AAS 12-125

NOVEL MULTIPLICATIVE UNSCENTED KALMAN FILTER FOR ATTITUDE ESTIMATION

Renato Zanetti,* **Kyle J. DeMars**,† and **Daniele Mortari**‡

A novel spacecraft attitude estimation algorithm is presented. The new algorithm utilizes unit vector measurements and is based on the unscented Kalman filter (UKF). The UKF, like the extended Kalman filter, employs a linear update in which an additive residual is formed. The residual is given by the difference between the measurement and its mean. This work utilizes a multiplicative residual in which the measurement and the mean are multiplied together using the vector cross product. Because of the nature of the problem, a multiplicative residual combined with a multiplicative update is a more natural solution.

INTRODUCTION

The Kalman filter (KF) [1, 2] and its nonlinear extension, 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]. Many approaches to enforce the unit-norm constraint of the quaternion-of-rotation in the Kalman filter exist, 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. While 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 done utilizing unit vector measurements. While direction measurements from attitude sensors are often provided as bearing angles, a unit vector can be 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 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

*Senior Member of the Technical Staff, Vehicle Dynamics and Control, The Charles Stark Draper Laboratory, 17629 El Camino Real, Suite 470, Houston, Texas, 77058. E-mail: rzanetti@draper.com

†National Research Council (NRC) Postdoctoral Fellow

‡Professor, 746C H.R. Bright Building, Aerospace Engineering, Texas A&M University, College Station, TX 77843-3141, Tel.: (979) 845-0734, Fax: (979) 845-6051. E-mail: mortari@tamu.edu

vector model because it more closely represents the actual errors of the measurements. This work presents a novel attitude UKF that utilizes both a multiplicative measurement model and a multiplicative residual [16]. 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 [17] which provides the estimate with the minimum attitude error.

THE UNSCENTED KALMAN FILTER

The EKF requires linearization of both the dynamics and measurement equations. However, the Kalman filtering paradigm 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 [18]. Under these three requirements, it can then be shown that for a dynamical system of the form

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

where $\mathbf{x}_{k-1}$ is the state at time t_{k-1}, $\mathbf{w}_{k-1}$ is the process noise input at time t_{k-1}, $\mathbf{x}$ is the state at time t_k, the state estimate evolves as

$$\hat{\mathbf{x}}^- = \mathrm{E}\left\{\mathbf{f}(\mathbf{x}_{k-1}, \mathbf{w}_{k-1})\right\} \,,$$

such that the second central moment can be computed via

$$\mathbf{P}^- = \mathrm{E}\left\{(\mathbf{x} - \hat{\mathbf{x}}^-)(\mathbf{x} - \hat{\mathbf{x}}^-)^{\mathrm{T}}\right\} \,.$$

Then, when measurement updates are considered, we have the predicted measurement as

$$\hat{\mathbf{y}}^- = \mathrm{E}\left\{\mathbf{h}(\mathbf{x}, \mathbf{v})\right\} \,,$$

where $\mathbf{v}$ represents measurement noise and $\mathbf{h}$ represents the nonlinear measurement function, which allows the state estimate and the covariance to be updated (assuming a linear scheme for the update), yielding

$$\begin{aligned} \hat{\mathbf{x}}^+ &= \hat{\mathbf{x}}^- + \mathbf{K}(\mathbf{y} - \hat{\mathbf{y}}^-) \\ \mathbf{P}^+ &= \mathbf{P}^- - \mathbf{K}\mathbf{P}_y\mathbf{K}^{\mathrm{T}} \,, \end{aligned}$$

where the Kalman gain is given by

$$\begin{aligned} \mathbf{K} &= \mathrm{E}\left\{(\mathbf{x} - \hat{\mathbf{x}}^-)(\mathbf{y} - \hat{\mathbf{y}}^-)^{\mathrm{T}}\right\} \mathrm{E}\left\{(\mathbf{y} - \hat{\mathbf{y}}^-)(\mathbf{y} - \hat{\mathbf{y}}^-)^{\mathrm{T}}\right\}^{-1} \\ &= \mathbf{P}_{xy}\mathbf{P}_y^{-1} \,. \end{aligned}$$

Note that if the linearization procedures described in the development of the EKF are implemented in the above relationships, then we recover the EKF. However, it is not necessary to consider linearizations. One such method which forgoes linearization in favor of a more accurate computation is the UKF.

The Unscented Transform

Consider a nonlinear function of the form

$$\mathbf{z} = \mathbf{g}(\mathbf{x}) \,,$$

where $\mathbf{x}$ is described by a known mean and covariance, respectively $\mathbf{m}_x$ and $\mathbf{P}_x$. The unscented transform (UT) seeks to approximate the transformation of the mean and covariance of the output, $\mathbf{z}$, which are denoted by $\mathbf{m}_z$ and $\mathbf{P}_z$.

Whereas linearization methods utilize a first-order Taylor series expansion to approximate the transformation of the mean and covariance through a nonlinear function, the UT approaches the problem under the philosophy that it is easier to approximate a probability distribution than it is to approximate an arbitrary nonlinear function [19]. To approximate the probability distribution, the UT considers a set of deterministically chosen weighted sigma-points which are selected such that $\mathbf{m}_x$ and $\mathbf{P}_x$ are exactly captured by the sigma-points. The sigma-points are then applied as inputs to the nonlinear function to yield nonlinearly transformed sigma-points, which can then be used to approximate a nonlinear transformation of the output mean and covariance, $\mathbf{m}_z$ and $\mathbf{P}_z$.

Let the set of sigma-points be denoted by the K values of $\mathcal{X}_i$ and the associated weights by w_i where $i \in \{1, \ldots, K\}$ and $\sum_{i=1}^{K} w_i = 1$. Then, the set of transformed sigma-points are given by

$$\mathcal{Z}_i = \mathbf{g}(\mathcal{X}_i) \quad \forall \;\, i \in \{1, \ldots, K\} \,,$$

which are then used to compute the transformed mean and covariance as

$$\begin{aligned}
\mathbf{m}_z &= \sum_{i=1}^{K} w_i \mathcal{Z}_i \\
\mathbf{P}_z &= \sum_{i=1}^{K} w_i (\mathcal{Z}_i - \mathbf{m}_z)(\mathcal{Z}_i - \mathbf{m}_z)^{\mathrm{T}} \,.
\end{aligned}$$

Additionally, the cross-covariance between the input and the output can be computed, if desired, as

$$\mathbf{P}_{xz} = \sum_{i=1}^{K} w_i (\mathcal{X}_i - \mathbf{m}_x)(\mathcal{Z}_i - \mathbf{m}_z)^{\mathrm{T}} \,.$$

Any selection of sigma-points that exactly describes the input mean and covariance guarantees that the transformed mean and covariance is correctly calculated to second order [12].

Many possibilities exist for the selection of the sigma-points and the associated weights, such as the simplex set, symmetric set, symmetric extended set, among others [12]. We restrict our attention to the symmetric extended set, which is given by the set of $K = 2n + 1$ points chosen as

$$\begin{aligned}
\mathcal{X}_0 &= \mathbf{m}_x \\
\mathcal{X}_i &= \mathbf{m}_x + \sqrt{n+\kappa}\,\mathbf{s}_{x,i} \\
\mathcal{X}_{i+n} &= \mathbf{m}_x - \sqrt{n+\kappa}\,\mathbf{s}_{x,i} \,,
\end{aligned}$$

with associated weights

$$\begin{aligned}
w_0 &= \kappa/(n+\kappa) \\
w_i &= 1/2(n+\kappa) \\
w_{i+n} &= 1/2(n+\kappa) \,,
\end{aligned}$$

for $i \in \{1, \ldots, n\}$, where n is the dimension of the input $\mathbf{x}$, $\mathbf{S}_x$ is a square-root factor of $\mathbf{P}_x$ such that $\mathbf{P}_x = \mathbf{S}_x \mathbf{S}_x^{\mathrm{T}}$, $\mathbf{s}_{x,i}$ is the i^{th} column of $\mathbf{S}_x$, and κ is a tuning parameter of the UT. It is easily verified that the symmetric extended set of sigma-points matches the mean and covariance of $\mathbf{x}$, that is

$$\begin{aligned}\mathbf{m}_x &= \sum_{i=1}^{K} w_i \mathcal{X}_i \\ \mathbf{P}_x &= \sum_{i=1}^{K} w_i (\mathcal{X}_i - \mathbf{m}_x)(\mathcal{X}_i - \mathbf{m}_x)^{\mathrm{T}} \,.\end{aligned}$$

Propagation

The purpose of the propagation step is to compute the mean and covariance at time t_k (denoted $\hat{\mathbf{x}}_{k-1}$ and $\mathbf{P}^-$, respectively) given the mean and covariance at time t_{k-1} (denoted $\mathbf{x}_{k-1}^-$ and $\mathbf{P}_{k-1}^-$, respectively). In order to apply the UT to the forward propagation of the mean and covariance, the first step is to determine the square-root factor of the covariance matrix, that is to find $\mathbf{S}_{k-1}$ such that

$$\mathbf{S}_{k-1}\mathbf{S}_{k-1}^{\mathrm{T}} = \begin{bmatrix} \mathbf{P}_{k-1} & \mathbf{O} \\ \mathbf{O} & \mathbf{Q} \end{bmatrix} ,$$

where $\mathbf{Q}$ is the process noise covariance matrix. The determination of $\mathbf{S}_{k-1}$ can be readily accomplished via a Cholesky factorization. Once the square-root factor is determined, the symmetric extended sigma-point selection scheme previously discussed is used to generate $2n+1$ sigma points, $\mathcal{Z}_{i,k-1}$, and their associated weights w_i, with the mean given by

$$\mathbf{m}_{k-1} = \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \mathbf{0} \end{bmatrix} ,$$

which reflects the zero-mean nature of the process noise. Note that n is now the dimension of the combination of the state dimension and the process noise dimension, i.e. $n = n_x + n_q$, where n_x is the state dimension and n_q is the process noise dimension. Let each of the sigma points be partitioned as

$$\mathcal{Z}_{i,k-1} = \begin{bmatrix} \mathcal{X}_{i,k-1} \\ \mathcal{W}_{i,k-1} \end{bmatrix} ,$$

where $\mathcal{Z}_{i,k-1}$ is n-dimensional, $\mathcal{X}_{i,k-1}$ is n_x-dimensional, and $\mathcal{W}_{i,k-1}$ is n_q-dimensional. Then, the propagated sigma points are obtained via application of the dynamical systems as

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

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

$$\begin{aligned}\hat{\mathbf{x}}^- &= \sum_{i=1}^{K} w_i \mathcal{X}_i \\ \mathbf{P}^- &= \sum_{i=1}^{K} w_i (\mathcal{X}_i - \hat{\mathbf{x}}^-)(\mathcal{X}_i - \hat{\mathbf{x}}^-)^{\mathrm{T}} \,.\end{aligned}$$

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 determine the square-root factor of the covariance matrix, that is to find $\mathbf{S}$ such that

$$\mathbf{S}\mathbf{S}^{\mathrm{T}} = \begin{bmatrix} \mathbf{P}^{-} & \mathbf{O} \\ \mathbf{O} & \mathbf{R} \end{bmatrix} ,$$

where $\mathbf{R}$ is now the measurement noise covariance matrix. Once the square-root factor is determined, the symmetric extended sigma-point selection scheme previously discussed is used to generate $2n+1$ sigma points, $\mathcal{X}_i$, and their associated weights w_i, with the mean given by

$$\mathbf{m} = \begin{bmatrix} \hat{\mathbf{x}}^{-} \\ \mathbf{0} \end{bmatrix} ,$$

which reflects the zero-mean nature of the measurement noise. Note that n is now the dimension of the combination of the state dimension and the measurement noise dimension, i.e. $n = n_x + n_v$, where n_x is the state dimension and n_v is the measurement noise dimension. Using the sigma-points at time t_k, the measurement-transformed sigma-points are given by evaluating the nonlinear measurement function $\mathbf{h}(\cdot)$ at each sigma-point, yielding the $K = 2n+1$ transformed sigma-points as

$$\mathcal{Y}_i = \mathbf{h}(\mathcal{X}_i) .$$

Note here that the first n_x elements of $\mathcal{X}_i$ represent the state and the last n_v elements of $\mathcal{X}_i$ represent the measurement noise. The expected value of the measurement, the measurement covariance, and the cross-covariance are found in terms of the transformed sigma-points as

$$\begin{aligned} \hat{\mathbf{y}}^{-} &= \sum_{i=1}^{K} w_i \mathcal{Y}_i \\ \mathbf{P}_y &= \sum_{i=1}^{K} w_i (\mathcal{Y}_i - \hat{\mathbf{y}}^{-})(\mathcal{Y}_i - \hat{\mathbf{y}}^{-})^{\mathrm{T}} \\ \mathbf{P}_{xy} &= \sum_{i=1}^{K} w_i (\mathcal{X}_i - \hat{\mathbf{x}}^{-})(\mathcal{Y}_i - \hat{\mathbf{y}}^{-})^{\mathrm{T}} . \end{aligned}$$

In terms of the measurement covariance and the cross-covariance between the state and the measurement, the Kalman gain is

$$\mathbf{K} = \begin{bmatrix} \mathbf{I}_{n_x \times n_x} & \mathbf{O}_{n_x \times n_v} \end{bmatrix} \mathbf{P}_{xy} \mathbf{P}_y^{-1} ,$$

and the associated updated state estimate and covariance are

$$\begin{aligned} \hat{\mathbf{x}}^{+} &= \hat{\mathbf{x}}^{-} + \mathbf{K}(\mathbf{y} - \hat{\mathbf{y}}^{-}) \\ \mathbf{P}^{+} &= \mathbf{P}^{-} - \mathbf{K}\mathbf{P}_y\mathbf{K}^{\mathrm{T}} . \end{aligned}$$

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 (scaled) covariance matrix. For states which 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 relies on a 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. 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. All of these needed modifications may be grouped together as the removal of additivity within the UKF in favor of multiplicative steps.

ATTITUDE FILTER

The proposed filter differs from that of Crassidis and Markley both in the update and propagation phases. The update differs because a multiplicative measurement model is used as well as a multiplicative residual. The measurement model is given by

$$\mathbf{y} = \mathbf{T}(\boldsymbol{\eta})\mathbf{T}_i^b\mathbf{r}\,, \tag{1}$$

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

$$\mathbf{y} = \mathbf{T}_i^b\mathbf{r} + \boldsymbol{\nu}. \tag{2}$$

The additive measurement model relies on linearization, even the large field-of-view model from Cheng *et al.* 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 $\mathbf{y}$ are of unit lengths we have that

$$\mathbf{y}^{\mathrm{T}}\mathbf{y} = 1 = \mathbf{r}^{\mathrm{T}}\mathbf{T}_b^i\mathbf{T}_i^b\mathbf{r} + 2\boldsymbol{\nu}^{\mathrm{T}}\mathbf{T}_i^b\mathbf{r} + \boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu} = 1 + 2\boldsymbol{\nu}^{\mathrm{T}}\mathbf{T}_i^b\mathbf{r} + \boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\nu}, \tag{3}$$

taking expected values and realizing that $\mathbf{r}$ is deterministic

$$2\mathbf{r}^{\mathrm{T}}\mathbf{T}_b^i\,\mathrm{E}\left\{\boldsymbol{\nu}\right\} = -\,\mathrm{trace}\,\mathrm{E}\left\{\boldsymbol{\nu}\boldsymbol{\nu}^{\mathrm{T}}\right\}. \tag{4}$$

The above equation implies that the measurement noise either has zero mean and zero covariance or has non-zero mean. Using the multiplicative measurement model in the UKF allows us to obtain an unbiased estimator. Furthermore one of the strengths of the UKF is that it avoids linearization, it therefore seems consistent to utilize a measurement model that also does not rely on linearization.

The state vector is given by

$$\mathbf{x} = \begin{bmatrix} \delta\boldsymbol{\varrho} \\ \mathbf{b} \end{bmatrix}\,, \tag{5}$$

where $\mathbf{b}$ is the gyro bias and $\delta\boldsymbol{\varrho}$ is the rotation between the estimated quaternion and the true quaternion, which is defined such that

$$\bar{\mathbf{q}}(\delta\boldsymbol{\varrho}) = \bar{\mathbf{q}}_i^b \otimes (\hat{\bar{\mathbf{q}}}_i^b)^*\,, \tag{6}$$

where the asterisk represents the quaternion conjugate and the quaternion multiplication $\otimes$ composes quaternions in the same order as attitude matrices. Many three-dimensional representations can be used to express $\delta\boldsymbol{\varrho}$, for example the rotation vector, the Gibbs vector, the modified Rodrigues parameters, etc. For this work the three-dimensional representation chosen is twice the vector component of the quaternion (where the scalar component must be non-negative). Notice that $\delta\hat{\boldsymbol{\varrho}}^- = \mathbf{0}$ by definition. The gyro bias can be replaced by the angular velocity for gyro-less systems; this change does not affect the update phase while it does affect the propagation phase.

The state update is given by

$$\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{K}(\mathbf{y} \times \hat{\mathbf{y}})\,, \tag{7}$$

where $\mathbf{y}$ is the actual measurement, which is one realization of the random vector $\mathbf{Y}$. The unit vector $\hat{\mathbf{y}}$ is the normalized mean of $\mathbf{Y}$. Notice that we could rewrite the above update as

$$\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{K}(\mathbf{z} - \hat{\mathbf{z}})\,, \tag{8}$$

where the auxiliary variable $\mathbf{z}$ is defined as $\mathbf{z} = \mathbf{y} \times \hat{\mathbf{y}}$ and has zero mean, $\hat{\mathbf{z}} = \mathbf{0}$, since $\mathrm{E}\{\mathbf{Y}\}$ and $\hat{\mathbf{y}}$ are parallel. The new update can be 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 following augmented covariance

$$\boldsymbol{\Sigma}_{XX}^U = \begin{bmatrix} \mathbf{P}^- & \mathbf{O} \\ \mathbf{O} & \mathbf{R} \end{bmatrix}\,, \tag{9}$$

where $\mathbf{P}^-$ is the *a priori* estimation error covariance and $\mathbf{R}$ is the measurement noise covariance. Because of the multiplicative measurement model of Eq. (1), $\mathbf{R}$ is chosen full-rank without any approximation. Additive measurement models inevitably possess a rank-deficient measurement error covariance. With the $n \times n$ matrix $\boldsymbol{\Sigma}_{XX}^U$ defined above, the $2n+1$ sigma points are given by

$$\mathcal{X}_0 = \begin{bmatrix} \hat{\mathbf{x}}^- \\ \mathbf{0} \end{bmatrix} \tag{10}$$

$$\mathcal{X}_i = \begin{bmatrix} \hat{\mathbf{x}}^- \\ \mathbf{0} \end{bmatrix} + \sqrt{(n+\kappa)\boldsymbol{\Sigma}_{XX}^U(i)} \tag{11}$$

$$\mathcal{X}_{i+n} = \begin{bmatrix} \hat{\mathbf{x}}^- \\ \mathbf{0} \end{bmatrix} - \sqrt{(n+\kappa)\boldsymbol{\Sigma}_{XX}^U(i)}\,, \tag{12}$$

where $\sqrt{\mathbf{A}(i)}$ is the i-th column of the matrix square root of $\mathbf{A}$. Along with the sigma points, weights are chosen as

$$w_0 = \kappa/(n+\kappa) \tag{13}$$

$$w_i = 1/2(n+\kappa)\,, \tag{14}$$

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

$$\mathcal{Y}_i = \mathbf{h}(\mathcal{X}_i, \mathbf{r}, \hat{\bar{\mathbf{q}}}_i^b)\,, \tag{15}$$

where

$$\mathbf{h}(\mathbf{X}, \mathbf{r}, \hat{\bar{\mathbf{q}}}_i^b) = \mathbf{T}(\boldsymbol{\eta})\mathbf{T}(\delta\boldsymbol{\varrho})\mathbf{T}(\hat{\bar{\mathbf{q}}}_i^b)\mathbf{r}\,, \tag{16}$$

where $\boldsymbol{\eta}$ and $\delta\boldsymbol{\varrho}$ are part of the augmented state vector (i.e. the input sigma points).

The mean and covariance of the transformed variables are found via

$$\hat{\mathbf{x}} = \sum_{i=0}^{2n} w_i \mathcal{X}_i \tag{17}$$

$$\hat{\mathbf{y}} = c \sum_{i=0}^{2n} w_i \mathcal{Y}_i \tag{18}$$

$$\boldsymbol{\Sigma}_{ZZ} = \sum_{i=0}^{2n} w_i (\mathcal{Y}_i \times \hat{\mathbf{y}})(\mathcal{Y}_i \times \hat{\mathbf{y}})^{\mathrm{T}} \tag{19}$$

$$\boldsymbol{\Sigma}_{XZ} = \sum_{i=0}^{2n} w_i (\mathcal{X}_i - \hat{\mathbf{x}})(\mathcal{Y}_i \times \hat{\mathbf{y}})^{\mathrm{T}}\,, \tag{20}$$

where c is a normalizing factor that ensures $\hat{\mathbf{y}}$ is a unit vector. The updated state and covariance are obtained from Eq. (7) and

$$\mathbf{K} = \begin{bmatrix} \mathbf{I}_{n\times n} & \mathbf{O}_{n\times 3} \end{bmatrix} \boldsymbol{\Sigma}_{XZ} \boldsymbol{\Sigma}_{ZZ}^{\dagger} \tag{21}$$

$$\mathbf{P}^{+} = \mathbf{P}^{-} - \mathbf{K}\boldsymbol{\Sigma}_{ZZ}\mathbf{K}^{\mathrm{T}}\,. \tag{22}$$

The superscript † represents the Moore-Penrose pseudoinverse which Catlin shows provides the minimum variance estimate [20, page 160]. When a matrix is non-singular its inverse and pseudoinverse coincide. After some manipulations, Eq. (19) can be rewritten as

$$\boldsymbol{\Sigma}_{ZZ} = [\hat{\mathbf{y}}\times] \left(\sum_{i=0}^{2n} W_i \mathcal{Y}_i \mathcal{Y}_i^{\mathrm{T}} \right) [\hat{\mathbf{y}}\times]^{\mathrm{T}}\,. \tag{23}$$

The pseudoinverse of $\boldsymbol{\Sigma}_{ZZ}$ is given by

$$\boldsymbol{\Sigma}_{ZZ}^{\dagger} = [\hat{\mathbf{y}}\times] \left(\sum_{i=0}^{2n} W_i \mathcal{Y}_i \mathcal{Y}_i^{\mathrm{T}} \right)^{-1} [\hat{\mathbf{y}}\times]^{\mathrm{T}}\,. \tag{24}$$

For small errors the matrix in parentheses can potentially be ill-conditioned and an alternative procedure can be used to calculate the pseudoinverse. Following [14] we notice that

$$\left(\boldsymbol{\Sigma}_{ZZ} + \hat{\mathbf{y}}\hat{\mathbf{y}}^{\mathrm{T}}\right)^{-1} = \boldsymbol{\Sigma}_{ZZ}^{\dagger} + \hat{\mathbf{y}}\hat{\mathbf{y}}^{\mathrm{T}}\,; \tag{25}$$

hence, the pseudoinverse is given by

$$\boldsymbol{\Sigma}_{ZZ}^{\dagger} = \left(\boldsymbol{\Sigma}_{ZZ} + \hat{\mathbf{y}}\hat{\mathbf{y}}^{\mathrm{T}}\right)^{-1} - \hat{\mathbf{y}}\hat{\mathbf{y}}^{\mathrm{T}}\,. \tag{26}$$

Eq. (26) is readily provable from the definition of pseudoinverse that must satisfy

1. $\mathbf{A}\mathbf{A}^{\dagger}\mathbf{A} = \mathbf{A}$

2. $\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$

3. $(\mathbf{A}\mathbf{A}^\dagger)^{\mathrm{T}} = \mathbf{A}\mathbf{A}^\dagger$

4. $(\mathbf{A}^\dagger\mathbf{A})^{\mathrm{T}} = \mathbf{A}^\dagger\mathbf{A}$.

Finally the quaternion is updated as

$$\hat{\bar{\mathbf{q}}}_i^b \leftarrow \bar{\mathbf{q}}(\delta\hat{\boldsymbol{\varrho}}^+) \otimes \hat{\bar{\mathbf{q}}}_i^b \,. \tag{27}$$

In the propagation we obtain a set of propagated quaternions $\hat{\bar{\mathbf{q}}}_i^b(i)$, $i = 0, 1, \ldots, 2N$ following the same procedure of [9]; however, we follow a different approach to obtain the propagated quaternion estimate. The desired estimate is the minimum mean-square error (MMSE) estimate. For a discrete random vector $\mathbf{X}$ with probability mass function p_i, the MMSE estimate $\hat{\mathbf{x}}$ minimizes

$$\hat{\mathbf{x}} = \min_{\hat{\mathbf{x}}} \sum_i p_i \|\mathbf{x}_i - \hat{\mathbf{x}}\|^2 \,, \tag{28}$$

since the estimation error is given by $\mathbf{x}_i - \hat{\mathbf{x}}$. The solution of Eq. (28) is the mean of the random vector

$$\hat{\mathbf{x}} = \sum_i p_i \mathbf{x}_i \,. \tag{29}$$

In this work the attitude estimation error is defined as twice the vector part of the error quaternion. We therefore follow [17] and we obtain the propagated quaternion estimate by minimizing

$$\sum_{i=0}^{2N} w_i \|\delta\hat{\boldsymbol{\varrho}}_i^-\|^2 \,, \tag{30}$$

where

$$\bar{\mathbf{q}}(\delta\hat{\boldsymbol{\varrho}}_i^-) = \hat{\bar{\mathbf{q}}}_i^b(i) \otimes (\hat{\bar{\mathbf{q}}}_i^b)^* \,, \tag{31}$$

where $\hat{\bar{\mathbf{q}}}_i^b$ is the propagated estimate we are solving for and is given by the unit eigenvector corresponding to the maximum eigenvalue of

$$\mathbf{M} = 4 \sum_{i=0}^{2N} \left(w_i \hat{\bar{\mathbf{q}}}_i^b(i) \, \hat{\bar{\mathbf{q}}}_i^b(i)^{\mathrm{T}} \right) - \mathbf{I}_{4\times 4} \,.$$

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

$$\begin{aligned}
\dot{\bar{\mathbf{q}}}_i^b &= \frac{1}{2} \bar{\boldsymbol{\omega}}_{b/i}^b \otimes \bar{\mathbf{q}}_i^b \\
\dot{\boldsymbol{\omega}}_{b/i}^b &= \mathbf{J}^{-1} \left(\sum \mathbf{m}^b - \boldsymbol{\omega}_{b/i}^b \times \mathbf{J}\boldsymbol{\omega}_{b/i}^b \right) \,,
\end{aligned}$$

where $\bar{\boldsymbol{\omega}}^b_{b/i}$ is the pure quaternion formed from the angular velocity vector $\boldsymbol{\omega}^b_{b/i}$. and $\otimes$ represents the quaternion multiplication operation, defined such that the quaternions are multiplied in the same order as the equivalent rotation matrices would be. Furthermore, $\mathbf{J}$ is the moment of inertia of the spacecraft and $\sum \mathbf{m}^b$ represents the summation of all active moments in the body frame. The active moments are assumed to be zero in this work, and the moment of inertia is taken to be a diagonal matrix with elements described along with the overall satellite geometry in Table 2. We assume that the initial attitude has a mean orientation given by the identity quaternion; that is, the mean quaternion represents a body frame that is exactly aligned to the inertial frame. Additionally, we take the initial mean angular velocity to be zero. True values are generated by sampling a Gaussian error distribution with a standard deviation of 10° in attitude and 0.1 rev/day in angular velocity. The above equations of motion are then 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 and then subjected to zero-mean white-noise sequences with standard deviations specified in Table 3.

The new attitude filter is then applied with a starting estimated quaternion equal to the identity quaternion and the estimated bias equal to zero. The initial attitude uncertainty and bias uncertainty which describe the elements of the initial covariance matrix are take to be 10° and 0.1°, 1σ, respectively. The UKF parameter is set to $\kappa = 3 - n$ and the filter's performance is shown in Figures 1 and 2. The blue line shows the estimation error while the red lines show the predicted 3σ error standard deviation.

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

CONCLUSIONS

This work presents a novel unit-vector quaternion unscented filter with multiplicative residual and multiplicative measurement error model. The aim of the work is to develop an algorithm that does not rely on linearization or small angles assumptions. Previous attitude UKF 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 in three-dimensional attitude deviations and simply averaged together. Such an average is

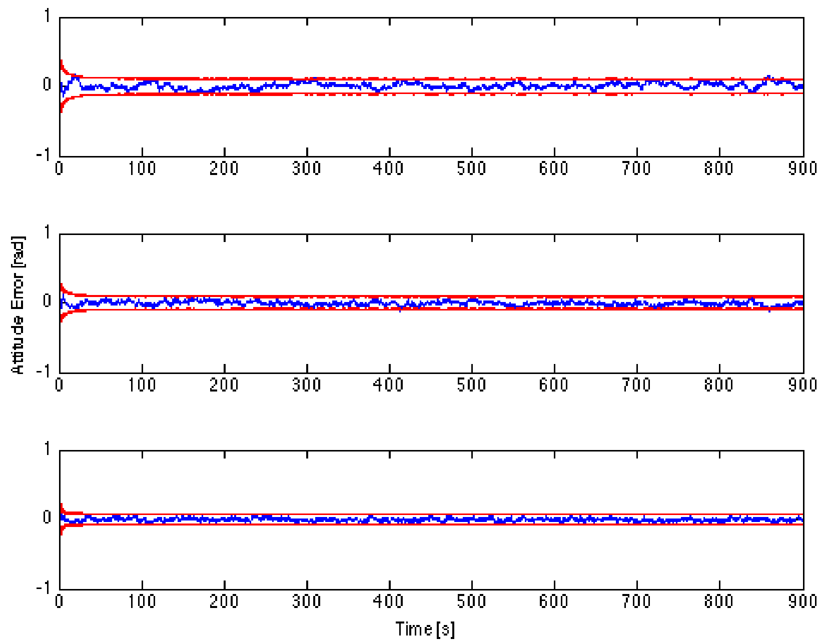

Figure 1. Attitude estimation error

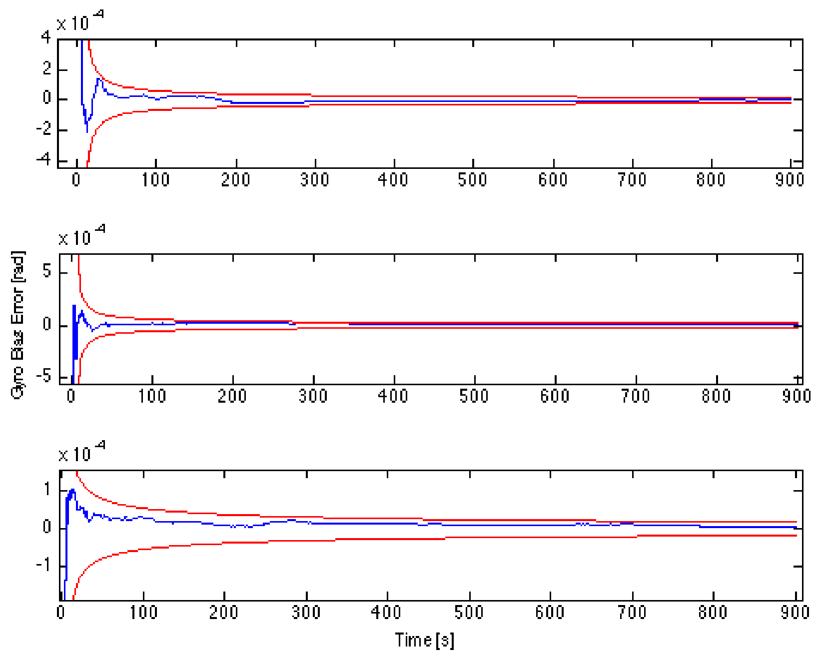

Figure 2. Gyro bias estimation error per IMU sample

Table 2. Satellite Geometry

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} Inertia	8100	kg m^2
I_{yy} Inertia	8100	kg m^2
I_{zz} Inertia	4500	kg m^2

Table 3. Sensors Specifications

Type	1σ Error	Units
Gyro Noise	0.1	deg/$\sqrt{\text{s}}$
Gyro Bias	0.1	deg/s
Sun Sensor Error	2	deg
Earth Sensor Error	5	deg

only valid for small angles, while the proposed algorithm averages the quaternions taking in full consideration the inherit nonlinear nature of the rotation group.

REFERENCES

[1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Journal of Basic Engineering*, Vol. 82, March 1960, pp. 35–45.

[2] R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction," *Journal of Basic Engineering*, Vol. 83, March 1961, pp. 95–108.

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

[4] F. L. Markley, "Attitude Estimation or Quaternion Estimation?," *The Journal of the Astronautical Sciences*, Vol. 52, January–June 2004, pp. 221–238.

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

[6] I. Y. Bar-Itzhack and Y. Oshman, "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transaction on Aerospace and Electronic Systems*, Vol. 21, January 1985, pp. 128–135.

[7] D. Simon and T. L. Chia, "Kalman Filtering with State Equality Constraints," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 38, January 2002, pp. 128–136.

[8] R. Zanetti, M. Majji, R. H. Bishop, and D. Mortari, "Norm-Constrained Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 32, September–October 2009.

[9] J. L. Crassidis and F. L. Markley, "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance Control and Dynamics*, Vol. 26, July - August 2003, pp. 536–541.

[10] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, Vol. 45, March 2000, pp. 477–482.

[11] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, "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.

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

[13] M. D. Shuster, "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences*, Vol. 38, July–September 1990, pp. 377–393.

[14] Y. Cheng, J. L. Crassidis, and F. L. Markley, “Attitude Estimation for Large Field-of-View Sensors,” *The Journal of the Astronautical Sciences*, Vol. 54, July–December 2006, pp. 433–448.

[15] D. Mortari and M. Majji, “Multiplicative Measurement Model,” AAS *Journal of the Astronautical Sciences*, Vol. 57, Nos. 1&2, January-June 2009, pp. 47-60.

[16] R. Zanetti, “A Multiplicative Residual Approach to Attitude Kalman Filtering with Unit-Vector Measurements,” *The Journal of the Astronautical Sciences*, Vol. 57, October–December 2009, pp. 793–801.

[17] F. L. Markley, Y. Cheng, J. L. Crassidis, and Y. Oshman, “Averaging Quaternions,” *Journal of Guidance, Control, and Dynamics*, Vol. 30, July–August 2007, pp. 1193–1197.

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

[19] J. Uhlmann, *Simultaneous Map Building and Localization for Real Time Applications*, PhD thesis, University of Oxford, 1994.

[20] D. E. Catlin, *Estimation, Control, and the Discrete Kalman Filter*. No. 71 in Applied Mathematical Sciences, New York: Springer-Verlag, 1989.