UNSCENTED FILTERING FOR SPACECRAFT ATTITUDE ESTIMATION

John L. Crassidis*
Department of Mechanical & Aerospace Engineering
University at Buffalo, The State University of New York
Amherst, NY 14260-4400

F. Landis Markley†
Guidance, Navigation & Control Systems Engineering Branch
NASA-Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

A new spacecraft attitude estimation approach based on the Unscented Filter is derived. For nonlinear systems the Unscented Filter uses a carefully selected set of sample points to more accurately map the probability distribution than the linearization of the standard Extended Kalman Filter, leading to faster convergence from inaccurate initial conditions in attitude estimation problems. The filter formulation is based on standard attitude-vector measurements using a gyro-based model for attitude propagation. The global attitude parameterization is given by a quaternion, while a generalized three-dimensional attitude representation is used to define the local attitude error. A multiplicative quaternion-error approach is derived from the local attitude error, which guarantees that quaternion normalization is maintained in the filter. Simulation results indicate that the Unscented Filter is more robust than the Extended Kalman Filter under realistic initial attitude-error conditions.

INTRODUCTION

The Extended Kalman Filter (EKF) is widely used in attitude estimation. Several parameterizations can be used to represent the attitude, such as Euler angles,[1] quaternions,[2] modified Rodrigues parameters,[3] and even the rotation vector.[4] Quaternions are especially appealing since no singularities are present and the kinematics equation is bilinear. However, the quaternion must obey a normalization constraint, which can be violated by the linear measurement-updates associated with the standard EKF approach. The most common approach to overcome this shortfall involves using a multiplicative error quaternion, where after neglecting higher-order terms the four-component quaternion can effectively be replaced by a three-component error vector.[2] Under ideal circumstances, such as small attitude errors, this approach works extremely well.

One interesting fact of the formulation presented in Ref. [2] is that the 4×4 quaternion covariance is assumed to have rank three, i.e., the 4×4 matrix can be projected onto a 3×3 matrix without any loss in information. But, this is only strictly valid when the constraint is linear, which is not the case for the quaternion.[5] However, the covariance is nearly singular, and a linear computation such as the EKF can make it exactly singular. This approach is justifiable for small estimation errors, but may cause difficulties outside the valid linear region, e.g., during the initialization stage of the EKF.

Several approaches have addressed the issue of initialization for attitude estimation. Reference [6] explicitly includes the quaternion constraint in the measurement update. This works well for large initial condition errors; however, linearizations about the previous error estimate are required for the covariance propagation and for the measurement updates, which may produce biased estimates. Reference [7] breaks the measurement update into two steps. The linear first step uses a transformation of the desired states, while the second step uses a non-recursive minimization to recover the desired states. This also works well; however, local iterations at each time-step on a constrained minimization problem are required in the second step.

In this paper a new attitude estimation approach, based on a filter developed by Julier, Uhlmann and Durrant-Whyte,[8] is shown as an alternative to the EKF. This filter approach, which they call the *Unscented Filter*[9] (UF), has several advantages over the EKF, including: 1) the expected error is lower than the EKF, 2) the new filter can be applied to non-differentiable functions, 3) the new filter avoids the derivation of Jacobian matrices, and 4) the new filter

*Associate Professor, Associate Fellow AIAA. Email: johnc@eng.buffalo.edu
†Aerospace Engineer, Fellow AIAA. Email: lmarkley@pop500.gsfc.nasa.gov

is valid to higher-order expansions than the standard EKF. The UF works on the premise that with a fixed number of parameters it should be easier to approximate a Gaussian distribution than to approximate an arbitrary nonlinear function. Also, the UF uses the standard Kalman form in the post-update, but uses a different propagation of the covariance and pre-measurement update with no local iterations.

As with the standard EKF, using the UF directly with a quaternion parameterization of the attitude yields a nonunit quaternion estimate. A linearized model, such as the one used in Ref. [2], does not take full advantage of the UF capabilities. An alternative approach presented in this paper uses a three-component attitude-error vector to represent the quaternion error vector. Several three-component representations are possible, including the Gibbs vector (also known as the vector of Rodrigues parameters), which has a singularity at 180° rotations, and the vector of modified Rodrigues parameters (MRPs), which has a singularity at 360°.[10] Reference [11] proves that four times the vector of MRPs and twice the Gibbs vector leads to the same second-order approximation. We choose a generalized representation where the singularity can be placed anywhere from 180° to 360°.[12] Since we only use a three-component representation for the attitude errors, the singularity is never encountered in practice. The updates are performed using quaternion multiplication, leading to a natural way of maintaining the normalization constraint.

The organization of this paper proceeds as follows. First, the UF is reviewed. Then, a brief review of the quaternion attitude kinematics and gyro-model is provided, followed by a review of the generalized three-component error-vector representation. Then, an UF is derived for attitude estimation using a multiplicative quaternion update with the three-component error-vector representation. Finally, the new filter is compared with the EKF using simulated three-axis magnetometer and gyro measurements of the Tropical Rainfall Measurement Mission (TRMM).

UNSCENTED FILTERING

In this section the UF is reviewed. The filter presented in Ref. [13] is derived for discrete-time nonlinear equations, where the system model is given by

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, k) + G_k \mathbf{w}_k \tag{1a}$$

$$\tilde{\mathbf{y}}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k \tag{1b}$$

where $\mathbf{x}_k$ is the $n \times 1$ state vector and $\tilde{\mathbf{y}}_k$ is the $m \times 1$ measurement vector. Note that a continuous-time model can always be expressed in the form of Eq. (1a) through an appropriate numerical integration scheme. We assume that the process noise $\mathbf{w}_k$ and measurement-error noise $\mathbf{v}_k$ are zero-mean Gaussian noise processes with covariances given by Q_k and R_k, respectively. The standard Kalman Filter update equations are first rewritten as[14]

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + K_k \boldsymbol{\upsilon}_k \tag{2a}$$

$$P_k^+ = P_k^- - K_k P_k^{\upsilon\upsilon} K_k^T \tag{2b}$$

where $\hat{\mathbf{x}}_k^-$ and P_k^- are the pre-update state estimate and covariance, respectively, and $\hat{\mathbf{x}}_k^+$ and P_k^+ are the post-update state estimate and covariance, respectively. The *innovation* $\boldsymbol{\upsilon}_k$ is given by

$$\boldsymbol{\upsilon}_k \equiv \tilde{\mathbf{y}}_k - \hat{\mathbf{y}}_k^- = \tilde{\mathbf{y}}_k - \mathbf{h}(\hat{\mathbf{x}}_k^-, k) \tag{3}$$

The covariance of $\boldsymbol{\upsilon}_k$ is denoted by $P_k^{\upsilon\upsilon}$. The gain K_k is computed by

$$K_k = P_k^{xy} (P_k^{\upsilon\upsilon})^{-1} \tag{4}$$

where P_k^{xy} is the cross-correlation matrix between $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{y}}_k^-$.

The UF uses a different propagation than the standard EKF. Given an $n \times n$ covariance matrix P, a set of $2n$ *sigma points* can be generated from the columns of the matrices $\pm\sqrt{(n+\lambda)P}$, where $\sqrt{M}$ is shorthand notation for a matrix Z such that $Z\,Z^T = M$. The set of points is zero-mean, but if the distribution has mean $\boldsymbol{\mu}$, then simply adding $\boldsymbol{\mu}$ to each of the points yields a symmetric set of $2n$ points having the desired mean and covariance.[13] Due to the symmetric nature of this set, its odd central moments are zero, so its first three moments are the same as the original Gaussian distribution. The scalar λ is a convenient parameter for exploiting knowledge (if available) about the higher moments of the given distribution.[14] In scalar systems (i.e., for $n = 1$), a value of $\lambda = 2$ leads to errors in the mean and variance that are sixth order. For higher-dimensional systems choosing $\lambda = 3 - n$ minimizes the mean-squared-error up to the fourth order.[13] However, caution should be exercised when λ is negative since a possibility exists that the predicted covariance can become non-positive semi-definite. If this is a major concern, then another approach can be used that allows for scaling of the sigma points, which guarantees a positive semi-definite covariance matrix.[15] Also, it can be shown that when $n + \lambda$ tends to zero the mean tends to that calculated by the truncated second-order filter.[16] This is the foundation for the UF.

A method for incorporating process noise in the UF is shown in Ref. [17]. This approach generates a set of points in $[\mathbf{x}_k,\ \mathbf{w}_k]$ space that has the correct mean and covariance, and propagates these points through the model in Eq. (1a). The predicted mean and covariance are also augmented to included the process noise, but the basic structure of the their calculations remain unchanged (see Ref. [17] for more details). Although this approach may more fully utilize the capability of the unscented transformation, it will be more computationally costly due to the extra required calculations

arising from the augmented system. For the attitude estimation problem, a set of six more sigma points is required to implement this approach. This significantly increases the computational burden, which may prohibit its use for actual onboard implementations. Another approach based on a trapezoidal approximation will be shown in this paper.

The general formulation for the propagation equations is given as follows. First, compute the following set of sigma points:

$$\boldsymbol{\sigma}_k \leftarrow 2n \text{ columns from } \pm\sqrt{(n+\lambda)[P_k^+ + \bar{Q}_k]} \quad (5a)$$

$$\boldsymbol{\chi}_k(0) = \hat{\mathbf{x}}_k^+ \quad (5b)$$

$$\boldsymbol{\chi}_k(i) = \boldsymbol{\sigma}_k(i) + \hat{\mathbf{x}}_k^+ \quad (5c)$$

where the matrix $\bar{Q}_k$ is related to the process noise covariance, which will be discussed shortly. One efficient method to compute the matrix square root is the Cholesky decomposition.[18] Alternatively, the sigma points can be chosen to lie along the eigenvectors of the covariance matrix. Note that there are a total of $2n$ values for $\boldsymbol{\sigma}_k$ (the positive and negative square roots). The transformed set of sigma points are evaluated for each of the points by

$$\boldsymbol{\chi}_{k+1}(i) = \mathbf{f}[\boldsymbol{\chi}_k(i), k] \quad (6)$$

The predicted mean is given by

$$\hat{\mathbf{x}}_{k+1}^- = \frac{1}{n+\lambda}\left\{\lambda\,\boldsymbol{\chi}_{k+1}(0) + \frac{1}{2}\sum_{i=1}^{2n}\boldsymbol{\chi}_{k+1}(i)\right\} \quad (7)$$

The predicted covariance is given by

$$P_{k+1}^- = \frac{1}{n+\lambda}\left\{\lambda\,[\boldsymbol{\chi}_{k+1}(0) - \hat{\mathbf{x}}_{k+1}^-]\,[\boldsymbol{\chi}_{k+1}(0) - \hat{\mathbf{x}}_{k+1}^-]^T + \frac{1}{2}\sum_{i=1}^{2n}[\boldsymbol{\chi}_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^-]\,[\boldsymbol{\chi}_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^-]^T\right\} + \bar{Q}_k \quad (8)$$

The mean observation is given by

$$\hat{\mathbf{y}}_{k+1}^- = \frac{1}{n+\lambda}\left\{\lambda\,\boldsymbol{\gamma}_{k+1}(0) + \frac{1}{2}\sum_{i=1}^{2n}\boldsymbol{\gamma}_{k+1}(i)\right\} \quad (9)$$

where

$$\boldsymbol{\gamma}_{k+1}(i) = \mathbf{h}[\boldsymbol{\chi}_{k+1}(i), k] \quad (10)$$

The output covariance is given by

$$P_{k+1}^{yy} = \frac{1}{n+\lambda}\left\{\lambda\,[\boldsymbol{\gamma}_{k+1}(0) - \hat{\mathbf{y}}_{k+1}^-]\,[\boldsymbol{\gamma}_{k+1}(0) - \hat{\mathbf{y}}_{k+1}^-]^T + \frac{1}{2}\sum_{i=1}^{2n}[\boldsymbol{\gamma}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^-]\,[\boldsymbol{\gamma}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^-]^T\right\} \quad (11)$$

Then, the innovation covariance is simply given by

$$P_{k+1}^{\upsilon\upsilon} = P_{k+1}^{yy} + R_{k+1} \quad (12)$$

Finally, the cross-correlation matrix is determined using

$$P_{k+1}^{xy} = \frac{1}{n+\lambda}\left\{\lambda\,[\boldsymbol{\chi}_{k+1}(0) - \hat{\mathbf{x}}_{k+1}^-]\,[\boldsymbol{\gamma}_{k+1}(0) - \hat{\mathbf{y}}_{k+1}^-]^T + \frac{1}{2}\sum_{i=1}^{2n}[\boldsymbol{\chi}_{k+1}(i) - \hat{\mathbf{x}}_{k+1}^-]\,[\boldsymbol{\gamma}_{k+1}(i) - \hat{\mathbf{y}}_{k+1}^-]^T\right\} \quad (13)$$

The filter gain is then computed using Eq. (4), and the state vector can now be updated using Eq. (2). Even though $2n+1$ propagations are required for the UF, the computations may be comparable to the EKF, especially if the continuous-time covariance equations need to be integrated and a numerical Jacobian matrix is evaluated. Since the propagations can be performed in parallel, the UF is ideally suited for parallel computation architectures. A square root implementation of the UF has been developed.[19] Although this promises to have improved numerical properties, it will not be considered in this paper.

Reference [17] states that if the process noise is purely additive in the model, then its covariance can simply be added using a simple additive procedure. In this paper we expand upon this concept by incorporating an approximation for the integration over the sampling interval, which more closely follows the actual process. Any process noise that is added to the state vector in the UF is effectively multiplied by the state transition matrix, $\Phi(\Delta t)$, which gives $\Phi(\Delta t)\,Q_k\Phi^T(\Delta t)$ at the end of the interval. This mapping is done automatically by the state propagation, and does not need to be explicitly accounted for in the propagation. However, adding equal process noise at the beginning and end of the propagation might yield better results. The total desired process noise follows

$$\Phi(\Delta t)\,\bar{Q}_k\Phi^T(\Delta t) + \bar{Q}_k = G_k Q_k G_k^T \quad (14)$$

where $\bar{Q}_k$ is used in Eq. (5a) and in the calculation of the predicted covariance in Eq. (8). This approach is similar to a trapezoid rule for integration. An explicit solution for $\bar{Q}_k$ in the attitude estimation problem depends on the attitude kinematics, which we next show.

ATTITUDE KINEMATICS

In this section a brief review of the attitude kinematics equation of motion using quaternions is shown. Also, a generalization of the Rodrigues parameters is briefly discussed. Finally, gyro and attitude-vector sensor models are shown. The quaternion is defined by $\mathbf{q} \equiv \begin{bmatrix}\boldsymbol{\varrho}^T & q_4\end{bmatrix}^T$, with $\boldsymbol{\varrho} \equiv [q_1 \;\; q_2 \;\; q_3]^T = \hat{\mathbf{e}}\sin(\vartheta/2)$ and $q_4 = \cos(\vartheta/2)$, where $\hat{\mathbf{e}}$ is the axis of rotation and

ϑ is the angle of rotation.[10] Since a four-dimensional vector is used to describe three dimensions, the quaternion components cannot be independent of each other. The quaternion satisfies a single constraint given by $\mathbf{q}^T\mathbf{q} = 1$. The attitude matrix is related to the quaternion by

$$A(\mathbf{q}) = \Xi^T(\mathbf{q})\Psi(\mathbf{q}) \tag{15}$$

with

$$\Xi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_{3\times 3} + [\boldsymbol{\varrho}\times] \\ -\boldsymbol{\varrho}^T \end{bmatrix} \tag{16a}$$

$$\Psi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_{3\times 3} - [\boldsymbol{\varrho}\times] \\ -\boldsymbol{\varrho}^T \end{bmatrix} \tag{16b}$$

where $I_{3\times 3}$ is a 3×3 identity matrix and $[\boldsymbol{\varrho}\times]$ is a cross product matrix since $\mathbf{a} \times \mathbf{b} = [\mathbf{a}\times]\mathbf{b}$, with

$$[\mathbf{a}\times] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{17}$$

Successive rotations can be accomplished using quaternion multiplication. Here we adopt the convention of Refs. [2] and [10] who multiply the quaternions in the same order as the attitude matrix multiplication: $A(\mathbf{q}')A(\mathbf{q}) = A(\mathbf{q}' \otimes \mathbf{q})$. The composition of the quaternions is bilinear, with

$$\mathbf{q}' \otimes \mathbf{q} = \begin{bmatrix} \Psi(\mathbf{q}') & \vdots & \mathbf{q}' \end{bmatrix} \mathbf{q} = \begin{bmatrix} \Xi(\mathbf{q}) & \vdots & \mathbf{q} \end{bmatrix} \mathbf{q}' \tag{18}$$

Also, the inverse quaternion is given by $\mathbf{q}^{-1} = \begin{bmatrix} -\boldsymbol{\varrho}^T & q_4 \end{bmatrix}^T$. The quaternion kinematics equation is given by

$$\dot{\mathbf{q}}(t) = \frac{1}{2}\Xi[\mathbf{q}(t)]\boldsymbol{\omega}(t) \tag{19}$$

where $\boldsymbol{\omega}$ is the 3×1 angular velocity vector.

The local error-quaternion, denoted by $\boldsymbol{\delta\mathbf{q}} \equiv \begin{bmatrix} \boldsymbol{\delta\varrho}^T & \delta q_4 \end{bmatrix}^T$, which will be defined in the UF formulation, is represented using a vector of generalized Rodrigues parameters:[12]

$$\boldsymbol{\delta\mathbf{p}} \equiv f\frac{\boldsymbol{\delta\varrho}}{a + \delta q_4} \tag{20}$$

where a is a parameter from 0 to 1, and f is a scale factor. Note when $a = 0$ and $f = 1$ then Eq. (20) gives the Gibbs vector, and when $a = f = 1$ then Eq. (20) gives the standard vector of MRPs. For small errors the attitude part of the covariance is closely related to the attitude estimation errors for any rotation sequence, given by a simple factor.[2] For example, the Gibbs vector linearize to half angles, and the vector of MRPs linearize to quarter angles. We will choose $f = 2(a+1)$ so that $||\boldsymbol{\delta\mathbf{p}}||$ is equal to ϑ for small errors. The inverse transformation from $\boldsymbol{\delta\mathbf{p}}$ to $\boldsymbol{\delta\mathbf{q}}$ is given by

$$\delta q_4 = \frac{-a\,||\boldsymbol{\delta\mathbf{p}}||^2 + f\sqrt{f^2 + (1-a^2)||\boldsymbol{\delta\mathbf{p}}||^2}}{f^2 + ||\boldsymbol{\delta\mathbf{p}}||^2} \tag{21a}$$

$$\boldsymbol{\delta\varrho} = f^{-1}(a + \delta q_4)\,\boldsymbol{\delta\mathbf{p}} \tag{21b}$$

Discrete-time attitude observations for a single sensor are given by

$$\tilde{\mathbf{b}}_i = A(\mathbf{q})\mathbf{r}_i + \boldsymbol{\nu}_i \tag{22}$$

where $\tilde{\mathbf{b}}_i$ denotes the i^{th} 3×1 measurement vector, $\mathbf{r}_i$ is the i^{th} known 3×1 reference vector, and the sensor error-vector $\boldsymbol{\nu}_i$ is Gaussian which satisfies

$$E\{\boldsymbol{\nu}_i\} = \mathbf{0} \tag{23a}$$

$$E\left\{\boldsymbol{\nu}_i\boldsymbol{\nu}_i^T\right\} = \sigma_i^2 I \tag{23b}$$

where $E\{\,\}$ denotes expectation. Note that if unit measurement vectors are used then Eq. (23b) should be appropriately modified.[20] Multiple (N) vector measurements can be concatenated to form

$$\tilde{\mathbf{y}}_k = \begin{bmatrix} A(\mathbf{q})\mathbf{r}_1 \\ A(\mathbf{q})\mathbf{r}_2 \\ \vdots \\ A(\mathbf{q})\mathbf{r}_N \end{bmatrix}_k + \begin{bmatrix} \boldsymbol{\nu}_1 \\ \boldsymbol{\nu}_2 \\ \vdots \\ \boldsymbol{\nu}_N \end{bmatrix}_k \tag{24a}$$

$$R_k = \text{diag}\begin{bmatrix} \sigma_1^2 & \sigma_2^2 & \dots & \sigma_N^2 \end{bmatrix} \tag{24b}$$

where diag denotes a diagonal matrix of appropriate dimension. We should note that any attitude sensor, such as a three-axis magnetometer, star tracker, sun sensor, etc., can be put into the form given by Eq. (22). However, most sensors only observe two quantities, such as two angles in star trackers. The resulting form in Eq. (22) for these type of sensors has a unity norm constraint in the observation.[20] However, the mean observation given by Eq. (9) may not produce an estimate with unit norm. Therefore, we recommend that the original two quantity observation model be used for these types of sensors in the UF.

A common sensor that measures the angular rate is a rate-integrating gyro. For this sensor, a widely used model is given by[21]

$$\tilde{\boldsymbol{\omega}}(t) = \boldsymbol{\omega}(t) + \boldsymbol{\beta}(t) + \boldsymbol{\eta}_v(t) \tag{25a}$$

$$\dot{\boldsymbol{\beta}}(t) = \boldsymbol{\eta}_u(t) \tag{25b}$$

where $\tilde{\boldsymbol{\omega}}(t)$ is the continuous-time measured angular rate, and $\boldsymbol{\eta}_v(t)$ and $\boldsymbol{\eta}_u(t)$ are independent zero-mean Gaussian white-noise processes with

$$E\left\{\boldsymbol{\eta}_v(t)\boldsymbol{\eta}_v^T(\tau)\right\} = I_{3\times 3}\sigma_v^2\delta(t-\tau) \tag{26a}$$

$$E\left\{\boldsymbol{\eta}_u(t)\boldsymbol{\eta}_u^T(\tau)\right\} = I_{3\times 3}\sigma_u^2\delta(t-\tau) \tag{26b}$$

where $\delta(t-\tau)$ is the Dirac delta function.

In the standard EKF formulation, given a post-update estimate $\hat{\boldsymbol{\beta}}_k^+$, the post-update angular velocity and propagated gyro bias follow

$$\hat{\boldsymbol{\omega}}_k^+ = \tilde{\boldsymbol{\omega}}_k - \hat{\boldsymbol{\beta}}_k^+ \tag{27a}$$

$$\hat{\boldsymbol{\beta}}_{k+1}^- = \hat{\boldsymbol{\beta}}_k^+ \tag{27b}$$

Given post-update estimates $\hat{\boldsymbol{\omega}}_k^+$ and $\hat{\mathbf{q}}_k^+$, the propagated quaternion is found from the discrete-time equivalent of Eq. (19):

$$\hat{\mathbf{q}}_{k+1}^- = \Omega(\hat{\boldsymbol{\omega}}_k^+)\hat{\mathbf{q}}_k^+ \tag{28}$$

with

$$\Omega(\hat{\boldsymbol{\omega}}_k^+) \equiv \begin{bmatrix} Z_k & \hat{\boldsymbol{\psi}}_k^+ \\ -\hat{\boldsymbol{\psi}}_k^{+T} & \cos\left(0.5||\hat{\boldsymbol{\omega}}_k^+||\Delta t\right) \end{bmatrix} \tag{29}$$

$$Z_k \equiv \cos\left(0.5||\hat{\boldsymbol{\omega}}_k^+||\Delta t\right) I_{3\times 3} - \left[\hat{\boldsymbol{\psi}}_k^+ \times\right] \tag{30a}$$

$$\hat{\boldsymbol{\psi}}_k^+ \equiv \sin\left(0.5||\hat{\boldsymbol{\omega}}_k^+||\Delta t\right) \hat{\boldsymbol{\omega}}_k^+ / ||\hat{\boldsymbol{\omega}}_k^+|| \tag{30b}$$

where Δt is the sampling interval in the gyro.

UNSCENTED ATTITUDE FILTER

In this section an UF is derived for attitude estimation. We call the new filter the UnScented QUaternion Estimator, or USQUE, which is Latin for "all the way." One approach for the design of this filter involves using the quaternion kinematics in Eq. (28) directly. However, this approach has a clear deficiency. Mainly, referring to Eq. (7), since the predicted quaternion mean is derived using an averaged sum of quaternions, no guarantees can be made that the resulting quaternion will have unit norm. This makes the straightforward implementation of the UF with quaternions undesirable. Still, a filter can be designed using this approach where the quaternion is normalized by brute force. An eigenvalue/eigenvector decomposition is recommended to decompose the 7×7 covariance matrix, since this decomposition produces orthogonal sigma points. Our experience has shown that this approach can be successfully accomplished using the aforementioned procedure. It works well for small bias errors since the eigenvectors of the covariance matrix are nearly parallel with the attitude update; however, we have found that this approach does not work well when large bias updates occur.

A better way involves using an unconstrained three-component vector to represent an attitude error quaternion. We begin by defining the following state vector:

$$\boldsymbol{\chi}_k(0) = \hat{\mathbf{x}}_k^+ \equiv \begin{bmatrix} \boldsymbol{\delta}\hat{\mathbf{p}}_k^+ \\ \hat{\boldsymbol{\beta}}_k^+ \end{bmatrix} \tag{31}$$

We will use $\boldsymbol{\delta}\hat{\mathbf{p}}_k$ from Eq. (20) to propagate and update a nominal quaternion. Since this three-dimensional representation is unconstrained, the resulting overall covariance matrix is a 6×6 matrix. Therefore, using Eq. (7) poses no difficulties, which makes this an attractive approach. First, the vector $\boldsymbol{\chi}_k(i)$ in Eq. (5) is partitioned into

$$\boldsymbol{\chi}_k(i) \equiv \begin{bmatrix} \boldsymbol{\chi}_k^{\delta p}(i) \\ \boldsymbol{\chi}_k^{\beta}(i) \end{bmatrix}, \quad i = 0,\, 1,\, \ldots,\, 12 \tag{32}$$

where $\boldsymbol{\chi}_k^{\delta p}$ is from the attitude-error part and $\boldsymbol{\chi}_k^{\beta}$ is from the gyro bias part. To describe $\boldsymbol{\chi}_k^{\delta p}$ we first define a new quaternion generated by multiplying an error quaternion by the current estimate. Using the notation in Eq. (5) we assume

$$\hat{\mathbf{q}}_k^+(0) = \hat{\mathbf{q}}_k^+ \tag{33a}$$

$$\hat{\mathbf{q}}_k^+(i) = \boldsymbol{\delta}\mathbf{q}_k^+(i) \otimes \hat{\mathbf{q}}_k^+, \quad i = 1,\, 2,\, \ldots,\, 12 \tag{33b}$$

where $\boldsymbol{\delta}\mathbf{q}_k^+(i) \equiv \left[\boldsymbol{\delta}\boldsymbol{\varrho}_k^{+T}(i) \;\; \delta q_{4_k}^+(i)\right]^T$ is represented by Eq. (21):

$$\delta q_{4_k}^+(i) = \frac{-a\,||\boldsymbol{\chi}_k^{\delta p}(i)||^2 + f\sqrt{f^2 + (1-a^2)||\boldsymbol{\chi}_k^{\delta p}(i)||^2}}{f^2 + ||\boldsymbol{\chi}_k^{\delta p}(i)||^2} \quad i = 1,\, 2\,\ldots,\, 12 \tag{34a}$$

$$\boldsymbol{\delta}\boldsymbol{\varrho}_k^+(i) = f^{-1}\left[a + \delta q_{4_k}^+(i)\right]\boldsymbol{\chi}_k^{\delta p}(i), \quad i = 1,\, 2,\, \ldots,\, 12 \tag{34b}$$

Equation (33a) clearly requires that $\boldsymbol{\chi}_k^{\delta p}(0)$ be zero. This is due to the reset of the attitude error to zero after the previous update, which is used to move information from one part of the estimate to another part.[11] This reset rotates the reference frame for the covariance, so we might expect the covariance to be rotated, even though no new information is added. But the covariance depends on the assumed statistics of the measurements, not on the actual measurements. Therefore, since the update is zero-mean, the mean rotation caused by the reset is actually zero, so the covariance is in fact not affected by the reset. Next, the updated quaternions are propagated forward using Eq. (28), with

$$\hat{\mathbf{q}}_{k+1}^-(i) = \Omega\left[\hat{\boldsymbol{\omega}}_k^+(i)\right]\hat{\mathbf{q}}_k^+(i), \quad i = 0,\, 1,\, \ldots,\, 12 \tag{35}$$

where the estimated angular velocities are given by

$$\hat{\boldsymbol{\omega}}_k^+(i) = \tilde{\boldsymbol{\omega}}_k - \boldsymbol{\chi}_k^{\beta}(i), \quad i = 0,\, 1,\, \ldots,\, 12 \tag{36}$$

Note that $\hat{\boldsymbol{\omega}}_k^+(0) = \tilde{\boldsymbol{\omega}}_k - \boldsymbol{\chi}_k^{\beta}(0)$, where $\boldsymbol{\chi}_k^{\beta}(0)$ is the zeroth-bias sigma point given by the current estimate, i.e., $\boldsymbol{\chi}_k^{\beta}(0) = \hat{\boldsymbol{\beta}}_k^+$ from Eq. (5b). The propagated error quaternions are computed using

$$\boldsymbol{\delta}\mathbf{q}_{k+1}^-(i) = \hat{\mathbf{q}}_{k+1}^-(i) \otimes \left[\hat{\mathbf{q}}_{k+1}^-(0)\right]^{-1}, \quad i = 0,\, 1,\, \ldots,\, 12 \tag{37}$$

Note that $\boldsymbol{\delta}\mathbf{q}_{k+1}^-(0)$ is the identity quaternion. Finally, the propagated sigma points can be computed using the representation in Eq. (20):

$$\boldsymbol{\chi}_{k+1}^{\delta p}(0) = \mathbf{0} \tag{38a}$$

$$\boldsymbol{\chi}_{k+1}^{\delta p}(i) = f\frac{\boldsymbol{\delta}\boldsymbol{\varrho}_{k+1}^-(i)}{a + \delta q_{4_{k+1}}^-(i)}, \quad i = 1,\, 2,\, \ldots,\, 12 \tag{38b}$$

with $\begin{bmatrix}\boldsymbol{\delta\varrho}_{k+1}^{-T}(i) & \delta q_{4_{k+1}}^{-}(i)\end{bmatrix}^T = \boldsymbol{\delta q}_{k+1}^{-}(i)$. Furthermore, from Eq. (27b) we have

$$\boldsymbol{\chi}_{k+1}^{\beta}(i) = \boldsymbol{\chi}_{k}^{\beta}(i), \quad i = 0, 1, \ldots, 12 \tag{39}$$

The predicted mean and covariance can now be computed using Eqs. (7) and (8).

We now derive expressions for the matrices G_k and $\bar{Q}_k$, which are used in Eq. (14). Our choice of $f = 2(a+1)$ leads to the following approximation of the kinematics of the true attitude error:[11]

$$\boldsymbol{\delta\dot{p}}(t) = -\hat{\boldsymbol{\omega}}(t)\times\boldsymbol{\delta p}(t) + \boldsymbol{\delta\omega}(t) - \frac{1}{2}\boldsymbol{\delta\omega}(t)\times\boldsymbol{\delta p}(t) \tag{40}$$

where $\boldsymbol{\delta\omega}(t) \equiv \boldsymbol{\omega}(t) - \hat{\boldsymbol{\omega}}(t)$. Equation (40) is valid to second-order terms; higher-order terms have not been shown since they all have terms that are functions of $\boldsymbol{\delta p}(t)$. Since $\boldsymbol{\chi}_k^{\delta p}(0) = \mathbf{0}$ for all k, then from Eq. (40) the matrix G_k is simply given by the identity matrix. We should note that the full version of the approximation in Eq. (40) can be used to directly propagate the sigma points; however, we choose to propagate the error quaternions since closed-form solutions exist, given by Eq. (35). The conversion back to the sigma points using Eq. (38) requires fewer computations than using a numerical integration routine to directly propagate the sigma points. Still, in some cases the first-order approximation, which has a closed-form solution, can yield accurate results in the propagation. However, this may not take full advantage of the true benefits of the UF formulation, and suboptimal results may be obtained by this approach.

The derivation of the matrix $\bar{Q}_k$ assumes that the approximation $||\Delta t\,\hat{\boldsymbol{\omega}}_k^+|| << 1$ is valid, which is generally adequate for computing process noise. With this assumption the state transition matrix can be approximated by

$$\Phi(\Delta t) = \begin{bmatrix} I_{3\times3} & -\Delta t I_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{41}$$

The discrete process noise covariance is given by[21]

$$Q_k = \begin{bmatrix} \left(\sigma_v^2\Delta t + \frac{1}{3}\sigma_u^2\Delta t^3\right) I_{3\times3} & -\left(\frac{1}{2}\sigma_u^2\Delta t^2\right) I_{3\times3} \\ -\left(\frac{1}{2}\sigma_u^2\Delta t^2\right) I_{3\times3} & \left(\sigma_u^2\Delta t\right) I_{3\times3} \end{bmatrix} \tag{42}$$

Solving Eq. (14) for $\bar{Q}_k$ gives

$$\bar{Q}_k = \frac{\Delta t}{2}\begin{bmatrix} \left(\sigma_v^2 - \frac{1}{6}\sigma_u^2\Delta t^2\right) I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & \sigma_u^2 I_{3\times3} \end{bmatrix} \tag{43}$$

The correction term $(1/6)\sigma_u^2\Delta t^2$ is needed because the trapezoid rule is exact for linear functions, but overestimates the integral of a quadratic, namely $\Delta t^2\sigma_u^2$. The new process noise in Eq. (43) is diagonal, which seems to represent the physics more accurately. Simpson's rule, which is exact for quadratics, would eliminate the need for the correction term. However, the new process noise would have to be added at the midpoint of the propagation, i.e., at $\Delta t/2$, as well as at the beginning and end. This requires us to recompute the sigma points at time $\Delta t/2$, which is less convenient to implement than the trapezoid rule.

An exact closed-form solution of the state transition matrix $\Phi(\Delta t)$ can be found using the discrete-time methods shown in Ref. [22]. This should be used if the sampling interval is above Nyquist's limit.[23] In this case $\bar{Q}_k$ can be determined by solving the general form of the Lyapunov equation (also known as the Sylvester equation[24]) in Eq. (14) using numerical methods. From our experiences with the UF developed in this paper, adding the matrix $\bar{Q}_k$ at the beginning and end of the propagation does not seem to produce better results than using Q_k at the beginning or end of the propagation alone. Still, we favor the aesthetically pleasing nature of the trapezoidal approximation.

The USQUE algorithm for attitude estimation comes with a computational cost. An increased amount of computation is required for the covariance decomposition and sigma point propagations. Our experience has shown that the USQUE algorithm is about 2.5 times slower than the EKF in Ref. [2]. One approach to help reduce the computational load in USQUE takes advantage of using a lower-triangular Cholesky decomposition for the square root in Eq. (5a). With a lower-triangular form, only the first three columns have nonzero entries in the attitude parts. Thus we can bypass the computation of $\hat{\mathbf{q}}_k^+(i)$ in Eq. (33b) at the beginning of the propagation for half of the sigma points. This results in a savings of about 15% in the computational load. More dramatic reductions would be seen for larger state vectors, e.g., for gyroless attitude filtering with estimation of system momentum, environmental torques, inertia components, etc.

The procedure in USQUE is as follows. We are given initial attitude quaternion ($\hat{\mathbf{q}}_0^+$) and gyro-bias ($\boldsymbol{\beta}_0^+$) estimates, as well as an initial 6×6 covariance (P_0^+), where the upper 3×3 partition of P_0^+ corresponds to attitude error angles, and gyro parameters σ_u and σ_v. The initial state vector in USQUE is set to $\hat{\mathbf{x}}_0^+ = \begin{bmatrix}\mathbf{0}^T & \boldsymbol{\beta}_0^{+T}\end{bmatrix}^T$. We choose the parameter a and set $f = 2(a+1)$. Then, $\bar{Q}_k$ is calculated using Eq. (43), which will be used in Eqs. (5a) and (8). The sigma points are then calculated using Eq. (5). Next, the corresponding error quaternions are calculated using Eq. (34), where Eq. (33) is used to compute the sigma-point quaternions from the error quaternions. The quaternions are subsequently propagated forward in time using Eq. (35). Then, the propagated error quaternions are determined using Eq. (37), and the propagated sigma points are calculated using Eqs. (38) and (39). The predicted mean and covariance can now be computed using Eqs. (7) and (8). Storing the prop-

agated quaternions from Eq. (35) we then calculate the mean observations using Eqs. (9) and (10), with

$$\boldsymbol{\gamma}_{k+1}(i) = \begin{bmatrix} A[\hat{\mathbf{q}}^-(i)]\mathbf{r}_1 \\ A[\hat{\mathbf{q}}^-(i)]\mathbf{r}_2 \\ \vdots \\ A[\hat{\mathbf{q}}^-(i)]\mathbf{r}_N \end{bmatrix}_{k+1} \quad , \quad i = 0,\, 1,\, \ldots,\, 12 \tag{44}$$

The output covariance, innovation covariance and cross-correlation matrix are computed using Eqs. (11), (12) and (13). Next, the state vector and covariance are updated using Eq. (2), with $\hat{\mathbf{x}}^+_{k+1} \equiv \begin{bmatrix} \boldsymbol{\delta}\hat{\mathbf{p}}^{+T}_{k+1} & \hat{\boldsymbol{\beta}}^{+T}_{k+1} \end{bmatrix}^T$. Then, the quaternion is updated using

$$\hat{\mathbf{q}}^+_{k+1} = \boldsymbol{\delta}\mathbf{q}^+_{k+1} \otimes \hat{\mathbf{q}}^-_{k+1}(0) \tag{45}$$

where $\boldsymbol{\delta}\mathbf{q}^+_{k+1} \equiv \begin{bmatrix} \boldsymbol{\delta}\boldsymbol{\varrho}^{+T}_{k+1} & \delta q^+_{4_{k+1}} \end{bmatrix}^T$ is represented by Eq. (21):

$$\delta q^+_{4_{k+1}} = \frac{-a\,||\boldsymbol{\delta}\hat{\mathbf{p}}^+_{k+1}||^2 + f\sqrt{f^2 + (1-a^2)||\boldsymbol{\delta}\hat{\mathbf{p}}^+_{k+1}||^2}}{f^2 + ||\boldsymbol{\delta}\hat{\mathbf{p}}^+_{k+1}||^2} \tag{46a}$$

$$\boldsymbol{\delta}\boldsymbol{\varrho}^+_{k+1} = f^{-1}\left[a + \delta q^+_{4_{k+1}}\right]\boldsymbol{\delta}\hat{\mathbf{p}}^+_{k+1} \tag{46b}$$

Finally, $\boldsymbol{\delta}\hat{\mathbf{p}}^+_{k+1}$ is reset to zero for the next propagation.

SIMULATION EXAMPLES

In this section several performance comparisons between USQUE and the EKF are made through simulated examples using a realistic spacecraft model. The TRMM spacecraft is a representative Earth-pointing spacecraft in a near-circular 90 min (350 km) orbit with an inclination of 35°. The nominal Earth-pointing mission mode requires a rotation once per orbit about the spacecraft's y-axis. The attitude determination hardware consists of an Earth sensor assembly, digital sun sensors, coarse sun sensors, a three-axis magnetometer (TAM), and gyroscopic rate sensors. All simulations shown here assume only TAM and gyro measurements. The magnetic field reference is modeled using a 10^{th}-order International Geomagnetic Reference Field model.[25] TAM sensor noise is modeled by a zero-mean Gaussian white-noise process with a standard deviation of 50 nT. We note that the actual magnetic field errors have systematic components,[26] but these are not relevant to the present filter comparisons. The gyro measurements are simulated using Eq. (25) with $\sigma_u = 3.1623 \times 10^{-4}$ μrad/sec$^{3/2}$ and $\sigma_v = 0.31623$ μrad/sec$^{1/2}$, and an initial bias of 0.1 deg/hr on each axis.

For the first simulation both USQUE and the EKF are executed with no initial attitude errors, but with initial bias estimates set to zero. The initial covariance P^+_0 is diagonal with attitude error elements set to $P_{\text{att}} = (0.5\text{ deg})^2$ and bias error elements set to $P_{\text{bias}} = (0.2\text{ deg/hr})^2$. The gyro and TAM measurements are both sampled at 10 sec intervals. Also, in USQUE $a = 1$ with $f = 4$, which gives four times the vector of MRPs for the error representation, and $\lambda = 1$ (the optimality of this value will be discussed shortly). For this simulated run both USQUE and the EKF attitude errors agree to within 1 μrad. This indicates that no advantages to using USQUE with small errors can be seen.

In the next simulation, errors of $-50°$, $50°$ and $160°$ for each axis, respectively, are added into the initial-condition attitude estimate, with initial bias estimates set to zero. The initial attitude covariance is set to $(50\text{ deg})^2$ and the initial bias covariance is unchanged. The chosen attitude covariance gives a one standard-deviation error of 50°. Choosing $a = 1$ with $f = 4$ gives an MRP error of $4\tan^{-1}[(50 \times \pi)/(4 \times 180)] = 49.23°$. Therefore, the same value for P_{att} can be used in USQUE and the EKF without loss in accuracy

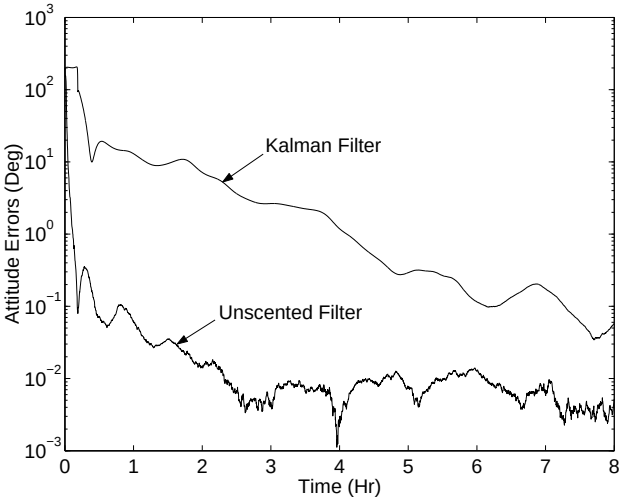

Fig. 1 Norm of Attitude Errors: Initial Attitude Errors Only

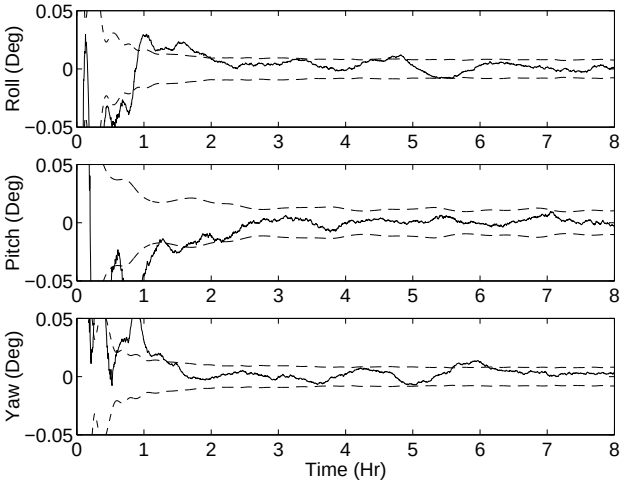

Fig. 2 Attitude Errors in USQUE with 3σ Error Bounds

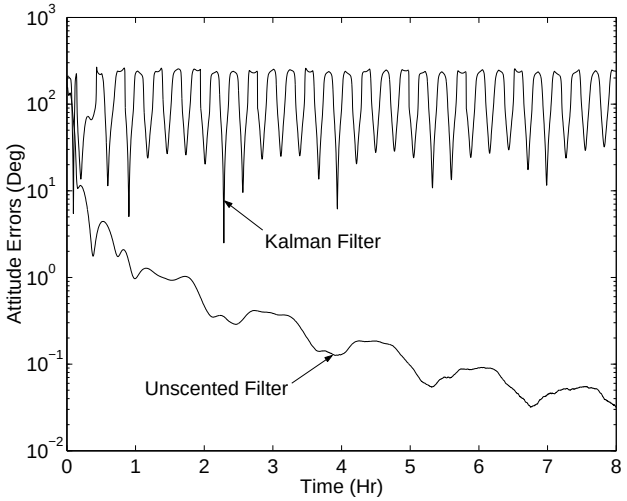

Fig. 3 Norm of Attitude Errors: Initial Attitude and Bias Errors

(note that two times the Gibbs vector gives 47.15°). This simulation illustrates a possible realistic scenario where only a TAM and a gyro exist with an unknown attitude estimate. A plot of the norm of the attitude errors for this simulation case is shown in Figure 1. The EKF takes almost 8 hours to converge to a value below 0.1°, while USQUE converges to this value in under 30 min. Also, the EKF attitude errors do not converge to within their respective 3σ error bounds until well after 8 hours. However, as shown in Figure 2, the USQUE attitude errors do converge to within their respective 3σ well within 8 hours, which indicates that USQUE is performing in a near optimal fashion.

A more dramatic situation is given by using the same attitude errors and P_{att}, but setting the initial y-axis bias estimate to 20 deg/hr. For both filters the initial bias covariance is set to $(20 \text{ deg/hr})^2$. A plot of the norm of the attitude errors for this simulation case is shown in Figure 3. The EKF never converges for this case since the first-order approximation cannot adequately capture the large initial condition errors. However, USQUE does converge to less than 0.1° in about 3.5 orbits. This clearly shows the importance of the higher-order terms used in the UF for convergence.

We now investigate the performance of USQUE for various values of a and λ. The first scenario includes initial attitude errors of −50°, 50° and 160° for each axis, respectively, with initial bias estimates set to zero. The second scenario also includes a large initial bias error, consistent with the previous simulations. For these simulations a is allowed to be greater than 1. Any a with absolute value greater than 1 risks the appearance of square roots of negative numbers, which is analogous to using the vector part of the quaternion error as the three-dimensional representation, i.e., the limit of infinite a. Still, values of a larger than 1 are worthy of study, due to the nonlinear nature of the estimation problem. In order to quantify the simulation

Table 1 Averaged J Values: Initial Attitude Errors Only

	$a=0$	$a=1$	$a=2$	$a=3$
$\lambda=-3$	18.98	14.98	10.02	10.06
$\lambda=-1$	5.97	4.28	4.63	4.70
$\lambda=0$	4.73	4.68	5.22	6.69
$\lambda=1$	5.30	**4.08**	9.29	6.60
$\lambda=3$	8.00	5.32	4.58	∞

Table 2 Averaged J Values: Initial Attitude and Bias Errors

	$a=0$	$a=1$	$a=2$	$a=3$
$\lambda=-3$	25.94	21.84	13.95	16.07
$\lambda=-1$	21.06	18.75	19.17	18.30
$\lambda=0$	20.71	19.49	20.59	22.97
$\lambda=1$	21.76	16.60	23.09	14.68
$\lambda=3$	25.77	13.87	**11.94**	∞

results, the norm of the attitude errors are numerically integrated over the 8 hour simulation run, given by

$$J = \int_0^8 ||\mathbf{e}(t)|| \, dt \tag{47}$$

where $\mathbf{e}(t)$ denotes the roll, pitch and yaw errors in degrees. Finally, Eq. (47) is averaged over 20 runs (i.e., a Monte-Carlo type simulation is performed).

Table 1 gives the averaged results of this integration for various values of a and λ using the first scenario of errors (i.e., large initial attitude errors only). The case with $a = 3$ and $\lambda = 3$ is unstable because the covariance does not remain positive definite. The best results are obtained with $a = 1$ and $\lambda = 1$. Other good results are obtained with $a = 1$ and $\lambda = 0$ or $\lambda = -1$, and $a = 2$ and $\lambda = -1$. The recommendation of using $\lambda = 3 - n$ is based an the approximation of the Gaussian distribution up to fourth-order terms. However, the Gaussian assumption is only valid through the first pass of the filter, since the nonlinear transformation causes a non-Gaussian distribution. Therefore, it is not surprising that values other than $\lambda = 3 - n$ yield better results.

Table 2 gives the averaged results of this integration for various values of a and λ using the second scenario of errors (i.e., large initial attitude and bias errors). For these initial errors, the best results are obtained with $a = 2$ and $\lambda = 3$. Even though the performance is better with $a > 1$, we do not recommend a value larger than 1, since the error vector has no physical meaning. A plot of the norm of the attitude errors for $a = 0$ to $a = 2$ with $\lambda = -3$ is shown in Figure 4. Note that the attitude errors asymptotically approach the same value. This is a typical result for various values

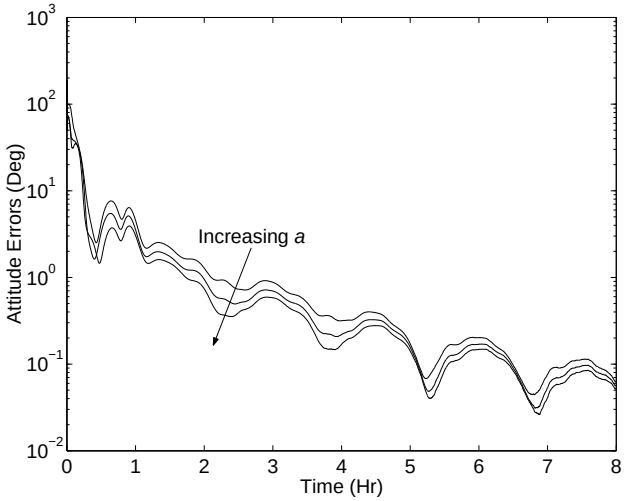

Fig. 4 Norm of Attitude Errors for $\lambda = -3$ and a Varying from 0 to 2

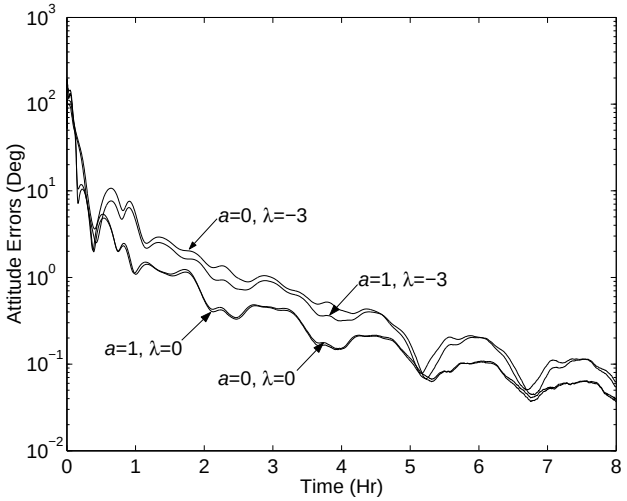

Fig. 5 Norm of Attitude Errors for Variations in Both a and λ

of a with constant values of λ. The performance then decreases as a becomes larger than 2 (not shown in the plot). The case with $a = 3$ and $\lambda = 3$ gives unstable filter results for this scenario, also. This can be overcome by using the scaled unscented transformation.[15] But, this is not required since $a = 3$ should not be used in practice.

From Tables 1 and 2 there exists a number of combinations for a and λ that yield approximately the same performance. The conclusion from these combinations is that the performance of USQUE is independent of the Gibbs vector or vector of MRPs formulations as long as the parameter λ is well chosen. Figure 5 shows the norm of the attitude errors with both large initial attitude and bias errors for various values of a and λ. Also, comparing Table 1 to Table 2 indicates a lack of some consistency for these two simulation scenarios. For example, from Table 1 when $a = 2$ the performance is better with $\lambda = -1$ than with $\lambda = -3$, while from Table 2 when $a = 2$ the performance is better with $\lambda = -3$ than with $\lambda = -1$. Also, a more dramatic performance increase is given when λ varies from -3 to -1 for all values of a in Table 1 than in Table 2. These findings are most likely due to the nonlinear nature of the problem. A further examination of Tables 1 and 2 indicates that $\lambda = 1$ seems to provide a good overall performance. Moreover, all reasonable values for a and λ outperform the standard EKF in all test scenarios for large initial condition errors.

CONCLUSIONS

In this paper a new approach for attitude estimation using the Unscented Filter formulation was derived. The new filter is based on a quaternion parameterization of the attitude. However, straightforward implementation of the Unscented Filter using quaternion kinematics did not produce a unit quaternion estimate. To overcome this difficulty the quaternion was represented by a three-dimensional vector of generalized Rodrigues parameters, which also reduced the size the covariance matrix. Simulation results indicated that the performance of the new filter far exceeds the standard Extended Kalman Filter for large initialization errors.

REFERENCES

[1]Farrell, J. L., "Attitude Determination by Kalman Filter," *Automatica*, Vol. 6, No. 5, 1970, pp. 419–430.

[2]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, Sept.-Oct. 1982, pp. 417–429.

[3]Crassidis, J. L. and Markley, F. L., "Attitude Estimation Using Modified Rodrigues Parameters," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, May 1996, pp. 71–83.

[4]Pittelkau, M. E., "Spacecraft Attitude Determination Using the Bortz Equation," *AAS/AIAA Astrodynamics Specialist Conference*, Quebec City, Quebec, August 2001, AAS 01-310.

[5]Gelb, A., editor, *Applied Optimal Estimation*, The MIT Press, Cambridge, MA, 1974, pp. 135–136.

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

[7]Kasdin, N. J., "Satellite Quaternion Estimation from Vector Measurements with the Two-Step Optimal Estimator," *AAS Guidance and Control Conference*, Breckenridge, CO, Feb. 2002, AAS 02-002.

[8]Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, Vol. AC-45, No. 3, March 2000, pp. 477–482.

[9]Julier, S. J. and Uhlmann, J. K., "A New Extension of the Kalman Filter to Nonlinear Systems," *Proceedings of the SPIE, Volume 3068, Signal Processing, Sensor Fusion, and Target Recognition VI*, edited by I. Kadar, Orlando, Florida, April 1997, pp. 182–193.

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

[11]Markley, F. L., "Attitude Representations for Kalman Filtering," *AAS/AIAA Astrodynamics Specialist Conference*, Quebec City, Quebec, Aug. 2001, AAS 01-309.

[12]Schaub, H. and Junkins, J. L., "Stereographic Orientation Parameters for Attitude Dynamics: A Generalization of the Rodrigues Parameters," *Journal of the Astronautical Sciences*, Vol. 44, No. 1, Jan.-March 1996, pp. 1–20.

[13]Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F., "A New Approach for Filtering Nonlinear Systems," *Proceedings of the American Control Conference*, Seattle, WA, June 1995, pp. 1628–1632.

[14]Bar-Shalom, Y. and Fortmann, T. E., *Tracking and Data Association*, Academic Press, Boston, MA, 1988, pp. 56–58.

[15]Julier, S. J., "The Scaled Unscented Transformation," *Proceedings of the American Control Conference*, Anchorage, AK, May 2002, pp. 1108–1114.

[16]Maybeck, P. S., *Stochastic Models, Estimation, and Control*, Vol. 2, Academic Press, New York, NY, 1982, pp. 217–221.

[17]Wan, E. and van der Merwe, R., "The Unscented Kalman Filter," *Kalman Filtering and Neural Networks*, edited by S. Haykin, chap. 7, John Wiley & Sons, New York, NY, 2001.

[18]Golub, G. H. and Van Loan, C. F., *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 2nd ed., 1989, pp. 145–146.

[19]van der Merwe, R. and Wan, E. A., "Square-Root Unscented Kalman Filter for State and Parameter-Estimation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, UT, May 2001, pp. 3461–3464.

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

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

[22]Markley, F. L., "Matrix and Vector Algebra," *Spacecraft Attitude Determination and Control*, edited by J. R. Wertz, Kluwer Academic Publishers, The Netherlands, 1978, pp. 754–755.

[23]Franklin, G. F., Powell, J. D., and Workman, M., *Digital Control of Dynamic Systems*, Addison Wesley Longman, Menlo Park, CA, 3rd ed., 1998, pp. 163–164.

[24]Maciejowski, J. M., *Multivariable Feedback Design*, Addison-Wesley Publishing Company, Wokingham, England, 1989, p. 385.

[25]Langel, R. A., "International Geomagnetic Reference Field: The Sixth Generation," *Journal of Geomagnetism and Geoelectricity*, Vol. 44, No. 9, 1992, pp. 679–707.

[26]Crassidis, J. L., Andrews, S. F., Markley, F. L., and Ha, K., "Contingency Designs for Attitude Determination of TRMM," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, May 1995, pp. 419–433.