Fully Multiplicative Unscented Kalman Filter for Attitude Estimation

Renato Zanetti[1]
The University of Texas at Austin, Austin, Texas 78712

Kyle J. DeMars[2]
Missouri University of Science and Technology, Rolla, Missouri, 65409

I. Introduction

The Kalman filter (KF) [1, 2] and its extension for nonlinear systems, the extended Kalman filter (EKF), are widely used algorithms in spacecraft navigation. To estimate spacecraft attitude, one favorite representation is the quaternion-of-rotation [3, 4]. Different approaches exist to enforce the unit-norm constraint of the quaternion-of-rotation (simply referred to as the quaternion from here on) in the Kalman filter, such as the multiplicative extended Kalman filter (MEKF) [5], the additive extended Kalman filter (AEKF) [6], as well as projection techniques [7], and constrained Kalman filtering [8]. Another extension of the Kalman filter for attitude estimation is the unscented quaternion estimator [9] that is based on the unscented Kalman filter (UKF) [10].

Like the EKF, the UKF is a linear estimator for nonlinear systems. Whereas the EKF employs linearization around the mean, the UKF utilizes stochastic linearization [11]. Stochastic linearization through a set of regression points employs the full set of nonlinear equations to estimate means and covariances for both the filter's propagation and update phases. As such, the UKF is capable of producing estimates of the means and covariances that are accurate to at least second order [12].

In this work, the filter update is performed utilizing unit vector measurements. Direction measurements from attitude sensors are often provided as bearing angles, but a unit vector can be easily derived from these angles. While it is possible to process the angular measurements directly, processing unit vectors is a widely adopted technique [6, 13]. The QUEST model [13] and the model

---

[1] Assistant Professor, Department of Aerospace Engineering and Engineering Mechanics, AIAA Associate Fellow
[2] Assistant Professor, Department of Mechanical and Aerospace Engineering, AIAA Senior Member

by Cheng *et al.* [14] are two unit vector measurement models in which the measurement error is represented as an additive contribution to the perfect measurement value. Mortari and Majji [15], on the other hand, introduced a multiplicative measurement model, which is the most natural unit vector model because it more closely represents the actual errors of the measurements. Contingent upon the linearization assumption, the multiplicative residual approach is equivalent to the additive residual approach [16], and the two are effectively identical to processing the bearing angles directly [17]. A key aspect of this work is the removal of the linearization assumption.

Both the EKF and the UKF are approximations of the optimal linear minimum mean square error (LMMSE) estimator. By performing a Taylor series and truncating to first order, the EKF defines the error as an additive quantity. This inconsistency with the nature of attitude is addressed by the MEKF by estimating attitude deviations, i.e. the series center of the attitude state is zero. By avoiding linearization around the mean, the UKF is capable of bypassing some of the limitations of linearization. At the heart of the UKF algorithm, however, the calculation of the sample mean and covariance from a set of sigma points is required. In every work to date, UKF-based attitude estimation approaches have computed these sample statistics by subtracting three-dimensional attitude deviations as if they were vectors.

This work presents a novel attitude UKF that has considerable conceptual and algorithmic differences from the attitude UKF of Crassidis and Markley [9]. Both a multiplicative measurement model and a multiplicative residual [16, 18] are utilized in this work, whereas Ref. [9] uses an additive measurement model and an additive residual. The propagation phase of the novel filter is also different from that of [9] in calculating the propagated estimated quaternion. Ref. [9] relies on the algebraic average of three-dimensional attitude parameterizations in order to compute the propagated quaternion. This work utilizes quaternion averaging [19], which provides the estimate with the minimum attitude error. Rather than minimizing the vector part of the attitude error as in Ref [19], this work minimizes the Gibbs error vector.

The consistent treatment of attitude and unit vectors results in robustness to extremely large measurement errors, validating the remarkable characteristic of the proposed algorithm of not relying on linearization or additivity but truly representing the nature of rotations. To demonstrate the robustness of the proposed scheme, a numerical example is provided with extremely large measurement errors, well beyond traditional sensor accuracy and even beyond their nominal field-of-view. Information is extracted through filtering many of these erratic measurements. Traditional attitude sensor models are not applicable to such large errors; hence, unit vectors and the multiplicative measurement model are employed [15].

II. The Unscented Kalman Filter

The EKF requires linearization of both the dynamics and measurement equations. The Kalman filtering paradigm, however, does not require that the models be linear. In fact, all that is required is that we have consistent, minimum variance estimates such that the distribution can be well-represented by its first two moments, that the measurement update be a linear scheme (that is, it is a linear combination of the prior state estimate and the measurement information), and that accurate predictions of the first two moments can be made [20]. The UKF works under similar assumptions.

This work considers a stochastic system whose state $\mathbf{x}$ evolves as

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_{k-1}) \,,$$

where $\mathbf{w}_{k-1}$ is a zero-mean, white sequence with covariance Q_{k-1} in conjunction with discrete measurements $\mathbf{y}_k$ that are modeled by

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k) \,,$$

where $\mathbf{v}_k$ is a zero-mean, white sequence with covariance R_k. The subscripts k and $k-1$ are used to denote the time index of the subscripted quantity.

A. Propagation

The UKF propagation step computes the *a priori* mean and covariance at time t_k (denoted $\hat{\mathbf{x}}_k^-$ and P_k^-, respectively) given the *a posteriori* mean and covariance at time t_{k-1} (denoted $\hat{\mathbf{x}}_{k-1}^+$ and P_{k-1}^+, respectively). An augmented state $\mathbf{z}_k$ is defined as

$$\mathbf{z}_{k-1}^{\mathrm{T}} = \begin{bmatrix} \mathbf{x}_{k-1}^{\mathrm{T}} & \mathbf{w}_{k-1}^{\mathrm{T}} \end{bmatrix} \,.$$

Let the set of sigma points for the augmented state be denoted by the N values of $\mathcal{Z}_{i,k-1}$ and the associated weights by w_i where $i \in \{1, \ldots, N\}$ and $\sum_{i=1}^{N} w_i = 1$. These sigma points are generated from the augmented mean and covariance given by

$$\mathbf{m}_{k-1} = \begin{bmatrix} \hat{\mathbf{x}}_{k-1}^{+} \\ \mathbf{0} \end{bmatrix} \qquad P_{k-1}^{\text{aug}} = \begin{bmatrix} P_{k-1}^{+} & \mathbf{O} \\ \mathbf{O} & Q_{k-1} \end{bmatrix},$$

and each of the sigma points is partitioned as

$$\mathcal{Z}_{i,k-1}^{\mathrm{T}} = \begin{bmatrix} \mathcal{X}_{i,k-1}^{\mathrm{T}} & \mathcal{W}_{i,k-1}^{\mathrm{T}} \end{bmatrix}.$$

The propagated sigma points are obtained via application of the nonlinear dynamical system, which gives

$$\mathcal{X}_{i,k} = \mathbf{f}(\mathcal{X}_{i,k-1}, \mathcal{W}_{i,k-1}).$$

These transformed sigma-points are then used to approximate the nonlinear transformation of the mean and the covariance via

$$\hat{\mathbf{x}}_k^- = \sum_{i=1}^{N} w_i \mathcal{X}_{i,k} \tag{1}$$

$$P_k^- = \sum_{i=1}^{N} w_i (\mathcal{X}_{i,k} - \hat{\mathbf{x}}_k^-)(\mathcal{X}_{i,k} - \hat{\mathbf{x}}_k^-)^{\mathrm{T}}. \tag{2}$$

While the effect of the process noise does not appear directly in these equations, it is captured through the propagated sigma points $\mathcal{X}_{i,k}$.

B. Update

Using the propagated mean and covariance at time t_k, a new set of sigma points is created. Again, the first step is to define an augmented state

$$\tilde{\mathbf{z}}_k^{\mathrm{T}} = \begin{bmatrix} \mathbf{x}_k^{\mathrm{T}} & \mathbf{v}_k^{\mathrm{T}} \end{bmatrix}$$

and generate sigma points, along with their associated weights $\tilde{w}_i$, with the mean and covariance

$$\tilde{\mathbf{m}}_k = \begin{bmatrix} \hat{\mathbf{x}}_k^{+} \\ \mathbf{0} \end{bmatrix} \qquad \tilde{P}_k^{\text{aug}} = \begin{bmatrix} P_k^{-} & O \\ O & R_k \end{bmatrix}.$$

In this case, the set of sigma points for the augmented state is denoted by the $\tilde{N}$ values of $\tilde{\mathcal{Z}}_{i,k}$ where there is no restriction that $N = \tilde{N}$; that is, the update step may employ a different number of sigma points than the propagation step. Similarly, there is no requirement that the weights of the sigma points in the update step are the same as the weights of the sigma points in the propagation step. Each of the sigma points is partitioned as

$$\tilde{\mathcal{Z}}_{i,k}^{\mathrm{T}} = \begin{bmatrix} \mathcal{X}_{i,k}^{\mathrm{T}} & \mathcal{V}_{i,k}^{\mathrm{T}} \end{bmatrix} ,$$

and the measurement-transformed sigma points are given by

$$\mathcal{Y}_{i,k} = \mathbf{h}(\mathcal{X}_{i,k}, \mathcal{V}_{i,k}) .$$

The expected value of the measurement, the measurement covariance, and the cross-covariance are found in terms of the transformed sigma-points as

$$\hat{\mathbf{y}}_k^- = \sum_{i=1}^{\tilde{N}} \tilde{w}_i \mathcal{Y}_{i,k} \tag{3}$$

$$P_{yy,k} = \sum_{i=1}^{\tilde{N}} \tilde{w}_i (\mathcal{Y}_{i,k} - \hat{\mathbf{y}}_k^-)(\mathcal{Y}_{i,k} - \hat{\mathbf{y}}^-)^{\mathrm{T}} \tag{4}$$

$$P_{xy,k} = \sum_{i=1}^{\tilde{N}} \tilde{w}_i (\mathcal{X}_{i,k} - \hat{\mathbf{x}}_k^-)(\mathcal{Y}_{i,k} - \hat{\mathbf{y}}^-)^{\mathrm{T}} . \tag{5}$$

The Kalman gain is $K_k = P_{xy,k}\ P_{yy,k}^{-1}$, and the associated updated state estimate and covariance are

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + K_k\ (\mathbf{y}_k - \hat{\mathbf{y}}_k^-) \tag{6}$$

$$P_k^+ = P_k^- - K_k\ P_{yy,k}\ K_k^{\mathrm{T}} . \tag{7}$$

C. Remarks

A few remarks regarding the nature of the UKF algorithm are in order. Firstly, it should be noted that sigma point generation relies on adding a deviation to the mean, where the deviation is generated from the covariance matrix. For states that utilize the quaternion description of attitude, this must be modified since the simple addition of a deviation to the quaternion will not, in general, result in a quaternion. Furthermore, the process of computing the propagated mean and covariance, Eqs. (1) and (2), relies on an averaging step and subtraction steps. Once again, when the state contains a quaternion, the averaging and subtraction steps need to be modified. Secondly, when considering the update stage of the UKF, vector subtraction is again utilized in Eq. (6). For situations in which unit vector measurements are to be processed, subtracting unit vectors will not yield a measurement residual which is also a unit vector. Therefore, when considering unit vector observations, the measurement update process of the UKF needs to be modified as well as the sample expectation calculations of Eqs. (3)–(5). All of these needed modifications may be grouped together as the removal of additivity within the UKF in favor of multiplicative steps.

III. Attitude Filter

In this section, we propose a novel attitude UKF; as is often done, a three-dimensional parameterization of the attitude is used to represent the estimation error. The proposed filter uses twice the Gibbs vector to represent the attitude error (denoted as $\delta\mathbf{g}$). The attitude covariance is obtained from this three-dimensional quantity. In developing the filtering equations, the nature of rotations is preserved by never adding or subtracting three-dimensional attitude parameterizations nor unit-vector direction measurements. The LMMSE estimate for nonlinear systems, of which the UKF is an approximation, seeks the estimate that minimizes the average of the square of the estimation error, which is usually defined as the Euclidean distance. The additive nature of the UKF described above is a direct result of the choice of the Euclidean distance. In this work, we minimize the error (defined as twice the Gibbs vector), obtaining the minimum mean-square Gibbs error attitude estimate rather than minimum Euclidean error estimate.

The proposed filter differs from that of Crassidis and Markley [9] both in the update and propagation phases, as the following sections describe. All "k" time subscripts are omitted for the remainder of the work for ease of notation. The development that follows only includes the attitude in the state vector; adding other estimated quantities, such as a gyro bias, follows from a simple extension of the resulting equations. The attitude state is the quaternion $\bar{\mathbf{q}}$ expressing the coordinate transformation from the inertial frame to the body-fixed frame. The attitude error is, as previously mentioned, twice the Gibbs vector representing the transformation from the estimated body frame to the true body frame.

A. Time Propagation

This portion of the algorithm calculates a propagated quaternion $\hat{\bar{\mathbf{q}}}$ and a propagated attitude error covariance matrix P^- that represent the estimates of how the attitude and its uncertainty evolve with time between measurement updates. During the propagation phase, a set of propagated sigma point quaternions $\hat{\bar{\mathbf{q}}}_j$, $j = 1, 2, \ldots, K$ is obtained following the same procedure of [9]; however, a different scheme is used to obtain the estimate. The desired estimate is the minimum mean-square error (MMSE) estimate. For a discrete random vector $\mathbf{X}$ with possible outcomes denoted by $\mathbf{x}_j$ and probability mass function p_j, the MMSE estimate $\hat{\mathbf{x}}$ minimizes

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \sum_j p_j \|\mathbf{x}_j - \mathbf{x}\|^2. \tag{8}$$

The solution of Eq. (8) is the mean of the random vector; that is,

$$\hat{\mathbf{x}} = \sum_j p_j \mathbf{x}_j \,. \tag{9}$$

Prior attitude UKF implementations took the three-dimensional attitude parameterizations of the propagated sigma point quaternions and performed their algebraic mean, effectively providing an estimate that minimizes the Euclidean distance between three-dimensional attitude errors. While this approach undoubtedly performs in a more than satisfactory fashion, it is more desirable to minimize an error defined as an attitude parameterization itself.

In this work, the attitude estimation error is defined as twice the Gibbs vector ($\delta\mathbf{g}$, also known as Rodrigues parameters). The goal is to obtain the attitude MMSE estimate, which means minimizing the performance index

$$\hat{\bar{\mathbf{q}}} = \min_{\hat{\bar{\mathbf{q}}}} \sum_{j=1}^{K} w_j \|\delta\hat{\mathbf{g}}_j^-\|^2 \,, \tag{10}$$

where

$$\bar{\mathbf{q}}(\delta\hat{\mathbf{g}}_j^-) = \delta\bar{\mathbf{q}} = \hat{\bar{\mathbf{q}}}_j \otimes \hat{\bar{\mathbf{q}}}^* \,, \tag{11}$$

the asterisk represents the quaternion conjugate, and the quaternion multiplication $\otimes$ composes quaternions in the same order as attitude matrices. Any other choice of attitude error representation will require a different performance index to be minimized and will produce a different estimate.

Besides the aforementioned Euclidean distance, the easiest choice of attitude error would be to minimize the vector part of the quaternion error. Such a choice would produce an estimated quaternion with a known solution obtained from solving a 4×4 eigenvalue problem [19]. The vector part of the quaternion is not a complete attitude parameterization; hence, in this work, we prefer to minimize an error that is physically a representation of attitude. Furthermore, the choice of the vector part of the quaternion as an error metric would produce undesirable effects during the measurement update phase of the algorithm, as detailed in the corresponding section of this paper.

The scaled Gibbs vector is given by

$$\delta\mathbf{g} = 2\delta\mathbf{q}_v/\delta q_s, \tag{12}$$

where the subscripts v and s indicate the vector and scalar parts of the quaternion, respectively. The propagated quaternion estimate $\hat{\bar{\mathbf{q}}}$ is obtained by solving Eq. (10) numerically with a simple recursion. In all numerical simulations, a Newton-Raphson method is used and it always converges in very few iterations. The initial guess is chosen as the average quaternion in terms of minimizing the vector part of the quaternion error rather than the Gibbs vector, which is obtained by calculating the unit eigenvector corresponding to the maximum eigenvalue of

$$\mathbf{M} = 4\sum_{j=1}^{K}\left(w_j\ \hat{\bar{\mathbf{q}}}_j\ \hat{\bar{\mathbf{q}}}_j^{\mathrm{T}}\right) - \mathbf{I}_{4\times 4}\,,$$

as shown in Ref. [19].

Particular care is also taken in computing the propagated attitude covariance. Once the estimated quaternion $\hat{\bar{\mathbf{q}}}$ is obtained from solving the eigenvalue problem, attitude deviations from the average quaternion are calculated for each sigma point with Eq. (11), and the three-dimensional deviations $\delta\hat{\mathbf{g}}_j^-$ are then calculated with Eq. (12); no algebraic mean is ever performed. The propagated covariance is given by

$$P^- = \sum_{j=1}^{K} w_j\ \delta\hat{\mathbf{g}}_j^-\ (\delta\hat{\mathbf{g}}_j^-)^{\mathrm{T}}\,. \tag{13}$$

Notice that this covariance calculation is also different from [9] since that work computes three-dimensional attitude deviations from the propagated mean quaternion and performs their algebraic mean, hence adding and subtracting attitudes. A common theme of the proposed algorithm is that rotations and unit vector (direction) measurements are never averaged as if they were Euclidean vectors.

Remark In order to reduce computations, it is possible to utilize the quaternion average proposed in [19] directly, as first done in [21], effectively producing the MMSE estimate where the error is defined as twice the vector part of the quaternion. In this work, however, we consistently define the error as the scaled Gibbs vector, which is a complete attitude parameterization.

B. Measurement Update

This proposed measurement update differs from [9] because a multiplicative measurement model is used as well as a multiplicative residual. The measurement model is given by

$$\mathbf{y} = T(\boldsymbol{\eta})\ T\ \mathbf{r}\,, \tag{14}$$

where $\boldsymbol{\eta}$ is a three-dimensional representation of the attitude error, for example a rotation vector, $T(\boldsymbol{\eta})$ represents the direction cosine matrix parameterization of $\boldsymbol{\eta}$, T is the inertial-to-body coordinate transformation matrix, and $\mathbf{r}$ is the true direction in the inertial frame. The classic additive measurement model is given by

$$\mathbf{y} = T\,\mathbf{r} + \mathbf{v}\,. \tag{15}$$

The additive measurement model relies on linearization (for example, the large field-of-view model from Cheng *et al.* [14] linearizes around the actual measurement). Therefore, for coarse sensors, a multiplicative measurement model is more accurate in representing the actual error. Since both the measurement $\mathbf{y}$ and the reference vector $\mathbf{r}$ are of unit length, it follows from Eq. (15) that

$$\mathbf{y}^{\mathrm{T}}\mathbf{y} = 1 = \mathbf{r}^{\mathrm{T}}T^{\mathrm{T}}T\mathbf{r} + 2\mathbf{v}^{\mathrm{T}}T\mathbf{r} + \mathbf{v}^{\mathrm{T}}\mathbf{v} = 1 + 2\mathbf{v}^{\mathrm{T}}T\mathbf{r} + \mathbf{v}^{\mathrm{T}}\mathbf{v}\,. \tag{16}$$

Taking expected values in Eq. (16) and using the fact that $\mathbf{r}$ is deterministic,

$$2\mathbf{r}^{\mathrm{T}}\ T\ \ \mathrm{E}\left\{\mathbf{v}\right\} = -\,\mathrm{trace}\,\mathrm{E}\left\{\mathbf{v}\mathbf{v}^{\mathrm{T}}\right\}. \tag{17}$$

Eq. (17) implies that, for the classic additive measurement model of Eq. (15), the measurement noise is either zero mean with zero covariance or not zero mean, i.e. the measurement is biased.

Using the multiplicative measurement model in the UKF overcomes the bias in the measurement error and allows an unbiased estimator to be obtained. Furthermore, one of the strengths of the UKF is that it avoids linearization around the mean; it is, therefore, more consistent to utilize a measurement model that also does not rely on linearization around the actual measurement.

The second feature of the proposed update methodology is a multiplicative residual; unit vectors representing directions are not subtracted as if they were vectors in $\Re^3$. The attitude update is given by

$$\delta\hat{\mathbf{g}}^+ = \delta\hat{\mathbf{g}}^- + \mathbf{K}\boldsymbol{\epsilon}\,, \tag{18}$$

where $\boldsymbol{\epsilon}$ is the multiplicative residual and, once again, the attitude error $\delta\mathbf{g}$ is twice the Gibbs vector defined as

$$\bar{\mathbf{q}}(\delta\mathbf{g}) = \bar{\mathbf{q}} \otimes \hat{\bar{\mathbf{q}}}^*\,, \tag{19}$$

where $\bar{\mathbf{q}}$ is the true (unknown) inertial-to-body quaternion. From Eq. (19) it follows immediately that $\delta\hat{\mathbf{g}}^- = \mathbf{0}$; therefore, in fact, attitudes are never added together.

The residual expresses the "distance" between the actual measurement and the expected measurement; the greater this distance, the greater the update. To be consistent with our approach, we define the residual $\boldsymbol{\epsilon}$ as the scaled Gibbs vector that expresses the rotation to take $\hat{\mathbf{y}}$ into $\mathbf{y}$. There are infinite such rotations, so the minimum one is chosen, which is to say that we choose the Gibbs vector to be perpendicular to both $\hat{\mathbf{y}}$ and $\mathbf{y}$, which yields

$$\boldsymbol{\epsilon} = 2\frac{\hat{\mathbf{y}} \times \mathbf{y}}{1 + \hat{\mathbf{y}} \cdot \mathbf{y}}\,, \tag{20}$$

where $\mathbf{y}$ is the unit vector measurement, which is one realization of the random vector $\mathbf{Y}$, and $\hat{\mathbf{y}}$ is the "average" measurement. Using the same logic employed before, the "average" measurement is the unit vector $\hat{\mathbf{y}}$ that minimizes the distance to all possible realizations of $\mathbf{Y}$. The distance is defined in terms of the Gibbs vector. Assuming a discrete distribution with possible outcomes denoted by $\mathbf{y}_j$ and probability mass function p_j

$$\hat{\mathbf{y}} = \min_{\hat{\mathbf{y}}} \sum_j p_j \|\boldsymbol{\epsilon}_j\|^2 = \min_{\hat{\mathbf{y}}} \sum_j p_j \,\frac{\|\hat{\mathbf{y}} \times \mathbf{y}_j\|^2}{(1 + \hat{\mathbf{y}} \cdot \mathbf{y}_j)^2} \qquad \text{subject to } \|\hat{\mathbf{y}}\| = 1\,.$$

The minimizing value of $\hat{\mathbf{y}}$ is obtained numerically with a simple recursion.

Notice that Eq. (18) can be re-written as

$$\delta\hat{\mathbf{g}}^{+} = \delta\hat{\mathbf{g}}^{-} + \mathbf{K}(\mathbf{z} - \hat{\mathbf{z}}) \,, \tag{21}$$

where the auxiliary variable $\mathbf{z}$ is defined as $\mathbf{z} = 2(\hat{\mathbf{y}} \times \mathbf{y})/(1 + \hat{\mathbf{y}} \cdot \mathbf{y})$ and has zero mean, $\hat{\mathbf{z}} = \mathbf{0}$. Therefore this approach effectively seeks the MMSE estimate of $\mathbf{x}$ given the measurement $\mathbf{z}$. The proposed update is rewritten in the standard UKF form utilizing the auxiliary variable $\mathbf{z}$ and all the UKF properties still hold.

The sigma points are obtained from the augmented covariance

$$P^{\text{aug}} = \begin{bmatrix} P^{-} & O \\ O & R \end{bmatrix} \,, \tag{22}$$

where P^{-} is the *a priori* estimation error covariance and R is the measurement noise ($\boldsymbol{\eta}$) covariance. Because of the multiplicative measurement model of Eq. (14), R is chosen full-rank without any approximation. Linearized additive measurement models, on the other hand, possess a rank-deficient measurement error covariance. With the $n \times n$ matrix P^{aug} defined above, the $2n+1$ sigma points are given by

$$\mathcal{X}_0 = \mathbf{0} \tag{23}$$

$$\mathcal{X}_i = \sqrt{(n+\kappa)\; P^{\text{aug}}_i} \tag{24}$$

$$\mathcal{X}_{i+n} = -\sqrt{(n+\kappa)\; P^{\text{aug}}_i} \,, \tag{25}$$

where $i = 1, \ldots, n$ and $\sqrt{\mathbf{A}_i}$ is the i^{th} column of the matrix square root of $\mathbf{A}$, which is defined such that $\mathbf{A} = (\sqrt{\mathbf{A}})(\sqrt{\mathbf{A}})^{\text{T}}$. Along with the sigma points, weights are chosen as

$$w_0 = \kappa/(n+\kappa) \qquad\qquad w_i = 0.5/(n+\kappa) \,, \tag{26}$$

where κ is a design parameter of the UKF. Once the sigma points are obtained, they are transformed through the nonlinear measurement function as

$$\mathcal{Y}_i = \mathbf{h}(\mathcal{X}_i, \mathbf{r}, \hat{\bar{\mathbf{q}}}) \,, \tag{27}$$

where

$$\mathbf{h}(\mathcal{X}_i, \mathbf{r}, \hat{\bar{\mathbf{q}}}) = T(\mathcal{N}_i)\; T(\delta\mathcal{G}_i)\; T(\hat{\bar{\mathbf{q}}})\; \mathbf{r} \,. \tag{28}$$

In Eq. (28), $\delta\mathcal{G}_i$ and $\mathcal{N}_i$ are the elements that compose the input sigma points; that is,

$$\mathcal{X}_i^{\mathrm{T}} = \left[\delta\mathcal{G}_i^{\mathrm{T}} \;\; \mathcal{N}_i^{\mathrm{T}}\right] .$$

The mean and covariance of the transformed variables are found via

$$\hat{\mathbf{y}} = \min_{\boldsymbol{\xi}} \sum_{i=0}^{2n} w_i \frac{\boldsymbol{\xi} \times \mathcal{Y}_i}{1 + \boldsymbol{\xi} \cdot \mathcal{Y}_i}, \quad \|\boldsymbol{\xi}\| = 1 \qquad\qquad \mathcal{Z}_i = 2\frac{\hat{\mathbf{y}} \times \mathcal{Y}_i}{1 + \hat{\mathbf{y}} \cdot \mathcal{Y}_i} \tag{29}$$

$$P_{zz} = \sum_{i=0}^{2n} w_i \; \mathcal{Z}_i \; \mathcal{Z}_i^{\mathrm{T}} \qquad\qquad P_{xz} = \sum_{i=0}^{2n} w_i \; \delta\mathcal{G}_i \; \mathcal{Z}_i^{\mathrm{T}} . \tag{30}$$

The updated state and covariance are obtained from Eq. (18) and

$$K = P_{xz} P_{zz}^{\dagger} \tag{31}$$

$$P^{+} = P^{-} - K \; P_{zz} \; K^{\mathrm{T}} , \tag{32}$$

where the pseudoinverse $\dagger$ provides the optimal estimate given the singular covariance P_{zz}. Finally, the quaternion is updated as

$$\hat{\bar{\mathbf{q}}} \leftarrow \bar{\mathbf{q}}(\delta\hat{\mathbf{g}}^{+}) \otimes \hat{\bar{\mathbf{q}}} . \tag{33}$$

Remark Once again, the system designer could choose to represent the error as twice the vector part of the quaternion rather than the scaled Gibbs vector, effectively minimizing $\sin^2\theta/2$ rather than $\tan^2\theta/2$, where θ is the Euler angle. For most, if not all, spacecraft applications, this alternative approach will work very well. However, the inherent constraint that each element of the vector part of the quaternion must be less than one can create problems in the presence of large attitude errors. A portion of the attitude sigma points are obtained as

$$\mathcal{X}_i = \sqrt{(n+\kappa) P_i^{\mathrm{aug}}} ,$$

and there is no guarantee that for very uncertain systems $\sqrt{(n+\kappa)}$ will not scale the components of the attitude sigma points beyond unity. By choosing to represent attitude errors as Gibbs vectors, not only do we employ a full attitude parameterization, but we take advantage that the Gibbs attitude error is (almost) a one-to-one parameterization of attitude with the only exception/singularity being the 180 degree error. In the context of generating sigma points, that singularity is actually very helpful, because it prevents attitude sigma-points from wrapping around, allowing an extremely robust UKF design even for extremely large attitude uncertainties.

IV. Numerical Results

To demonstrate the validity of the proposed approach, we consider a satellite attitude tracking problem in which the orbit is perfectly known, but the attitude is not. The satellite is taken to be in near-geosynchronous orbit with Keplerian elements as shown in Table 1.

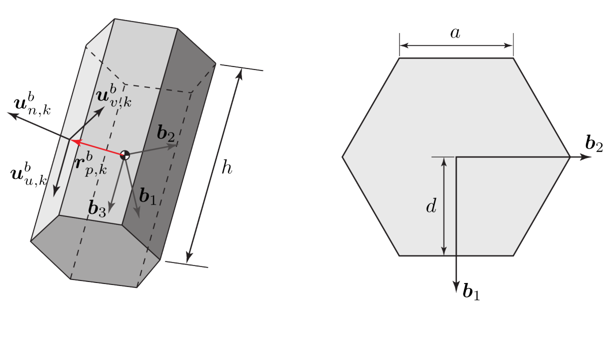
Table 1 Satellite Orbit

Type	Value	Units
Semi-Major axis	43000	km
Eccentricity	0.03	nd
Inclination	3	deg
RAAN	0	deg
Argument of Periapsis	0	deg
Mean Anomaly	0	deg

To generate a true attitude profile, we take the rotational dynamics to be

$$\dot{\bar{\mathbf{q}}} = \frac{1}{2}\bar{\boldsymbol{\omega}} \otimes \bar{\mathbf{q}}$$
$$\dot{\boldsymbol{\omega}} = J^{-1}\left(\sum \mathbf{m} - \boldsymbol{\omega} \times J\boldsymbol{\omega}\right),$$

where $\bar{\boldsymbol{\omega}}$ is the pure quaternion formed from the angular velocity vector $\boldsymbol{\omega}$, J is the moment of inertia of the spacecraft, and $\sum \mathbf{m}$ represents the summation of all active moments in the body frame. The active moments are assumed to be zero in this work. The computation of the moment of inertia depends upon the mass distribution of the object. In this work, the object is assumed to be a hexagonal prism, as shown in Figure 1. This is an 8-plate model with the body-frame unit



(a) 3D view (b) End view

Fig. 1 Satellite hexagonal prism flat plate model.

vectors defined by the unit vector triad $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. Additionally, the plate normal, denoted for the

k^{th} plate by $\mathbf{u}^b_{n,k}$, is depicted in Figure 1. The area, A_k, and position from the object center, $\mathbf{r}^b_{p,k}$, of each plate are fully determined by specifying the side length, a, and the prism height, h. The distance of the side from the center, d, can be determined from the length of the side, a. The size parameters are chosen so as to represent a typical spacecraft size; the values used, along with the total mass of the object, are presented in Table 2. Based upon the mass, side length, prism height, and the distance from the center to the side, the moment of inertia can be found to be a diagonal matrix of the form

$$J = \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix},$$

where the elements of J are given by

$$I_{xx} = I_{yy} = m\left(\frac{a^2}{6} + \frac{d^2}{3} + \frac{h^2}{12}\right)$$
$$I_{zz} = m\left(\frac{a^2}{6} + \frac{d^2}{3}\right).$$

The resulting inertia values are also summarized in Table 2.

Table 2 Satellite Geometry and Characteristics

Type	Value	Units
Length of side	2	m
Height of side	4	m
Distance of side from center	1.7	m
Mass	2688	kg
I_{xx} and I_{yy} Inertia	8100	kg m^2
I_{zz} Inertia	4500	kg m^2

The proposed algorithm is tested with a case where most existing linear filters for attitude fail to provide a consistent estimate. It is assumed that the initial attitude has a mean orientation given by the identity quaternion and that the initial mean angular velocity is taken to be zero. True values are generated by sampling a Gaussian error distribution with a standard deviation of 50° in attitude and 0.1°/s in angular velocity. The equations of motion are applied to generate a true attitude and angular velocity profile.

The satellite is equipped with a three-axis rate-integrating gyro that provides incremental angular changes at 100 Hz. The gyro measurements are generated by integrating the true angular velocity signal at the 100 Hz frequency and then subjecting the true integrated signal to a zero-mean bias and a zero-mean white-noise sequence. The statistics of the gyro bias and noise are given in Table 3. In addition to the gyro, the satellite is equipped with a sun sensor and an Earth sensor operating at 1 Hz, which provide unit vector measurements that point to the sun and Earth, respectively. The pointing vectors are generated based on the specified (known) orbit and the uncertain attitude; that is, the unit vector measurement for each of the sensors follows Eq. (14) and is computed as

$$\mathbf{y} = T(\boldsymbol{\eta})\ T\ \mathbf{r}\,, \tag{34}$$

where $\mathbf{r}$ is the reference vector in the inertial (i) frame, T is the true inertial-to-body direction cosine matrix (DCM), and $T(\boldsymbol{\eta})$ is the DCM parameterization of the three-dimensional attitude measurement error $\boldsymbol{\eta}$, which is taken to be zero mean. The measurement error $\boldsymbol{\eta}$ has covariance $\sigma^2 I_{3\times3}$, where the standard deviation, σ, is specified for each sensor in Table 3.

Table 3 Sensor Specifications

Type	1σ Error	Units
Gyro Noise	1	deg/$\sqrt{\text{s}}$
Initial Gyro Bias	1	deg/s
Sun Sensor Error	50	deg
Earth Sensor Error	50	deg

The proposed attitude filter is initialized with a starting estimated quaternion equal to the identity quaternion and the estimated bias equal to zero. The UKF parameter is set to $\kappa = 0$ to avoid any issues with a lack of positive definiteness in the filter's covariance matrix that sometimes plague UKF designs that choose $\kappa = 3 - n$ when $n > 3$ and large covariances. The resulting attitude and gyro bias estimation performance is summarized in Figures 2 and 3, where the gray line shows the estimation error while the black lines show the predicted 3σ error standard deviation. It can be seen that the filter is capable of providing a consistent estimate and reducing the estimation error to below 100° (3σ) given the very erratic measurements. Typical of highly nonlinear and highly uncertain systems (the measurement noise can easily jump by more than 90°), the predicted covariance is highly dependent on the value of the estimated measurement (and hence on the value of the estimated state). Different measurement sequences produce different values of the estimated state and of the estimated covariance.

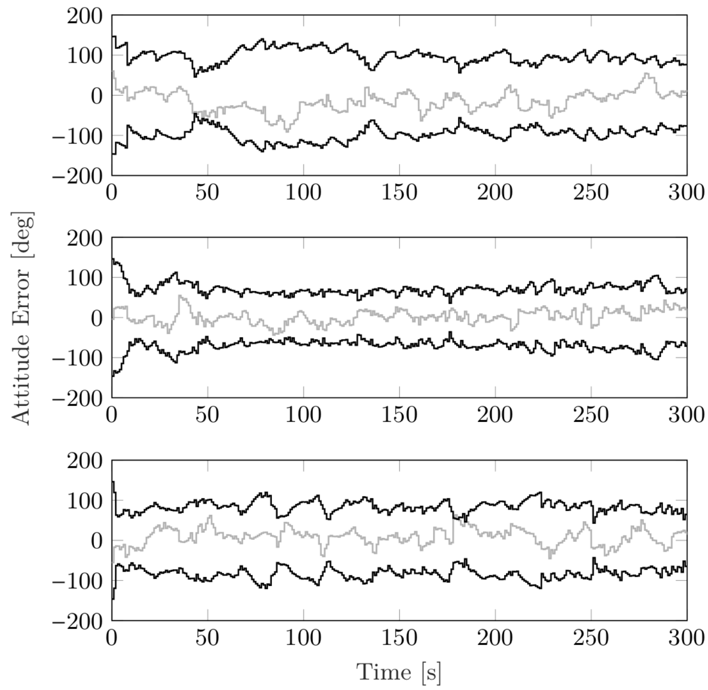


Fig. 2 Attitude estimation error - single run

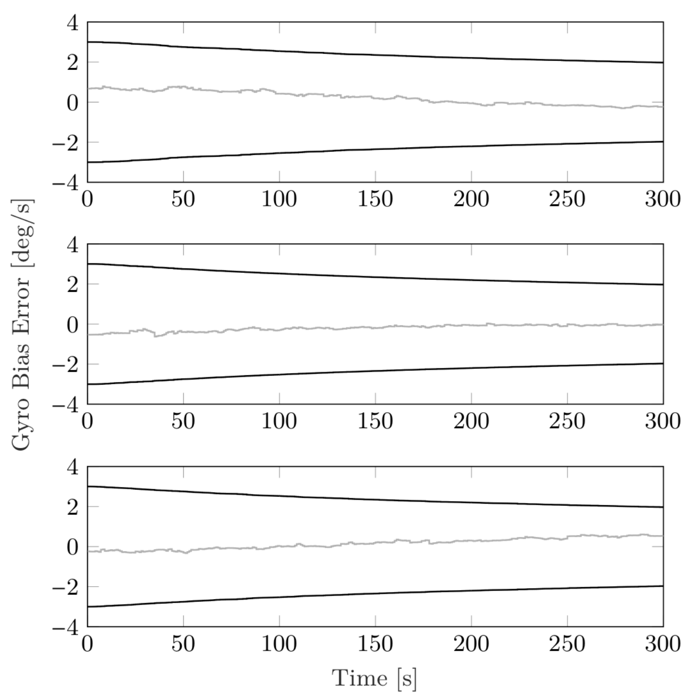


Fig. 3 Gyro bias estimation error - single run

To further investigate the performance of the filter and to demonstrate the consistency of the filter's covariance, a Monte Carlo simulation consisting of 100 runs is considered. On each run of the Monte Carlo simulation, a new true attitude and angular velocity profile is created by sampling the Gaussian distribution assumed for the initial attitude uncertainty and angular velocity uncertainty. Each run of the Monte Carlo simulation is also characterized by a different true gyro bias. At each time step of each simulation, the noises for the gyro measurements and the pointing vector measurements are re-sampled according to the statistics provided in Table 3. While unscented transformations provide exact mean and covariance random variable transformations of linear functions, they are an approximation for nonlinear functions and some tuning is sometimes required. In highly nonlinear systems (either nonlinearity is significant in a small interval around the mean, or the uncertainty is large such that the nonlinearity is overall significant in the large domain of interest), tuning is used to achieve the desired performance of the filter. In UKF implementations one usually tweaks the two values of κ independently for both the propagation and update phases in order to spread the sigma points the desired amount and obtain good performance. We choose an alternative approach. As previously noted, a negative value of κ can create non-positive definite posterior covariances; we therefore chose to utilize $2n$ sigma points by setting $\kappa = 0$. The tuning of the spread of the sigma points is therefore achieved in a similar fashion as EKFs are usually tuned: by tweaking the noise covariance matrices. The process noise and measurement noise are therefore altered so that the Monte Carlo statistics are brought into agreement with the filter statistics. As such, the process and measurement noise covariances are parameterized as

$$Q_{filter} = f_q Q_{true} \qquad\qquad R_{filter} = f_r R_{true}\,,$$

and it is found that $f_q = 2$ and $f_r = 1.2^2$ yield excellent performance. The results of a single filter run are compared against the sample covariance results from the Monte Carlo simulation in Figures 4 and 5, where the single filter run 3σ curves are plotted in black and the sample covariance 3σ curves are plotted as gray dashed lines. The figures show a very good match between the sample covariances and the predicted covariance of the filter, demonstrating the filter is performing as expected. In order to show the performance of every Monte Carlo run, Figure 6 shows the absolute value of the attitude error scaled by the filter's predicted standard deviation for each Monte Carlo simulation (gray lines) and the root-sum-square (RSS) of these quantities (black line). From Figure 6, it can be seen that the RSS is very close to one, further indicating that the filter is correctly

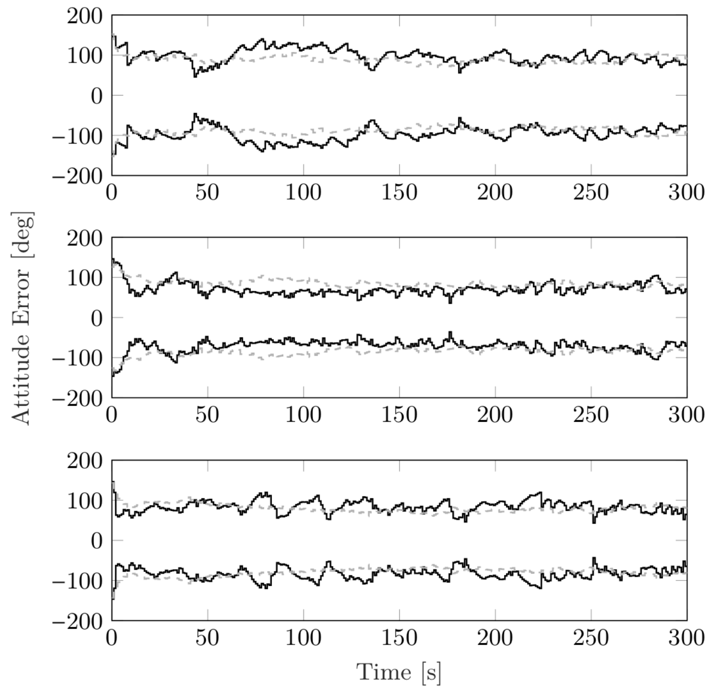


Fig. 4 Attitude estimation error - Monte Carlo

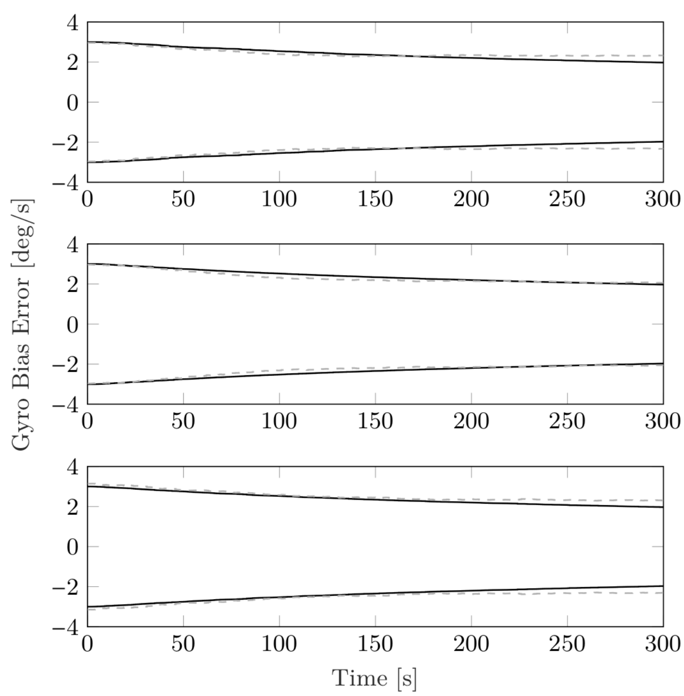


Fig. 5 Gyro bias estimation error - Monte Carlo

predicting its performance.

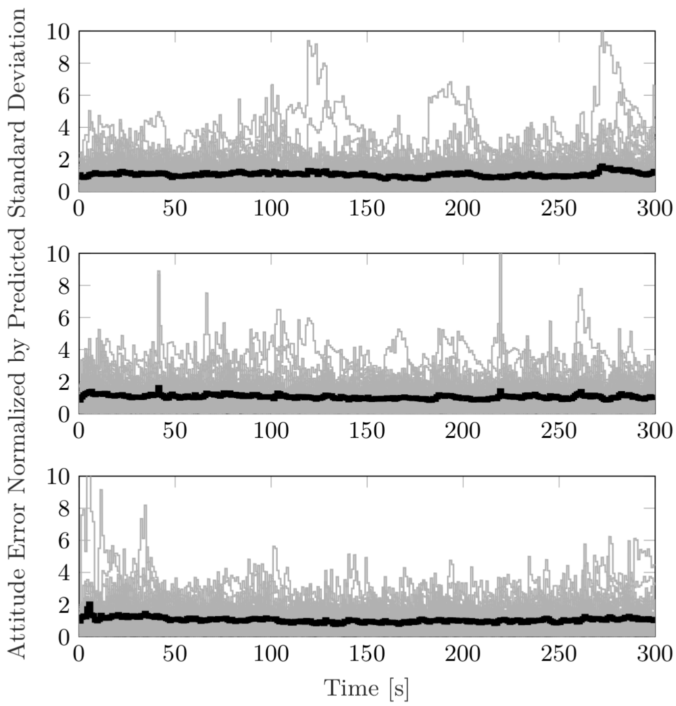


Fig. 6 Normalized attitude estimation error

V. Conclusions

This work presents a novel unit-vector quaternion unscented filter with multiplicative residual and multiplicative measurement error models. The proposed algorithm treats all attitude errors consistently as Gibbs vectors and consistently with the nature of rotations; no additions, subtractions, or algebraic averages are ever performed. The work develops an algorithm that does not rely on linearization nor small angle assumptions. Previous attitude unscented Kalman filter works relied on an additive measurement model that requires linearization during the update phase of the algorithm. During propagation, the various quaternions obtained from the propagated sigma points are transformed into three-dimensional attitude deviations and simply averaged together. Such an average is only valid for small angles, whereas the proposed algorithm averages the quaternions taking in full consideration the inherent non-Euclidean nature of the rotation group. Similar steps are taken to produce a linear minimum mean square Gibbs attitude error estimator when processing unit vector measurements. While the Gibbs vector is singular for rotations of 180 degrees, it possesses very desirable properties to represent errors. Because of the aforementioned characteristics, the proposed algorithm is robust to very large attitude estimation errors, as shown in simulations.

References

[1] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, Vol. 82, No. Series D, March 1960, pp. 35–45, doi:10.1115/1.3662552.

[2] Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering and Prediction," *Journal of Basic Engineering*, Vol. 83, No. Series D, March 1961, pp. 95–108, doi:10.1115/1.3658902.

[3] Kuipers, J. B., *Quaternions and Rotation Sequences*, Princeton University Press, Princeton, NJ, 1999.

[4] Markley, F. L., "Attitude Estimation or Quaternion Estimation?" *The Journal of the Astronautical Sciences*, Vol. 52, No. 1/2, January–June 2004, pp. 221–238.

[5] Lefferts, E. J., Markley, F. L., and Shuster, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429, doi:10.2514/6.1982-70.

[6] Bar-Itzhack, I. Y. and Oshman, Y., "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 21, No. 1, January 1985, pp. 128–135, doi: 10.1109/TAES.1985.310546 .

[7] Simon, D. and Chia, T. L., "Kalman Filtering with State Equality Constraints," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, No. 1, January 2002, pp. 128–136, doi: 10.1109/7.993234.

[8] Zanetti, R., Majji, M., Bishop, R. H., and Mortari, D., "Norm-Constrained Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 5, September–October 2009, pp. 1458–1465, doi: 10.2514/1.43119.

[9] Crassidis, J. L. and Markley, F. L., "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance Control and Dynamics*, Vol. 26, No. 4, July - August 2003, pp. 536–541, 10.2514/2.5102.

[10] 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. 45, No. 3, March 2000, pp. 477–482, doi: 10.1109/9.847726.

[11] Lefebvre, T., Bruyninckx, H., and Schutter, J. D., "Comment on "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators"," *IEEE Transactions on Automatic Control*, Vol. 47, No. 8, 2002, pp. 1406–1408, doi: 10.1109/TAC.2002.800742.

[12] Julier, S. J. and Uhlmann, J. K., "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, Vol. 92, No. 3, March 2004, pp. 401–422, doi: 10.1109/JPROC.2003.823141.

[13] Shuster, M. D., "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences*, Vol. 38, No. 3, July–September 1990, pp. 377–393.

[14] Cheng, Y., Crassidis, J. L., and Markley, F. L., "Attitude Estimation for Large Field-of-View Sensors,"

The Journal of the Astronautical Sciences, Vol. 54, No. 3 and 4, July–December 2006, pp. 433–448, doi:10.1007/BF03256499.

[15] Mortari, D. and Majji, M., "Multiplicative Measurement Model and Single-Point Attitude Estimation," *Proceedings of the F. Landis Markley Astronautics Symposium held June 30 – July 2, 2008, Cambridge, Maryland*, Vol. 132 of *Advances in the Astronautical Sciences*, 2008, pp. 51–69, AAS 08-263.

[16] Zanetti, R., "A Multiplicative Residual Approach to Attitude Kalman Filtering with Unit-Vector Measurements," *The Journal of the Astronautical Sciences*, Vol. 57, No. 4, October–December 2009, pp. 793–801, doi: 10.1007/BF03321530.

[17] Hinks, J. C. and Crassidis, J. L., "Covariance Analysis of Maximum Likelihood Attitude Estimation," *The Journal of the Astronautical Sciences*, Vol. 60, No. 2, Jun 2013, pp. 186–210, doi:10.1007/s40295-014-0028-7.

[18] Crassidis, J. L., Andrews, S. F., Markley, F. L., and Ha, K., "Contingency Designs for Attitude Determination of TRMM," *NASA/GSFC Flight Mechanics/Estimation Theory Symposium*, NASA/CP 3299, Greenbelt, MD, May 1995, pp. 419–433.

[19] Markley, F. L., Cheng, Y., Crassidis, J. L., and Oshman, Y., "Averaging Quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 4, July–August 2007, pp. 1193–1197, doi: 10.2514/1.28949.

[20] van der Merwe, R., *Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*, Ph.D. thesis, Oregon Health and Science University, 2004.

[21] Zanetti, R., DeMars, K. J., and Mortari, D., "Novel Multiplicative Unscented Kalman Filter for Attitude Estimation," *Proceedings of the 2012 AAS/AIAA Space Flight Mechanics Meeting*, Vol. 143 of *Advances in the Astronautical Sciences*, Charleston, South Carolina, January 31-February 4 2012, pp. 337–348, AAS 12-125.