parameters' a priori 1-σ uncertainties are 2.5% of their a priori values, the third parameter's 1-σ uncertainty is 0.025% of its a priori value, and the principal axis direction a priori uncertainties are 15 deg per axis. A θ_0 grid spacing of 30 deg has been used. The θ_1 grid spacing parameters are $f = 0.2$, $d\Omega_{\min} = 0.0017$ rad/s, and $\sigma_{\theta 1\,\max} = 0.2618$ rad. These values result in a total of 230 (θ_0, θ_1) grid points. The initialization procedure runs all 230 filters for the first 61 samples (0.17 h) and then runs the best filter for an additional 440 samples (1.2 h).

The results for this case are reasonable, though not great. The peak attitude error per axis after transients have died out is 18 deg. The peak per-axis attitude rate error is 0.5 deg/s. This performance is not impressive; the peak spin-rate errors are as much as 25% of the actual spin rate. This level of accuracy is reasonable given the challenges of this estimation problem.

The inertia matrix estimate slowly improves in comparison to its initial value. Suppose that one uses the induced matrix 2-norm to define the following relative inertia matrix error metric: $\|\hat{I}_m tr(I_m)/tr(\hat{I}_m) - I_m\|/\|I_m\|$, where $\hat{I}_m$ is the estimated inertia matrix, I_m is the true inertia matrix, and $tr(\,)$ is the trace operator. This metric is sensible because the trace ratio in the first term eliminates the effect of the unobservable (and ignorable) absolute scaling of the inertia matrix. This metric starts out at 5.6%, but after 1.4 h of filtering the relative error gets reduced to 2.9%. Thus, inertia parameter estimation is possible with this filter.

The nonlinear smoother is much more efficacious than the filter for this problem. Its attitude estimates are accurate to within 0.6 deg for the entire 1.4-h data interval, and the relative error metric for its inertia matrix estimate is 0.1%. These results suggest that the filter's performance would improve after a longer interval, when inertia matrix estimation transients had died out. They also imply that a filter which used fixed-lag smoothing[15] with nonlinear iteration, a sort of superiterated EKF, would produce performance improvements for this case.

D. Algorithm Execution Speed

The algorithms have been run on a 3.06-GHz Xeon Processor using Matlab® in its normal interpretive mode. They require about 0.16 s per magnetometer sample when doing the full nonlinear outer-loop optimization with two iterations of the inner optimization. Use of the rapid outer-loop optimization of Section III.E will decrease the execution time by a factor of seven or more, and use of compiled code will further decrease the execution time, perhaps by a factor of 10 or more. These speeds indicate that the basic algorithm could run on a less powerful flight processor in real time if the magnetometer sampling rate were not too high, say 1 Hz or slower. The global initialization procedure of Section IV, however, probably will not be executable in real time. Fortunately, this procedure will not need to be executed often. It might be acceptable to execute it in a flight computer in a non-real-time mode and, after it completes, temporarily run the best filter in a faster-than-real-time mode in order to catch up with the magnetometer data.

VI. Conclusions

A new attitude determination filter has been developed that can estimate three-axis attitude and rate based solely on magnetometer measurements. It uses a new attitude parameterization that is based on perturbations to the minimum quaternion that aligns the magnetic field in reference coordinates with the field in spacecraft coordinates. The filter model includes Euler dynamics and attitude kinematics, and its least-squares cost function penalizes magnetometer measurement errors and torque process noise. The new attitude representation is used to develop a generalized iterated extended Kalman filter that estimates the unknown rotation about the measured magnetic field by performing a global nonlinear minimization of the squared-error cost function. The nominal filter relies only on magnetometer data, but other types of attitude and rate data can be incorporated if available. The filter can be augmented to estimate parameters that define the spacecraft's moment-of-inertia matrix.

The new algorithm includes an initialization procedure, which automatically finds a reasonable initial estimate that causes the extended Kalman filter to converge. This procedure forms a grid in a two-dimensional box within the space of rotations about the first two measured magnetic field vectors. Multiple filters are run, one for each initialization grid point, and a statistical hypothesis test is used to pick the filter that has the best initialization.

The new attitude estimation algorithm has been tested on flight data and using a truth-model simulation. The method has yielded per-axis accuracies of better than 5 deg in attitude and better than 0.03 deg/s in rate when operating on flight data from the Far-Ultraviolet Spectroscopic Explorer mission and from the Hubble Space Telescope mission. Truth model simulation results have validated the method's ability to estimate moment-of-inertia parameters, and they have demonstrated the efficacy of its initialization procedure for a spacecraft that is tumbling at 2 deg/s.

Acknowledgments

This work was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) Grant Number F49620-01-1-0117 and by NASA Grant Number NAG5-11919. Belinda King was the AFOSR Grant Monitor, and Richard Harman was the NASA Grant Monitor. Richard Harman also supplied the Hubble Space Telescope and Far-Ultraviolet Spectroscopic Explorer Mission data that has been used.

References

[1]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.

[2]Hashmall, J. A., Liu, K., and Rokni, M., "Accurate Spacecraft Attitudes from Magnetometer Data," *CNES Spaceflight Dynamics (Colloque 95)*, Cépaduès-Éditions, Toulouse, France, 1995, pp. 169–179.

[3]Hashmall, J., and Sedlak, J., "The Use of Magnetometers for Accurate Attitude Determination," *Proceedings of the 12th International Symposium on Spaceflight Dynamics*, ESTEC, Noordwijk, The Netherlands, 1997, pp. 179–184.

[4]Sedlak, J., and Hashmall, J., "Accurate Magnetometer/Gyroscope Attitudes Using a Filter with Correlated Sensor Noise," *Proceedings of the NASA/GSFC Flight Mechanics Symposium*, NASA, Goddard Space Flight Center, Greenbelt, MO, 1997, pp. 293–303.

[5]Sedlak, J. E., "Improved Spacecraft Attitude Filter Using a Sequentially Correlated Magnetometer Noise Model," *Proceedings of the AIAA/IEEE 16th Digital Avionics Systems Conference*, AIAA, Reston, VA, 1997, pp. 8.4-9–8.4-16.

[6]Challa, M., Kotaru, S., and Natanson, G., "Magnetometer-Only Attitude and Rate Estimates During the Earth Radiation Budget Satellite 1987 Control Anomaly," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Vol. 2, AIAA, Reston, VA, 1997, pp. 830–840.

[7]Challa, M., Natanson, G., and Ottenstein, N., "Magnetometer-Only Attitude and Rates for Spinning Spacecraft," *Proceedings of the AIAA/AAS Astrodynamics Specialists Conference*, AIAA, Reston, VA, 2000, pp. 311–321.

[8]Reynolds, R. G., "Quaternion Parameterization and a Simple Algorithm for Global Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 4, 1998, pp. 669–671.

[9]Psiaki, M. L., and Oshman, Y., "Spacecraft Attitude Rate Estimation from Geomagnetic Field Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, 2003, pp. 244–252.

[10]Psiaki, M. L., "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 2, 2000, pp. 206–214.

[11]Kasdin, N. J., and Weaver, T. J. M., "Recursive Satellite Attitude Estimation with the Two-Step Optimal Estimator," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, AIAA, Reston, VA, 2002.

[12]Wertz, J. R. (ed.), *Spacecraft Attitude Determination and Control*, D. Reidel Pub. Co., Boston, 1978, pp. 414–416.

[13]Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977, pp. 69–76, 115–122, 214–217.

[14]Gill, P. E., Murray, W., and Wright, M. H., *Practical Optimization*, Academic Press, New York, 1981, pp. 88–92, 133–140.

[15]Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation*, Wiley, New York, 2001, pp. 213, 214, 236, 237, 333, 441, 442.

[16]Markley, F. L., and Nelson, J. D., "Zero-Gyro Safemode Controller for the Hubble Space Telescope," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 4, 1994, pp. 815–822.