The Journal of the Astronautical Sciences, Vol. 57, Nos. 1 & 2, January–June 2009, pp. 93–111

Additive Divided Difference Filtering for Real-Time Spacecraft Attitude Estimation Using Modified Rodrigues Parameters[1]

Deok-Jin Lee[2] and Kyle T. Alfriend[3]

Abstract

In this paper, a real-time attitude estimation algorithm is derived by using an additive divided difference filter as an efficient alternative to the extended Kalman filter. To make the attitude filtering algorithm suitable for real-time applications and to minimize the computational load, a square-root sigma point attitude filter is designed by integrating the divided difference filter with the additive noise concept using the modified Rodrigues attitude parameters. The new attitude filter provides numerically stable and accurate estimates of the state and covariance, but the computational workload of the new estimator is almost identical to the computational complexity of the extended Kalman attitude filter. For performance evaluation the new sigma point attitude filter is compared with the unscented attitude filter and the extended Kalman filter. The sensor measurements include a three-axis magnetometer and rate-gyros. Simulation results indicate that the proposed additive divided difference attitude filter shows faster convergence with accurate and reliable estimation.

Introduction

Real-time attitude estimation plays a major role in the navigation and control of autonomous systems, such as unmanned aerial vehicles, spacecraft, and autonomous underwater vehicles. For autonomous systems estimation algorithms must carry out the determination of the attitude by fusing information from onboard sensors, such as rate gyros, three-axis magnetometers, GPS sensors and star sensors, etc. However, attitude determination in the presence of system uncertainty and a high degree

[1]Presented at the F. Landis Markley Astronautics Symposium, Cambridge, Maryland, June 29–July 2, 2008.
[2]Adjunct Research Professor/NRC Research Associate, Center for Autonomous Vehicle Research, Department of Mechanical & Astronautical Engineering, Naval Postgraduate School, Monterey, 93943-5100. E-mail: djlee@nps.edu or apollo17@gmail.com.
[3]TEES Distinguished Research Chair Professor, Center for Mechanics & Control, Department of Aerospace Engineering, Texas A&M University, College Station, TX, 77843-3131, Fellow, AAS.

of nonlinearity leads to degraded estimation accuracy with slower convergence [1]. Therefore, robust and reliable filtering algorithms with lower computational complexity are essential in many real-time fault-tolerant estimation applications [2].

The extended Kalman filter (EKF) has been widely implemented for nonlinear estimation problems [2, 3]. The filter involves the recursive estimation of the first two moments of the state distribution under a Gaussian assumption based on the first-order linearization of the nonlinear dynamic and measurement equations. The series approximation can, however, introduce large errors in the true mean and covariance of the posterior distribution, which can lead to biased sub-optimal estimation, and even divergence of the filter [3].

To compensate for the drawbacks of the EKF and improve estimation performance, several alternative nonlinear filtering algorithms, such as particle filters (PFs) [4–7], the unscented Kalman filter (UKF) [8], the central difference filter (CDF) [9], and the divided difference filter (DDF) [10], have been proposed over the past decade. The particle filter, also known as the bootstrap filter [4], provides a tractable solution to nonlinear and/or non-Gaussian systems as a general new class of estimation algorithms. Particle filtering for spacecraft attitude estimation has been applied in [11] and [12]. Although it compensates for the difficulties of the EKF, it faces the practical issue that the number of random samples required to achieve an accurate and reliable estimation solution increases dramatically in a three or higher dimensional estimation problem [6, 7]. This drawback can be mitigated by minimizing the number of random samples, or by reducing the dimension of the estimated state vector known as the Rao-Blackwellization approach [7]. In reference [13] the dimensionality problem was alleviated by using an efficient initialization method along with a modified importance weight sampling approach. On the other hand, the UKF, CDF, and DDF, called sigma point Kalman filters (SPKFs) [14], or sigma point filters (SPFs) [15], work on the principle that a set of deterministically sampled sigma points can be used to parameterize the mean and covariance of the conditional probability density of the state given the observations, and the posterior mean and covariance are propagated through the true nonlinear function without the linearization step. One of the differences between the SPFs and the PFs is that the SPFs utilize a set of deterministically sampled sigma points from a Gaussian distribution to construct the first two moments, and follow the Kalman predictor-corrector structure, whereas the PFs use recursively a set of random samples of the state variables from an assumed probability distribution to solve the on-line Bayesian filtering problem [4]. The deterministically obtained samples of the SPFs avoid random sampling errors caused by the Monte-Carlo sampling methods, thus the number of sample points required is reduced dramatically compared to the PFs. Recently, the SPFs have been used in a wide field of applications such as integrated navigation [14], satellite orbit estimation [15], and spacecraft attitude estimation [16]. There have also been alternative extensions of the standard SPFs to increase their performance. For example, the suboptimal adaptive UKF and adaptive DDF were proposed in order to take into account unknown system and measurement noise uncertainties [17]. For a numerically stabilized estimation algorithm, the square-root forms of the UKF and the CDF [18] were proposed by applying factorization methods in the terms of the augmented state vector and covariance matrix. The square-root filters increase the numerical stability properties of the filtering algorithm by producing a positive definite covariance matrix at the added computational cost.

The purpose of this paper is to propose an efficient attitude estimation algorithm that not only provides reliable estimation accuracy, but also alleviates the computational load with an enhanced computational stability. In this paper, two tactical approaches are introduced to bring these objectives into reality. First, a new attitude estimation algorithm based on the divided difference filter which is reformulated in terms of sigma point vectors with a square-root structure is derived to enhance the estimation performance. Second, in order to alleviate the computational load of the standard DDF an additive process and measurement noise concept is integrated, which leads to an additive DDF algorithm. Therefore, the proposed new sigma point attitude filter, called the "additive divided difference filter" (ADDF), substantially reduces the number of sigma points required from the $2n_a + 1$ symmetric sigma points with a n_a-dimensional augmented state vector including process and measurement noise vectors to the sigma points with an n-dimensional state vector without augmenting the noise vectors. This leads to a computationally efficient real-time estimation algorithm.

There are several parameterization approaches for representing the attitude, such as Euler angles, quaternions, and modified Rodrigues parameters [19]. Appealing advantages of using quaternions are that they are singularity-free, and the kinematics equation is bilinear. However, since the quaternion parameterization uses four components to represent the attitude motion, the quaternion components are non-minimal [20]. This leads to the unit norm constraint that the quaternion must satisfy, and this constraint restricts direct application of the standard extended Kalman filter to the quaternion-based attitude estimation because the linear measurement updates in the EKF structure violate the unit norm condition. One common approach for overcoming this difficulty is to use a multiplicative error quaternion, where the four quaternion components can be replaced by a three-component incremental error vector [20]. On the other hand, minimal three-dimensional parameterization of attitude motion is useful for many control applications because the control algorithms require estimates of the attitude and/or rate in real-time with minimum computational complexity. Since the modified Rodrigues parameters (MRP) allow for rotation up to 360 degrees, they are used in many attitude estimation and control applications [21, 22]. This paper also utilizes the modified Rodrigues parameters in the divided difference filtering formulation in order to develop a computational efficient attitude estimation algorithm.

The proposed additive divided difference attitude filter (ADDAF) has an advantage over the UKF in that it can lead to a more robust and stable estimation solution. As indicated in reference [16], the quaternion-based unscented attitude filter is subject to the variations of the tuning parameters related to not only the weight scale factors, but also the generalized MRP parameter. Thus, it is necessary to optimize the filter performance by tuning the weight scale factor and the attitude parameter with trial and error operations. However the new sigma point attitude filter (SPAF) based on the additive divided difference filter is robust against the variations of the weight scale factor, the interval length h.

The remainder of this paper is organized as follows: i) development of a sigma point divided difference filtering algorithm for nonlinear estimation, ii) review of attitude system models, iii) derivation of an additive divided difference attitude filtering algorithm using the MRP attitude parameters, iv) comparison of the performance of the proposed sigma point attitude filter with the UKF and the EKF, and v) discussion of the simulation results.

Sigma Point Divided Difference Filtering

In this section a sigma point divided difference filtering algorithm for discrete-time nonlinear estimation is derived by embedding scaled sigma point vectors into the second-order divided difference filter [10] in square-root forms. Consider the discrete-time nonlinear equations

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

$$\tilde{\mathbf{y}}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k, k) \tag{2}$$

where $\mathbf{x}_k \in \mathcal{R}^{n_x}$ is the state vector, $\tilde{\mathbf{y}}_k \in \mathcal{R}^{n_y}$ is the observation vector, $\mathbf{w}_k \in \mathcal{R}^{n_w}$ is the state noise vector and $\mathbf{v}_k \in \mathcal{R}^{n_v}$ is the measurement noise vector. It is assumed that the noise vectors are uncorrelated white Gaussian processes with $\mathbf{w}_k \sim \mathcal{N}(\bar{\mathbf{w}}_k, Q_k)$ and $\mathbf{v}_k \sim \mathcal{N}(\bar{\mathbf{v}}_k, R_k)$, where $Q_k \in \mathcal{R}^{n_x \times n_x}$ and $R_k \in \mathcal{R}^{n_v \times n_v}$ are the process and measurement noise covariance matrices and $\bar{\mathbf{w}}_k \in \mathcal{R}^{n_x}$ and $\bar{\mathbf{v}}_k \in \mathcal{R}^{n_v}$ are the corresponding mean values. First, the augmented state vector $\hat{\mathbf{x}}_k^{xw} \in \mathcal{R}^{(n_x+n_w)}$ and the augmented square-root matrix $S_k^{xw} \in \mathcal{R}^{(n_x+n_w)\times(n_x+n_w)}$ are constructed by

$$\hat{\mathbf{x}}_k^{xw} = \begin{bmatrix} \hat{\mathbf{x}}_k^+ \\ \bar{\mathbf{w}}_k \end{bmatrix}, \qquad S_k^{xw} = \begin{bmatrix} S_{x,k}^+ & \mathbf{0} \\ \mathbf{0} & S_{w,k} \end{bmatrix} \tag{3}$$

where the square-root matrices $S_{x,k} \in \mathcal{R}^{n_x \times n_x}$ and $S_{w,k} \in \mathcal{R}^{n_w \times n_w}$ are obtained from the Cholesky factorization method [23], $S_{x,k}^+ (S_{x,k}^+)^{\mathrm{T}} \equiv \mathrm{chol}\,(P_k^+)$ and $S_{w,k} S_{w,k}^{\mathrm{T}} \equiv \mathrm{chol}\,(Q_k)$, and $\hat{\mathbf{x}}_k^+$ and P_k^+ denote the initial state estimate and the initial covariance of the state estimate error, respectively. Then, a set of weighted scaled symmetric sigma point vectors $\boldsymbol{\chi}_{i,k}^{xw} \equiv [(\boldsymbol{\chi}_{i,k}^{x})^{\mathrm{T}}\,(\boldsymbol{\chi}_{i,k}^{w})^{\mathrm{T}}]^{\mathrm{T}}$ is constructed by

$$\begin{aligned} \boldsymbol{\chi}_{0,k}^{xw} &= \hat{\mathbf{x}}_k^{xw} \\ \boldsymbol{\chi}_{i,k}^{xw} &= \hat{\mathbf{x}}_k^{xw} + h\mathbf{S}_{i,k}^{xw}, \quad i = 1,\ldots,n_{xw} \\ \boldsymbol{\chi}_{i,k}^{xw} &= \hat{\mathbf{x}}_k^{xw} - h\mathbf{S}_{i,k}^{xw}, \quad i = n_{xw}+1,\ldots,2n_{xw} \end{aligned} \tag{4}$$

where $n_{xw} \equiv n_x + n_w$ is the dimension of the augmented state vector, $\mathbf{S}_{i,k}^{xw}$ is the ith sigma point column vector of the compound square-root matrix $\mathbf{S}_k^{xw}$ and h is a scaling weight factor coming from the divided difference step size or interval length [10]. Then, the sigma point vectors are propagated forward in time through the nonlinear dynamic equation as

$$\boldsymbol{\chi}_{i,k+1|k}^{x} = \mathbf{f}\,(\boldsymbol{\chi}_{i,k}^{x}, \boldsymbol{\chi}_{i,k}^{w}, k) \tag{5}$$

where $\boldsymbol{\chi}_{i,k}^{x}$ is a weighted sigma point vector of the first n_x elements of the ith augmented sigma point column vector $\boldsymbol{\chi}_{i,k}^{xw}$ and $\boldsymbol{\chi}_{i,k}^{w}$ is a weighted sigma point vector of the next n_w elements of $\boldsymbol{\chi}_{i,k}^{xw}$, respectively. The predicted state vector $\hat{\mathbf{x}}_{k+1}^-$ is computed by using the propagated sigma points $\boldsymbol{\chi}_{i,k+1|k}^{x}$ and their associated weight parameters

$$\hat{\mathbf{x}}_{k+1}^- = \sum_{i=0}^{2n_{xw}} w_{x,i}^{(m)}\, \boldsymbol{\chi}_{i,k+1|k}^{x} \tag{6}$$

where the weights for the expected means are given by

$$w_{x,0}^{(m)} = \frac{h^2 - (n_x + n_w)}{h^2}, \quad w_{x,i}^{(m)} = \frac{1}{2h^2}, (i > 0) \tag{7}$$

The predicted state error covariance matrix $P_{k+1}^- \in \mathcal{R}^{n_x \times n_x}$ is computed by

$$P_{k+1}^{-} = S_{x,k+1}^{-} (S_{x,k+1}^{-})^{\mathrm{T}} \tag{8}$$

where $S_{x,k+1}^{-} \in \mathcal{R}^{n_x \times 2n_{xw}}$ is the predicted square root compound matrix obtained by

$$S_{x,k+1}^{-} = \mathcal{HT}([S_{x,k+1}^{(1)}(i) \quad S_{x,k+1}^{(2)}(i)]) \tag{9}$$

where $\mathcal{HT}(\cdot)$ denotes the Householder triangularization [23], and $S = \mathcal{HT}(A)$ decomposes a general rectangular matrix A into a square and upper triangular matrix S with the property $SS^{\mathrm{T}} = AA^{\mathrm{T}}$. Each component matrix is given by

$$\begin{aligned} S_{x,k+1}^{(1)}(i) &= w_{d,i}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} - \boldsymbol{\chi}_{i+n_{xw},k+1|k}^{x}], \quad i = 1,\ldots,n_{xw} \\ S_{x,k+1}^{(2)}(i) &= w_{d,0}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} + \boldsymbol{\chi}_{i+n_{xw},k+1|k}^{x} - 2\boldsymbol{\chi}_{0,k+1|k}^{x}], \quad i = 1,\ldots,n_{xw} \end{aligned} \tag{10}$$

where the weight factors for the covariance matrix are given by

$$w_{d,0}^{(c)} = \frac{\sqrt{h^2 - 1}}{2h^2}, \quad w_{d,i}^{(c)} = \frac{1}{2h}, (i > 0) \tag{11}$$

Note the above square-root compound matrix $S_{x,k+1}^{-}$ is reformulated in terms of sigma point vectors after manipulating the original compound matrix in reference [10], and derivations are described in the Appendix. Note also that instead of the *HT* factorization method, the *QR* decomposition [23] can be used to produce a triangular matrix C and an orthogonal matrix T, i.e., $CT = QR(A)$ and $AA^{\mathrm{T}} = CC^{\mathrm{T}}$.

For the measurement update, the augmented state $\hat{\mathbf{x}}_{k+1|k}^{xv} \in \mathcal{R}^{n_{xv}}$ and the augmented square-root matrix $S_{k+1|k}^{xv} \in \mathcal{R}^{n_{xv} \times n_{xv}}$ are reconstructed by

$$\hat{\mathbf{x}}_{k+1|k}^{xv} = \begin{bmatrix} \hat{\mathbf{x}}_{k+1}^{-} \\ \bar{\mathbf{v}}_{k+1} \end{bmatrix}, \quad S_{k+1|k}^{xv} = \begin{bmatrix} S_{x,k+1}^{-} & \mathbf{0} \\ \mathbf{0} & S_{v,k+1} \end{bmatrix} \tag{12}$$

where $S_{x,k+1}^{-} \in \mathcal{R}^{n_x \times n_x}$ is given in equation (9) and $S_{v,k+1} \in \mathcal{R}^{n_v \times n_v}$ is obtained from the Cholesky factorization of the measurement noise covariance matrix, $S_{v,k+1}\, S_{v,k+1}^{\mathrm{T}}$ chol (R_{k+1}). For the measurement update the ith column vector of the scaled symmetric sigma point matrix, $\boldsymbol{\chi}_{i,k+1|k}^{xv} \equiv [(\boldsymbol{\chi}_{i,k+1|k}^{x^*})^{\mathrm{T}} (\boldsymbol{\chi}_{i,k+1|k}^{v})^{\mathrm{T}}]^{\mathrm{T}}$, generated by

$$\begin{aligned} \boldsymbol{\chi}_{0,k+1}^{xv} &= \hat{\mathbf{x}}_{k+1|k}^{xv} \\ \boldsymbol{\chi}_{i,k+1}^{xv} &= \hat{\mathbf{x}}_{k+1|k}^{xv} + hS_{i,k+1|k}^{xv}, \quad i = 1,\ldots,n_{xv} \\ \boldsymbol{\chi}_{i,k+1}^{xv} &= \hat{\mathbf{x}}_{k+1|k}^{xv} - hS_{i,k+1|k}^{xv}, \quad i = n_{xv} + 1,\ldots,2n_{xv} \end{aligned} \tag{13}$$

where $S_{i,k+1|k}^{xv}$ is the ith sigma point column vector of $S_{k+1|k}^{xv}$. Then, the sigma points are propagated forward in time through the nonlinear measurement equation as

$$\mathcal{Y}_{i,k+1|k} = \mathbf{h}\,(\boldsymbol{\chi}_{i,k+1|k}^{x^*}, \boldsymbol{\chi}_{i,k+1|k}^{v}) \tag{14}$$

The predicted observation vector $\hat{\mathbf{y}}_{k+1}^{-} \in \mathcal{R}^{n_y}$ is computed by

$$\hat{\mathbf{y}}_{k+1}^{-} = \sum_{i=0}^{2n_{xv}} w_{y,i}^{(m)}\, \mathcal{Y}_{i,k+1|k} \tag{15}$$

where the weight factors are given by

$$w_{y,0}^{(m)} = \frac{h^2 - (n_x + n_v)}{h^2}, \quad w_{y,i}^{(m)} = w_{x,i}^{(m)} = \frac{1}{2h^2}, (i > 0) \tag{16}$$

The innovation covariance matrix is computed by

$$P_{k+1}^{vv} = S_{k+1}^{vv} (S_{k+1}^{vv})^{\mathrm{T}} \tag{17}$$

where $S^{vv}_{i,k+1}$ is the square root factor of the innovation compound matrix

$$S^{vv}_{k+1} = \mathcal{HT}([S^{(1)}_{v,k+1}(i) \;\; S^{(2)}_{v,k+1}(i)]) \tag{18}$$

and each matrix is given by

$$\begin{aligned} S^{(1)}_{v,k+1}(i) &= w^{(c)}_{d,i}\,[\mathcal{Y}_{i,k+1|k} - \mathcal{Y}_{i+n_{xv},k+1|k}], \quad i = 1,\ldots,n_{xv} \\ S^{(2)}_{v,k+1}(i) &= w^{(c)}_{d,0}\,[\mathcal{Y}_{i,k+1|k} + \mathcal{Y}_{i+n_{xv},k+1|k} - 2\mathcal{Y}_{0,k+1|k}], \quad i = 1,\ldots,n_{xv} \end{aligned} \tag{19}$$

The weight factors $w^{(c)}_{d,i}$ and $w^{(c)}_{d,0}$ are given in equation (11), and the dimension of each compound matrix is $S^{(1)}_{v,k+1}(i) = S^{(2)}_{v,k+1}(i) \in \mathcal{R}^{n_y \times n_{xv}}$. Similarly, the cross-correlation matrix is calculated by

$$P^{xy}_{k+1} = S^{-}_{x,k+1}\,(S^{(1)}_{v,k+1}(i))^{\mathrm{T}}, \quad i = 1,\ldots,n_x \tag{20}$$

where $(S^{(1)}_{v,k+1}\,(i)) \in \mathcal{R}^{n_y \times n_x}$ is the component matrix of the compound matrix in equation (19) with the $n_y \times n_x$ dimension, and $S^{-}_{x,k+1}$ is obtained from equation (9) through the Householder triangularization. Now the filter gain matrix $\mathcal{K}_{k+1}$ is computed by

$$\mathcal{K}_{k+1} = P^{xy}_{k+1}\,[S^{vv}_{k+1}(S^{vv}_{k+1})^{\mathrm{T}}]^{-1} \tag{21}$$

Then, the estimated update state vector $\hat{\mathbf{x}}^{+}_{k+1}$ is given by

$$\hat{\mathbf{x}}^{+}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} + \mathcal{K}_{k+1}\,\boldsymbol{v}_{k+1} \tag{22}$$

where $\boldsymbol{v}_{k+1}$ is the innovation vector, which is equal to the difference between the actual and the predicted observations

$$\mathbf{v}_{k+1} \equiv \tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}^{-}_{k+1} = \tilde{\mathbf{y}}_{k+1} - \sum_{i=0}^{2n_{xv}} w^{(m)}_{y,i}\,\mathbf{Y}_{i,k+1|k} \tag{23}$$

Finally, the square root form of the updated state covariance matrix is expressed by

$$P^{+}_{k+1} = S^{+}_{x,k+1}\,(S^{+}_{x,k+1})^{\mathrm{T}} \tag{24}$$

$$S^{+}_{x,k+1} = \mathcal{HT}([S^{-}_{x,k+1} - \mathcal{K}_{k+1}S^{vv}_{k+1,x} \;\; \mathcal{K}_{k+1}S^{vv}_{k+1,xv}]) \tag{25}$$

where $S^{vv}_{k+1,x} \in \mathcal{R}^{n_y \times n_x}$ and $S^{vv}_{k+1,xv} \in \mathcal{R}^{n_y \times (n_x + 2n_v)}$ consist of the portion of the square root innovation compound matrix as

$$S^{vv}_{k+1,x} = [S^{(1)}_{v,k+1}(1),\ldots,S^{(1)}_{v,k+1}(n_x)] \tag{26}$$

$$S^{vv}_{k+1,xv} = [S^{(1)}_{v,k+1}(n_x + 1),\ldots,S^{(1)}_{v,k+1}(n_{xv}),\, S^{(2)}_{v,k+1}(1),\ldots,S^{(2)}_{v,k+1}(n_{xv})] \tag{27}$$

Attitude System Models

In this section, a brief review of the attitude kinematics and sensor measurement equations based on the modified Rodrigues parameters is presented.

Attitude Kinematics Equations

The quaternion-based attitude kinematics equation of motion is described by [20]

$$\dot{\mathbf{q}} = \frac{1}{2}\,\Omega\,(\boldsymbol{\omega})\,\mathbf{q} = \frac{1}{2}\,\Xi\,(\mathbf{q})\,\boldsymbol{\omega} \tag{28}$$

where $\Omega\,(\boldsymbol{\omega}) \in \mathcal{R}^{4\times4}$ and $\Xi(\mathbf{q}) \in \mathcal{R}^{4\times3}$ are expressed as

$$\Omega(\boldsymbol{\omega}) = \begin{bmatrix} -[\boldsymbol{\omega}\times] & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^{\mathrm{T}} & 0 \end{bmatrix}, \quad \Xi\,(\mathbf{q}) = \begin{bmatrix} q_4 I_{3\times3} + [\mathcal{Q}\times] \\ -\mathcal{Q}^{\mathrm{T}} \end{bmatrix} \tag{29}$$

In equations (28) and (29), $\boldsymbol{\omega} = [\omega_1 \ \omega_2 \ \omega_3]^T$ is the body angular velocity vector, and $\mathbf{q} \equiv [\mathcal{Q}^T \ q_4]^T$ is the quaternion attitude parameters with $\mathcal{Q} \equiv [q_1 \ q_2 \ q_3]^T$ representing directions and q_4 describing the angle of rotation. The terms $[\boldsymbol{\omega}\times]$ and $[\mathcal{Q}\times]$ represent a cross product as a matrix operation defined by

$$[\mathcal{Q}\times] \equiv \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}, \quad [\boldsymbol{\omega}\times] \equiv \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{30}$$

Note that a four-dimensional redundant parameter is used for a three-dimensional attitude system, the quaternions are not independent, and this leads to the normalization constraint given by $\mathbf{q}^T\mathbf{q} = 1$. The modified Rodrigues parameter (MRP) vector $\boldsymbol{\sigma} \in \mathcal{R}^{3\times 1}$ is defined in terms of the quaternions as the transformation [22]

$$\sigma_i = \frac{q_i}{1 + q_4}, \quad i = 1, 2, 3 \tag{31}$$

The MRP vector can be expressed in terms of the principal rotation elements as

$$\boldsymbol{\sigma} \equiv \tan\left(\frac{\Phi}{4}\right)\hat{\mathbf{e}} \tag{32}$$

where Φ is the principal rotation angle and $\hat{\mathbf{e}}$ is the principal axis unit vector. As the principal angle approaches $\pm 2\pi$ (i.e., $q_4 \rightarrow -1$), the MRP becomes singular. This singularity can be avoided by implementing a shadow mapping defined by [22] as

$$\boldsymbol{\sigma}^s \equiv \tan\left(\frac{\Phi - 2\pi}{4}\right)\hat{\mathbf{e}} \tag{33}$$

The shadow MRP vector $\boldsymbol{\sigma}^s$ has a singular orientation at $\Phi = 0°$ as compared to the MRP vector which is singular at $\Phi = \pm 360°$. This allows the MRPs to avoid singularities by switching between the original and shadow MRP as the MRP vector approaches a singular orientation. The components of the shadow MRPs can be expressed by

$$\sigma_i^s = \frac{-\sigma_i}{\sigma^T\sigma} = \frac{-q_i}{1 + q_4}, \quad i = 1, 2, 3 \tag{34}$$

The kinematics differential equation of motion in terms of the MRP vector is expressed by [21] as

$$\dot{\boldsymbol{\sigma}} = \frac{1}{4}\mathcal{B}(\boldsymbol{\sigma})\boldsymbol{\omega} \tag{35}$$

where $\mathcal{B}(\boldsymbol{\sigma}) \in \mathcal{R}^{3\times 3}$ is given by

$$\mathcal{B}(\boldsymbol{\sigma}) = [(1 - \boldsymbol{\sigma}^T\boldsymbol{\sigma})I_{3\times 3} + 2[\boldsymbol{\sigma}\times] + 2\boldsymbol{\sigma}\boldsymbol{\sigma}^T] \tag{36}$$

In many applications, angular rates are provided from a triad of integrating rate gyroscopes. For this gyro-sensor measurement [16], the following equation is applied

$$\bar{\boldsymbol{\omega}}(t) = \boldsymbol{\omega}(t) + \boldsymbol{\beta}(t) + \boldsymbol{\eta}_v(t) \tag{37}$$

$$\dot{\boldsymbol{\beta}}(t) = \boldsymbol{\eta}_u(t) \tag{38}$$

where $\tilde{\boldsymbol{\omega}}(t)$ denotes the measured angular velocity vector, $\boldsymbol{\beta}(t)$ is the gyro bias vector, and $\boldsymbol{\eta}_v(t)$ and $\boldsymbol{\eta}_u(t)$ are the rate-gyro measurement and bias noise processes with zero mean and covariance, $E\{\boldsymbol{\eta}_v(t)\boldsymbol{\eta}_v(\tau)\} = \boldsymbol{\sigma}_v^2\delta(t-\tau)I_{3\times3}$ and $E\{\boldsymbol{\eta}_u(t)\boldsymbol{\eta}_u(\tau)\} = \boldsymbol{\sigma}_u^2\delta(t-\tau)I_{3\times3}$, respectively. Substituting the rate-gyro sensor model in equation (37) into equation (35) gives the nonlinear attitude kinematics equations as

$$\dot{\boldsymbol{\sigma}}(t) = \frac{1}{4}\mathcal{B}(\boldsymbol{\sigma}(t))(\tilde{\boldsymbol{\omega}} - \boldsymbol{\beta}(t)) - \frac{1}{4}\mathcal{B}(\boldsymbol{\sigma}(t))\boldsymbol{\eta}_v(t) \tag{39}$$

Sensor Measurement Model

Suppose a three-axis magnetometer (TAM) is mounted along the body fixed reference frame and measures the geomagnetic field vector in the body fixed reference frame, and the Earth geomagnetic field vector resolved in the inertial reference frame is represented by $\mathbf{r}_k \in \mathcal{R}^{3\times1}$. Then the measurement equation can be written as [21]

$$\tilde{\mathbf{y}}_k = A(\boldsymbol{\sigma}_k)\mathbf{r}_k + \mathbf{v}_k \tag{40}$$

where $\mathbf{v}_k \in \mathcal{R}^{3\times1}$ denotes the measurement noise with mean and covariance matrix $E\{\mathbf{v}_k\,\mathbf{v}_k^{\mathrm{T}}\} = \mathcal{N}(0, \mathbf{R}_k)$, $\mathbf{r}_k \in \mathcal{R}^{3\times1}$ is a reference attitude vector in a reference coordinate system, $\tilde{\mathbf{y}}_k \in \mathcal{R}^{3\times1}$ is the measurement in the spacecraft body coordinate system, and the attitude direction cosine matrix $A(\boldsymbol{\sigma}_k)$ is expressed in terms of the MRPs as

$$A(\boldsymbol{\sigma}_k) = I_{3\times3} + \frac{8[\boldsymbol{\sigma}_k\times]^2 - 4(1-\boldsymbol{\sigma}_k^{\mathrm{T}}\boldsymbol{\sigma}_k)[\boldsymbol{\sigma}_k\times]}{(1+\boldsymbol{\sigma}_k^{\mathrm{T}}\boldsymbol{\sigma}_k)} \tag{41}$$

which is the 3×3 dimensional orthogonal attitude matrix, and satisfies the identity, $A^{\mathrm{T}}(\boldsymbol{\sigma}_k) = A(-\boldsymbol{\sigma}_k)$.

Additive Divided Difference Attitude Filtering

In this section an efficient additive divided difference attitude filtering (ADDAF) algorithm is derived for realtime attitude estimation by integrating the proposed sigma point divided difference filter (SP-DDF) with an additive process and measurement noise concept to minimize the computational workload. The three-dimensional attitude representation is made in terms of the modified Rodrigues parameters.

Consider the nonlinear attitude kinematics and measurement equations with an additive noise vector

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) + \mathbf{g}(\mathbf{x}(t))\mathbf{w}(t) \tag{42}$$

$$\tilde{\mathbf{y}}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k \tag{43}$$

where

$$\mathbf{f}(\mathbf{x}(t)) \equiv \begin{bmatrix} \frac{1}{4}\mathcal{B}(\boldsymbol{\sigma}(t))(\tilde{\boldsymbol{\omega}} - \boldsymbol{\beta}(t)) \\ \mathbf{0}_{3\times1} \end{bmatrix}, \qquad g(\mathbf{x}(t)) \equiv \begin{bmatrix} -\frac{1}{4}\mathcal{B}(\boldsymbol{\sigma}(t)) & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & I_{3\times3} \end{bmatrix} \tag{44}$$

The state vector $\mathbf{x}(t) \equiv [\boldsymbol{\sigma}^{\mathrm{T}}(t)\ \boldsymbol{\beta}^{\mathrm{T}}(t)]^{\mathrm{T}} \in \mathcal{R}^{6\times1}$ includes the three MRP components and the gyro drift-rate bias vector, and the process noise vector $\mathbf{w}(t)$ is defined by $\mathbf{w}(t) \equiv [\boldsymbol{\eta}_v^{\mathrm{T}}(t)\boldsymbol{\eta}_n^{\mathrm{T}}(t)]^{\mathrm{T}}$. The nonlinear measurement equation is defined as $\mathbf{h}(\mathbf{x}_k, k) \equiv A(\boldsymbol{\sigma}_k)\mathbf{r}_k$, and the noise vectors are assumed to be Gaussian and have zero-mean and covariance, $\mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, Q(t))$, and $\mathbf{v}(t) \sim \mathcal{N}(\mathbf{0}, R_k)$, respectively.

Since the nonlinear kinematics and measurement models with additive noises are defined, the number of required sigma point column vectors $\boldsymbol{\chi}_{i,k}^{x}$ is reduced substantially from $i = 2(n_x + n_w)$ to $i = 2n_x$, which alleviates the computational workload in the attitude estimation and makes its real-time implementation suitable. First, applying the additive noise property, a set of weighted symmetric sigma point vectors, $\boldsymbol{\chi}_{i,k}^{x} \in \mathcal{R}^{n_x \times 1}$, $i = 1, \ldots, 2n_x$, is generated by using the non-augmented initial state $\hat{\mathbf{x}}_k^+ \in \mathcal{R}^{6\times 1}$ and the square-root matrix $S_{x,k}^+ \in \mathcal{R}^{6\times 6}$

$$\begin{aligned} \boldsymbol{\chi}_{0,k}^{x} &= \hat{\mathbf{x}}_k^+ \\ \boldsymbol{\chi}_{i,k}^{x} &= \hat{\mathbf{x}}_k^+ + hS_{x,k}^+(i), \quad i = 1, \ldots, n_x \\ \boldsymbol{\chi}_{i,k}^{x} &= \hat{\mathbf{x}}_k^+ - hS_{x,k}^+(i), \quad i = n_x + 1, \ldots, 2n_x \end{aligned} \tag{45}$$

where n_x is the dimension of the state vector ($n_x = 6$), $\hat{\mathbf{x}}_k^+ \equiv [(\hat{\boldsymbol{\sigma}}_k^+)^{\mathrm{T}}\ (\hat{\boldsymbol{\beta}}_k^+)^{\mathrm{T}}]^{\mathrm{T}} \in \mathcal{R}^{6\times 1}$ is the initial state estimate vector at time t_0, and $S_{x,k}^+(i)$ is the ith column vector of the square root matrix $S_{x,k}^+ \in \mathcal{R}^{n_x \times n_x}$ obtained from the Cholesky factorization of the initial state covariance matrix $S_{x,k}^+(S_{x,k}^+)^{\mathrm{T}} \equiv \mathrm{chol}(P_k^+) \in \mathcal{R}^{n_x \times n_x}$ at time t_0. Note the sigma point vector $\boldsymbol{\chi}_{i,k}^{x}$ is generated by using only the state and covariance matrix without augmenting the noise terms. Now, the sigma point vectors are propagated forward in time by numerically integrating the nonlinear differential equation in equation (42) as

$$\boldsymbol{\chi}_{i,k+1|k}^{x} = \boldsymbol{\chi}_{i,k}^{x} + \int_{t_k}^{t_{k+1}} \mathbf{f}(\boldsymbol{\chi}_{i,k}^{x})\mathrm{d}t, \qquad i = 0, \ldots, 2n_x \tag{46}$$

where $\boldsymbol{\chi}_{i,k}^{x}$ is the ith sigma point column vector composed of the three elements $\boldsymbol{\chi}_{i,k}^{\sigma}$ for the MPRs and the next three elements $\boldsymbol{\chi}_{i,k}^{\beta}$ for the gyro bias, respectively, and the nonlinear state mapping function is expressed by

$$\mathbf{f}(\boldsymbol{\chi}_{i,k}^{x}) \equiv \begin{bmatrix} \dfrac{1}{4}\mathcal{B}(\boldsymbol{\chi}_{i,k}^{\sigma})(\tilde{\boldsymbol{\omega}} - \boldsymbol{\chi}_{i,k}^{\beta}) \\ \mathbf{0}_{3\times 1} \end{bmatrix} \tag{47}$$

The predicted MRP and the bias estimate vector $\hat{\mathbf{x}}_{k+1}^- = [(\hat{\boldsymbol{\sigma}}_{k+1}^-)^{\mathrm{T}}\ (\hat{\boldsymbol{\beta}}_{k+1}^-)^{\mathrm{T}}]^{\mathrm{T}}$ is computed by using the propagated sigma points $\boldsymbol{\chi}_{i,k+1|k}^{x} \equiv [(\boldsymbol{\chi}_{i,k+1|k}^{\sigma})^{\mathrm{T}}\ (\boldsymbol{\chi}_{i,k+1|k}^{\beta})^{\mathrm{T}}]^{\mathrm{T}}$ and their associated weight parameters

$$\hat{\mathbf{x}}_{k+1}^- = \sum_{i=0}^{2n_x} w_{x,i}^{(m)} \boldsymbol{\chi}_{i,k+1|k}^{x} \tag{48}$$

where $w_{x,i}^{(m)} = 1/2h^2$, $(i > 0)$ and $w_{x,0}^{(m)}$ given by

$$w_{x,0}^{(m)} = \frac{h^2 - n_x}{h^2} = \frac{h^2 - 6}{h^2} \tag{49}$$

The predicted state covariance matrix $P_{k+1}^- \in \mathcal{R}^{n_x \times n_x}$ is computed in terms of the predicted square root factor and the matrix square root factor of the process noise matrix as

$$P_{k+1}^- = S_{x,k+1}^-(S_{x,k+1}^-)^{\mathrm{T}} \tag{50}$$

where $S_{x,k+1}^- = \mathcal{HT}([S_{x,k+1}^{(1)}(i)\ S_{x,k+1}^{(2)}(i)\ S_{Q,k+1}])$ is the predicted square root factor obtained from the Householder triangularization of the predicted compound matrix with each element formulated in terms of the predicted sigma point vectors

$$S^{(1)}_{x,k+1}(i) = w^{(c)}_{d,i}[\boldsymbol{\chi}^{x}_{i,k+1|k} - \boldsymbol{\chi}^{x}_{i+n_x,k+1|k}], \qquad i = 1, \ldots, n_x$$
$$S^{(2)}_{x,k+1}(i) = w^{(c)}_{d,0}[\boldsymbol{\chi}^{x}_{i,k+1|k} + \boldsymbol{\chi}^{x}_{i+n_x,k+1|k} - 2\boldsymbol{\chi}^{x}_{0,k+1|k}], \qquad i = 1, \ldots, n_x \tag{51}$$

The term $S_{Q,k+1} \equiv \sqrt{Q_k}$ is the square root Cholesky factor of the process noise covariance $Q_k \equiv G(t)Q(t)G^{\mathrm{T}}(t)\Delta t$ which is computed by using $Q(t)$ from the noise vector $\mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, Q(t))$, $G(t) \equiv \mathbf{g}(\mathbf{x}(t))$ in equation (44), and time step Δt.

Note the idea behind the additive sigma point filter is to separate the process noise term from the augmented state vector adopted in the original divided difference filter in order to reduce the computational load in the prediction step. To take into account the process noise uncertainty, a set of sigma point vectors is regenerated by combining in the predicted state vector with the square root of either the process noise covariance matrix or the predicted state covariance matrix prior to the measurement update step [24]. In the measurement prediction step, a set of new sigma point vectors is obtained by combining the estimated state $\hat{\mathbf{x}}^{-}_{k+1} = [(\hat{\boldsymbol{\sigma}}^{-}_{k+1})^{\mathrm{T}}\ (\hat{\boldsymbol{\beta}}^{-}_{k+1})^{\mathrm{T}}]^{\mathrm{T}}$ with the square root factor of the predicted state covariance matrix P^{-}_{k+1} as

$$\chi^{*}_{k+1|k} = [\hat{\mathbf{x}}^{-}_{k+1} \quad \hat{\mathbf{x}}^{-}_{k+1} + hS^{-}_{x,k+1} \quad \hat{\mathbf{x}}^{-}_{k+1} - hS^{-}_{x,k+1}] \tag{52}$$

where $S^{-}_{x,k+1}$ is the square root factor obtained from equation (50).

Now, the new sigma points are propagated forward in time through the nonlinear measurement equation as

$$\mathcal{Y}_{i,k+1|k} = \mathbf{h}(\boldsymbol{\chi}^{*}_{i,k+1|k}, k+1) = A(\boldsymbol{\chi}^{\sigma^*}_{i,k+1|k})\mathbf{r}_{k+1}, \qquad i = 0, \ldots, 2n_x \tag{53}$$

where the nonlinear measurement equation $\mathbf{h}(\boldsymbol{\chi}^{*}_{i,k+1}, k+1)$ is expressed in terms of the direction cosine matrix $A(\boldsymbol{\chi}^{\sigma^*}_{i,k+1|k})$ and the Earth geomagnetic field vector $\mathbf{r}_{k+1}$ resolved in the inertial reference frame. Then, the predicted observation vector $\hat{\mathbf{y}}^{-}_{k+1} \in \mathcal{R}^{n_y}$ is computed by

$$\hat{\mathbf{y}}^{-}_{k+1} = \sum_{i=0}^{2n_x} w^{(m)}_{y,i}\mathcal{Y}_{i,k+1|k} = \sum_{i=0}^{2n_x} w^{(m)}_{y,i}A(\boldsymbol{\chi}^{*}_{i,k+1|k})\mathbf{r}_{k+1} \tag{54}$$

where $w^{(m)}_{y,i}$ is identical to $w^{(m)}_{x,i} = 1/2h^2$ and $w^{(m)}_{y,0}$ is given by

$$w^{(m)}_{y,0} = \frac{h^2 - n_x}{h^2} = \frac{h^2 - 6}{h^2} \tag{55}$$

The innovation covariance matrix is formulated by

$$P^{vv}_{k+1} = S^{vv}_{k+1}(S^{vv}_{k+1})^{\mathrm{T}} \tag{56}$$

where S^{vv}_{k+1} is the square-root factor from the Householder triangularization of the innovation compound matrix

$$S^{vv}_{k+1} = \mathcal{HT}([S^{(1)}_{v,k+1}(i) \quad S^{(2)}_{v,k+1}(i) \quad S_{R,k+1}]) \tag{57}$$
$$S^{(1)}_{v,k+1}(i) = w^{(c)}_{d,i}[\mathcal{Y}_{i,k+1|k} - \mathcal{Y}_{i+n_x,k+1|k}], \qquad i = 1, \ldots, n_x$$
$$S^{(2)}_{v,k+1}(i) = w^{(c)}_{d,0}[\mathcal{Y}_{i,k+1|k} + \mathcal{Y}_{i+n_x,k+1|k} - 2\mathcal{Y}_{0,k+1|k}], \qquad i = 1, \ldots, n_x \tag{58}$$

and $S_{R,k+1} \equiv \sqrt{R_{k+1}}$ is the square-root Cholesky factor of the measurement noise covariance R_{k+1}. The weights $w^{(c)}_{d,i}$ and $w^{(c)}_{d,0}$ are given in equation (11), and the dimension of each compound matrix is $S^{(1)}_{v,k+1}(i) = S^{(2)}_{v,k+1}(i) \in \mathcal{R}^{n_y \times n_x}$. Similarly, the cross correlation matrix is calculated by

$$P^{xy}_{k+1} = S^{-}_{x,k+1}(S^{(1)}_{v,k+1})^{\mathrm{T}} \tag{59}$$

where $S^{(1)}_{v,k+1} \in R^{n_y \times n_x}$ is the component matrix in equation (58), and $S^{-}_{x,k+1}$ is obtained from equation (50) through the Householder triangularization. Now, the filter gain matrix $\mathcal{K}_{k+1}$ is computed from equation (22), and the innovation vector $\boldsymbol{v}_{k+1}$ is given by

$$\boldsymbol{v}_{k+1} \equiv \tilde{\mathbf{y}}_{k+1} - \hat{\mathbf{y}}^{-}_{k+1} = \tilde{\mathbf{y}}_{k+1} - \sum_{i=0}^{2n_x} w^{(m)}_{y,i} \mathcal{Y}_{i,k+1|k} \tag{60}$$

where $\tilde{\mathbf{y}}_{k+1}$ is the TAM measurement vector in the spacecraft body coordinate system. Then, the estimated update state vector $\hat{\mathbf{x}}^{+}_{k+1} \equiv [(\hat{\boldsymbol{\sigma}}^{+}_{k+1})^{\mathrm{T}} (\hat{\boldsymbol{\beta}}^{+}_{k+1})^{\mathrm{T}}]^{\mathrm{T}} \in R^{6\times 1}$ is calculated by $\hat{\mathbf{x}}^{+}_{k+1} = \hat{\mathbf{x}}^{-}_{k+1} + \mathcal{K}_{k+1}\boldsymbol{v}_{k+1}$. Finally, the square root form of the updated state covariance matrix is expressed by

$$P^{+}_{k+1} = S^{+}_{x,k+1}(S^{+}_{x,k+1})^{\mathrm{T}} \tag{61}$$

$$S^{+}_{x,k+1} = \mathcal{HT}([S^{-}_{x,k+1} - \mathcal{K}_{k+1}S^{(1)}_{v,k+1} \quad \mathcal{K}_{k+1}S^{vv}_{k+1,xv}]) \tag{62}$$

where $S^{(1)}_{v,k+1} = [S^{(1)}_{v,k+1}(1), \ldots, S^{(1)}_{v,k+1}(n_x)]$, and $S^{vv}_{k+1,xv} = [S^{(2)}_{v,k+1}(1), \ldots, S^{(2)}_{v,k+1}(n_{xv})S_{R,k+1}]$.

Note that the proposed ADDAF requires only one single scalar parameter, the divided difference step size h, and for Gaussian distribution, its optimal value is $h = \sqrt{3}$ [10]. In contrast, the unscented Kalman filter based attitude estimation algorithm uses at least three parameters that need to be tuned for optimal performance. The sensitivity of the scale parameters in the SR-DDF and the UKF is investigated through simulation.

Simulation Results

In this section the performance of the MRP-based additive divided difference attitude filter is compared with the quaternion-based unscented attitude filter and the quaternion-based extended Kalman filter through simulation examples in terms of the estimation accuracy and convergence properties using realistic satellite and observation models. The Earth-orbiting spacecraft has the following near-circular orbit parameters: $a = 6728.136$ km, $e = 1.0 \times 10^{-5}$, $i = 35°$, and $\Omega = 25°$. In this paper, the spacecraft's orbit is propagated by using the multiple sequential compression (MSC) orbit propagator [25]. It is assumed that the Earth-pointing satellite under consideration is equipped with a three-axis magnetometer and integrating rate gyros. The TAM measurements are used to test the convergence properties of the proposed sigma point attitude filter because the TAM yields slower convergence than using two or more attitude sensors such as GPS, Sun sensor, and star tracker. Data information used in this simulation study is based on reference [21]. The noise of the TAM sensor is modeled by a Gaussian white noise with a standard deviation of 0.3 mG (30 nT). Thus, its measurement covariance matrix is given by $R_k = 0.3^2 I_{3\times 3}$. The magnetic field of the Earth is represented by a spherical harmonics approximation and it is modeled by using the coefficients of the 10th-order of the international geomagnetic reference field model [26]. The gyro measurements are simulated with a gyro noise standard deviation of $\sigma_v = 0.062$ deg/hr, the drift bias noise is modeled with a standard deviation of $\sigma_u = 0.235$ deg/hr/hr, and an initial drift of 0.1 deg/hr on each axis. The sampling rate of the TAM and rate gyro measurements is 0.1 Hz. In this study, the average root mean square (ARMS) of the magnitude of the Euler estimation errors is used for the quantitative performance comparison of the proposed filters along with plots of the estimated state errors with the 3-sigma outlier.

In the first simulation, the performance of two divided difference attitude filters (DDAFs) with the MRP attitude parameterization is compared. One is the augmented DDAF with the noise vectors augmented in the state vector, and the other is the additive DDAF with only the attitude state estimates. Initial roll, pitch, and yaw angle errors of 20°, 70°, and −20° are used, respectively. These simulation conditions illustrate a realistic scenario with unknown or disturbed attitude information due to sensor and/or actuator failure, or represent the typical cases during satellite in-orbit commissioning phase and after computer reset. The initial attitude covariance is diagonal and set to $P_{att} = (10 \text{ deg})^2$ and bias errors are set to $P_{bias} = (0.2 \text{ deg/hr})^2$ The interval length $h = \sqrt{3}$ is set for both the DDAFs, and it is the optimal value for the Gaussian distribution. Figure 1 depicts the absolute magnitude of the attitude estimation errors. As can be seen, there is not much difference between them for the attitude estimate errors. The estimation errors for each axis in terms of the Euler angles with the 3-sigma outlier are shown in Fig. 2. The attitude errors of the additive DDAF do converge to within the 3-sigma bound within eight hours. This result indicates the DDAF is performing in a near optimal fashion. The computational workload of the augmented DDAF is more than that of the additive DDAF. The average computational time per one cycle of the augmented DDAF is 0.095 sec, while the additive DDAF requires 0.078 sec, and in terms of the normalized factor the augmented DDAF is about 1.2, which means it requires about 20% more computational time. This is because the augmented DDAF with noise vectors augmented into the state vector implements a 12-dimensional state vector to achieve attitude estimates, while the additive DDAF utilizes only the six-dimensional state vector. In this first simulation, it is seen that the additive divided difference filter as a MRP-based attitude filter achieves not only a reliable

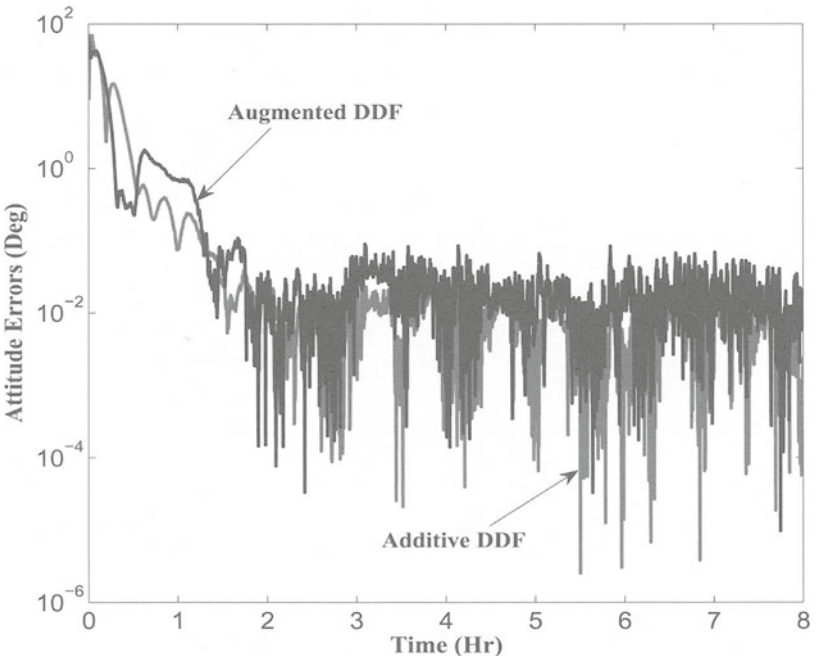

FIG. 1. Absolute Magnitude of Position Errors from Distributed Estimation.

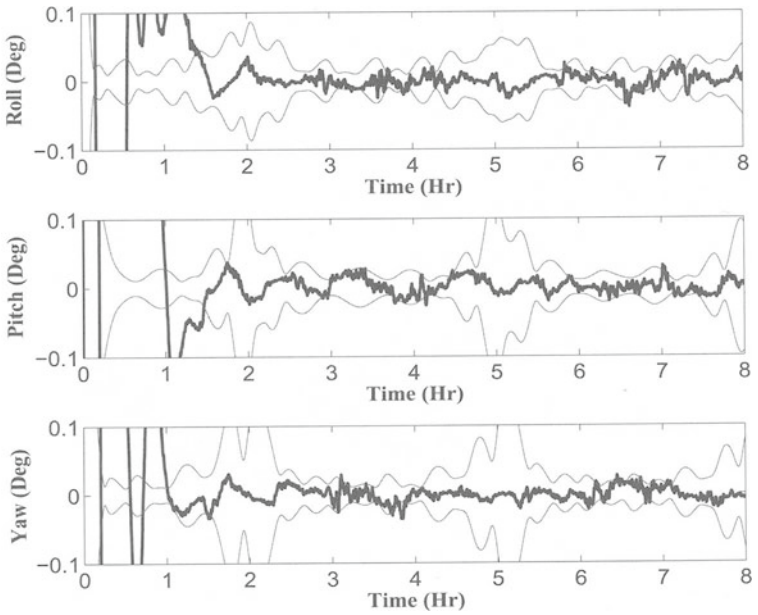

FIG. 2. Attitude Estimation Errors with 3σ Error Outlier.

degree of estimation accuracy, but its computational efficiency is suitable for real-time applications.

In the next example, the performance between the MRP-based additive divided difference attitude filter and the unscented filter and the extended Kalman filter, that are parameterized in terms of the quaternions, are compared. The same values of the initial Euler angle errors for roll, pitch, and yaw are considered. The same initial attitude covariance is used as in the first example. The interval length $h = \sqrt{3}$ is set for a Gaussian distribution in the DDF. For the UKF attitude estimator, the generalized parameters are set to $a = 1$ with $f = 2(a + 1) = 4$, and the scale factor is set to $\lambda = 1$.

In Fig. 3 the norms of the attitude estimation errors are depicted. There is not much difference between the additive DDAF and the UKF. However, the EKF takes almost six hours to converge to a value below 0.1°, while the DDAF and the UKF converge to this value within 1.2 hours, and take less than two hours to achieve a converged solution. As can been seen in Fig. 3, the sigma point attitude filters, the DDAF, and UKF converge faster than the EKF, which indicates the nonlinearities from the higher-order terms are not negligible, but important in robust and reliable estimation.

Figure 4 shows the performance comparison between the three attitude filters in terms of the computational speed. A scale factor that is normalized with respect to the additive DDAF is used to check the computational speed. The EKF has a 0.7 normalized scale factor with respect to the DDAF, and the UKF has almost identical computational speed with the DDAF. Even though the standard type EKF has less computational workload than the SR-DDAF, there is a marginal stability issue of the numerical solution of the associated Riccati equation in the standard EKF application. Thus, it is necessary to implement a numerical stabilization routine with a square-root type factorization for the covariance propagation and update equations.

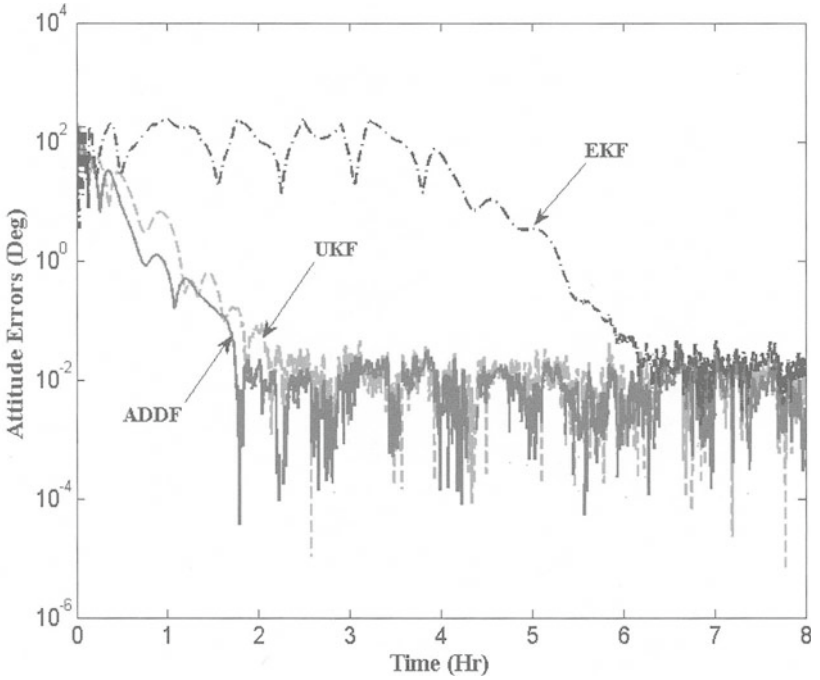

FIG. 3. Absolute Magnitude of Attitude Estimation Errors and Performance Comparison.

Since the numerical stability of the EKF implementation is obtained at the expense of added computational workload, it is expected that the computational workload between the SR-DDAF and the SR-EKF is almost identical. Therefore, the sigma

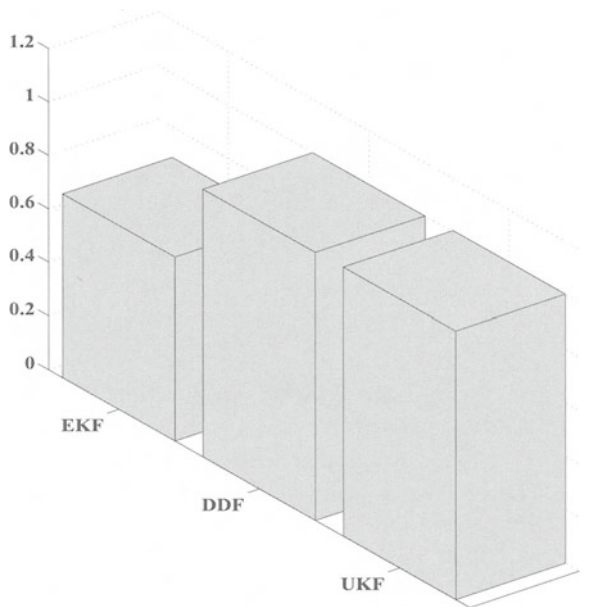

FIG. 4. Comparison for Normalized Computational Speed.

point attitude filters, the DDAF and the UKF, are also suitable for real-time applications as well as the SR-EKF.

Now, the performance of the DDAF and the UKF is investigated with respect to variations of the weight scale factors required to optimize their estimation accuracy. The same initial conditions are used, but the weight scale factor for each sigma point filter is changed to check the sensitivity. Figure 5 depicts the norm of the attitude estimation errors of the UKF with respect to the variation of the generalized attitude parameter a and the weight scale factor $\lambda = \alpha^2(n_x + \kappa) - n_x$ which is a function of the control parameters, $0 < \alpha < 1$ and κ. In this simulation, the weight scale factor λ is selected directly instead of computing through the control parameters, α and κ. Four different cases are tested with values $a = 0.1$ and $\lambda = 0, 1, -3$. The best result is obtained with $a = 1$, $(f = 2(a + 1) = 4)$ and $\lambda = 0$. The result with $a = 1$ and $\lambda = 1$ has almost the same accuracy as the first case. However, the cases with $a = 0$ and $\lambda = 1$, and $a = 1$ and $\lambda = -3$ have degraded performance with slower convergence than the first two cases. Through this simulation, it is recognized that the scale factors of the UKF play an important role in producing a reliable estimate when the Gaussian approximation is not valid due to the nonlinear transformation, that is, the UKF performance is sensitive to the values of the parameters.

In Fig. 6 the sensitivity analysis of the DDAF to the variations of the interval-length parameter h is investigated. The interval length is usually set to $h = \sqrt{3}$ for a Gaussian distribution in the DDF in order to obtain an optimal solution. The interval length is changed with a range from $1 < h \leq \sqrt{3.5}$. The best performance is obtained with $h = \sqrt{3}$, and the case with $h = \sqrt{2}$ gives the worst performance.

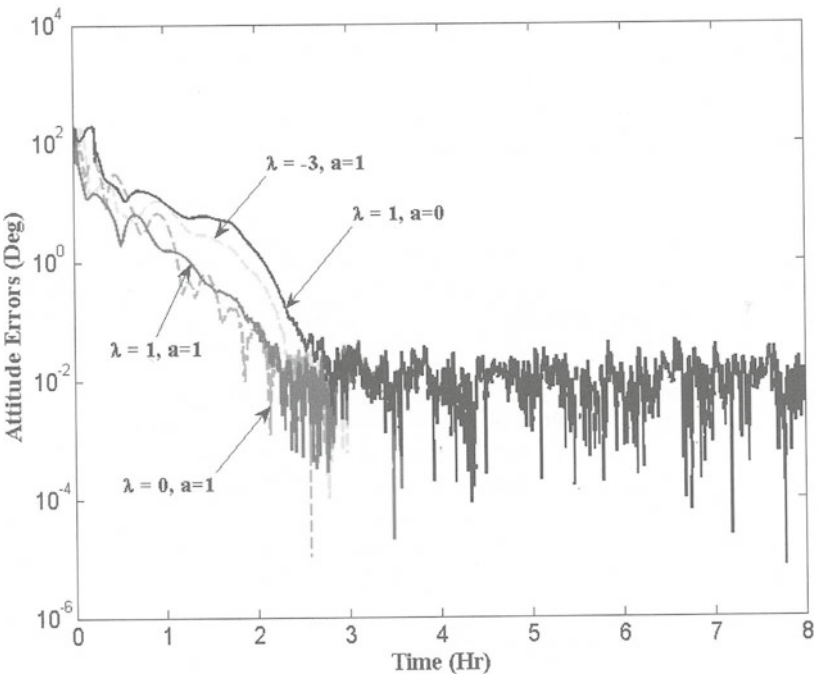

FIG. 5. Absolute Magnitude of Attitude Estimation Errors with Variations of Parameters a and λ.

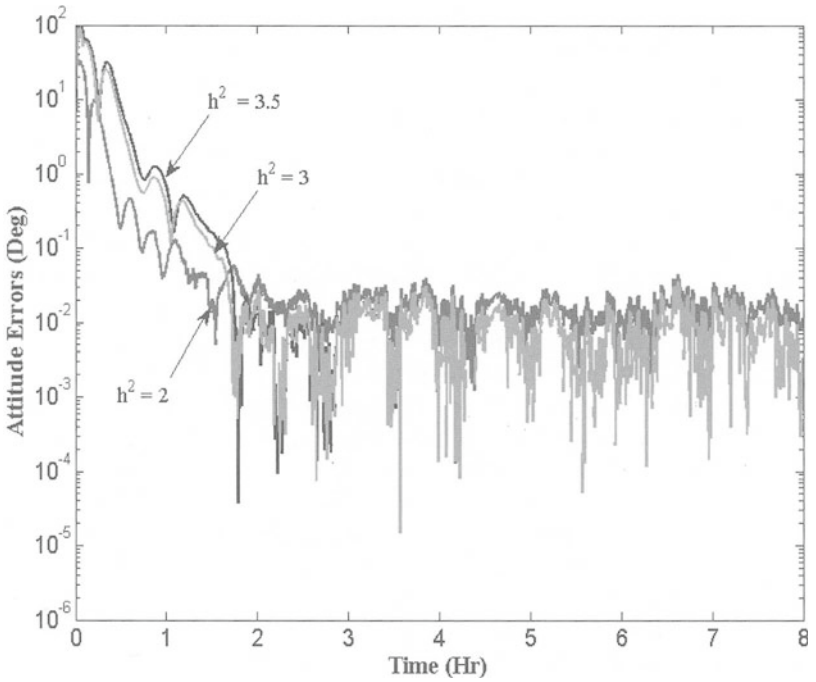

FIG. 6. Absolute Magnitude of Attitude Estimation Errors with Variations of Scale Factor h.

However, the variations of the interval length do not have much effect on the performance of the additive DDAF, which means the additive DDAF is robust to the variation of the interval length. This makes its implementation (tuning) easy in real-time applications. It is seen that the main advantage of the additive DDAF over the UKF is its robustness to the variations of the scale factors. For the ADDAF the scale factor h does not require burdensome tuning and its optimal value is $h = \sqrt{3}$, which is the value for a Gaussian distribution.

Conclusion

The purpose of this paper is to propose an efficient attitude estimation algorithm that not only provides reliable estimation accuracy, but also alleviates the computational load with an enhanced computational stability. Two tactical approaches are introduced to achieve the objectives. First, a sigma point form of the divided difference filter is formulated in terms of sigma point vectors with a square-root structure to enhance the computational efficiency. Second, in order to alleviate the computational workload of the standard divided difference filter an additive process and measurement noise concept is integrated, which leads to an additive divided difference filtering algorithm. The additive divided difference attitude filter based on the modified Rodrigues parameters has a computational workload comparable to the extended Kalman filter, but it is superior in terms of the estimate accuracy and convergence efficiency. In particular, the robustness of the additive divided difference filter to the variation of the weight scale factors makes its real-time implementation easier without requiring the burdensome tuning step as opposed to the unscented Kalman filter that needs to tune three parameters. The advantages

of the proposed nonlinear attitude filtering algorithm make it suitable for efficient nonlinear estimation, not only in real-time attitude estimation, but also in other application areas.

References

[1] WERTZ, J.R. *Mission Geometry: Orbit and Constellation Design and Management*, Microcosm Press, El Segundo, CA, 2001, Chapter 1.

[2] BROWN, R.G. and HWANG, P.Y.C. *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley and Sons, New York, NY, 1997.

[3] JAZWINSKI, A.H. *Stochastic Processes and Filtering Theory*, Academic Press, San Diego, CA, 1970.

[4] GORDON, N.J., SALMON, D.J., and SMITH, A.F.M. "A Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE Proceedings on Radar and Signal Processing*, Vol. 140, 1993, pp. 107–113.

[5] ARULAMPALAM, M.S., MASKELL, S., GORDON, N.J., and CLAPP, T. "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking," *IEEE Transactions on Signal Processing*, Vol. 50, No. 2, Feb. 2002, pp. 174–188.

[6] RISTIC, B., ARULAMPALAM, M.S., and GORDON, N.J. *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, Boston, MA, 2004.

[7] FOX, D. "Adapting the Sample Size in Particle Filters through KLD-Sampling," *International Journal of Robotics Research*, Vol. 22, 2003, pp. 985–1003.

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

[9] SCHEI, T.S. "A Finite-Difference Method for Linearization in Nonlinear Estimation Algorithms," *Automatica*, Vol. 33, No. 11, November 1997, pp. 2053–2058.

[10] NORGAARD, M., POULSEN, N.K., and RAVN, O. "New Developments in State Estimation for Nonlinear Systems," *Automatica*, Vol. 36, No. 11, November 2000, pp. 1627–1638.

[11] OSHMAN, Y. and CARMI, A. "Estimating Attitude from Vector Observations Using Genetic-Algorithm-Embedded Quaternion Particle Filter," *Journal of Guidance, Control, and Dynamics*, Vol. 29, No. 4, 2006, pp. 879–891.

[12] CHENG, Y. and CRASSIDIS, J.L. "Particle Filtering for Sequential Spacecraft Attitude Estimation," presented as paper AIAA 2004-5337 at the AIAA Guidance, Navigation, and Control Conference, Providence, Rhode Island, Aug. 2004.

[13] CARMI, A. and OSHMAN, Y. "Fast Particle Filtering for Attitude and Angular-Rate Estimation from Vector Observations," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 1, 2009, pp. 70–78.

[14] VAN DER MERWE, R., WAN, E.A., and JULIER, S.J. "Sigma-Point Kalman Filters for Nonlinear Estimation and Sensor Fusion: Applications to Integrated Navigation," presented as paper AIAA 2004-5120 at the AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, Rhode Island, August 2004.

[15] LEE, D.-J. and ALFRIEND, K.T. "Sigma Point Filtering for Sequential Orbit Estimation and Prediction," *Journal of Spacecraft and Rockets*, Vol. 44, No. 2, 2007, pp. 388–398.

[16] CRASSIDIS, J.L. and MARKLEY, F.L. "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, July–Aug. 2003, pp. 536–542.

[17] LEE, D.-J. and ALFRIEND, K.T. "Adaptive Sigma Point Filtering for State and Parameter Estimation," presented as paper AIAA 2004-5101 at the AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island, August 2004.

[18] VAN DER MERWE, R. and WAN, E.A. "The Square-Root Unscented Kalman Filter for State and Parameter Estimation," *IEEE Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, Utah, May 2001.

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

[20] 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, Sep.–Oct. 1982, pp. 417–429.

[21] 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, 1996, pp. 71–83.

[22] SCHAUB, H. and JUNKINS, J.L. *Analytical Mechanics of Space Systems*, AIAA Education Series, Chapter 3, pp. 107–115.

[23] GREWAL, M.S. and ANDREWS, A.P. *Kalman Filtering: Theory and Practice, 2nd* ed, John Wiley & Sons, Inc., New York, NY, 2001, Chapter 6.

[24] VAN DER MERWE, R. *Sigma Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models*, Ph.D. Dissertation, OGI School of Science & Engineering at Oregon Health & Science University, Portland, Oregon, April 2004.

[25] LEE, D.-J., NO, T.S., LEE, S.-R., KIM, H.-J., and ALFRIEND, K.T. "Precise Ephemeris Reconstruction Using the Clohessy-Wiltshire Frame and Multiple Sequential Compressions," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 5, Sep.–Oct. 2003, pp. 781–785.

[26] IAGA International Geomagnetic Reference Field Model. Available at http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html.

Appendix: Sigma Point Formulation

In this appendix, it is shown that the computation of the second-moment of the predicted state defined in equation (9) is simplified by expressing the covariance matrix in terms of the predicted sigma point vectors.

First, the original predicted state covariance matrix $P_{k+1}^{-} \in \mathcal{R}^{n_x \times n_x}$ in reference [10] is obtained by

$$P_{k+1}^{-} = S_{x,k+1}^{-}(S_{x,k+1}^{-})^{\mathrm{T}} \tag{63}$$

where the compound matrix $S_{x,k+1}^{-} \in \mathcal{R}^{n_x \times 2n_x}$ can be computed by using sigma points vectors

$$S_{x,k+1}^{-} = [S_{xx,k+1}^{(1)}(i)\, S_{xw,k+1}^{(1)}(i)\, S_{xx,k+1}^{(2)}(i)\, S_{xw,k+1}^{(2)}(i)] \tag{64}$$

$$S_{xx,k+1}^{(1)}(i) = w_{d,i}^{(c)}[\mathbf{f}(\boldsymbol{\chi}_{i,k}^{x}, \boldsymbol{\chi}_{0,k}^{w}) - \mathbf{f}(\boldsymbol{\chi}_{i+n_{xw},k}^{x}, \boldsymbol{\chi}_{0,k}^{w})], \qquad i = 1, \ldots, n_x$$

$$S_{xw,k+1}^{(1)}(i) = w_{d,i}^{(c)}[\mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{i,k}^{w}) - \mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{i+n_{xw},k}^{w})], \qquad i = 1, \ldots, n_w \tag{65}$$

$$S_{xx,k+1}^{(2)}(i) = w_{d,0}^{(c)}[\mathbf{f}(\boldsymbol{\chi}_{i,k}^{x}, \boldsymbol{\chi}_{0,k}^{w}) + \mathbf{f}(\boldsymbol{\chi}_{i+n_{xw},k}^{x}, \boldsymbol{\chi}_{0,k}^{w}) - 2\mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{0,k}^{w})], \quad i = 1, \ldots, n_x$$

$$S_{xw,k+1}^{(2)}(i) = w_{d,0}^{(c)}[\mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{i,k}^{w}) + \mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{i+n_{xw},k}^{w}) - 2\mathbf{f}(\boldsymbol{\chi}_{0,k}^{x}, \boldsymbol{\chi}_{0,k}^{w})], \quad i = 1, \ldots, n_w \tag{66}$$

Second, the square root compound matrix in equation (64) can be simplified as

$$S_{x,k+1}^{-} = [S_{x,k+1}^{(1)}(i) \;\; S_{x,k+1}^{(2)}(i)] \tag{67}$$

$$S_{x,k+1}^{(1)}(i) = w_{d,i}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} - \boldsymbol{\chi}_{i+n_x,k+1|k}^{x}], \quad i = 1, \ldots, n_{xw}$$

$$S_{x,k+1}^{(2)}(i) = w_{d,0}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} - \boldsymbol{\chi}_{i+n_x,k+1|k}^{x} - 2\boldsymbol{\chi}_{0,k+1|k}^{x}], \quad i = 1, \ldots, n_{xw} \tag{68}$$

To prove the above equations, suppose a predicted sigma point vector is obtained from

$$\boldsymbol{\chi}_{i,k+1|k}^{x} = \mathbf{f}(\boldsymbol{\chi}_{i,k}^{x}, \boldsymbol{\chi}_{i,k}^{w}, k), \qquad i = 0, 1, \ldots, (n_x + n_w) \tag{69}$$

Then, each compound matrix in equations (65) and (66) can be written in terms of the predicted sigma point vectors as

$$S_{xx,k+1}^{(1)}(i) = w_{d,i}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} - \boldsymbol{\chi}_{i+n_{xw}+n_{xw},k+1|k}^{x}], \qquad i = 1, \ldots, n_x$$

$$S_{xw,k+1}^{(1)}(i) = w_{d,i}^{(c)}[\boldsymbol{\chi}_{i+n_x,k+1|k}^{x} - \boldsymbol{\chi}_{i+n_{xw}+n_x,k+1|k}^{x}], \qquad i = 1, \ldots, n_w \tag{70}$$

$$S_{xx,k+1}^{(2)}(i) = w_{d,i}^{(c)}[\boldsymbol{\chi}_{i,k+1|k}^{x} + \boldsymbol{\chi}_{i+n_{xw}+n_x,k+1|k}^{x} - 2\boldsymbol{\chi}_{0,k+1|k}^{x}], \qquad i = 1, \ldots, n_x$$

$$S_{xw,k+1}^{(2)}(i) = w_{d,i}^{(c)}[\boldsymbol{\chi}_{i+n_x,k+1|k}^{x} + \boldsymbol{\chi}_{i+n_{xw}+n_x,k+1|k}^{x} - 2\boldsymbol{\chi}_{0,k+1|k}^{x}], \qquad i = 1, \ldots, n_w \tag{71}$$

Now, integration by augmenting two matrices in equation (70) and equation (71), respectively, leads to

$$S^{(1)}_{x,k+1}(i) = [S^{(1)}_{xx,k+1}(l) S^{(1)}_{xw,k+1}(j)], \qquad l = 1, \ldots, n_x, j = 1, \ldots, n_w \tag{72}$$

$$S^{(2)}_{x,k+1}(i) = [S^{(2)}_{xx,k+1}(l) S^{(2)}_{xw,k+1}(j)], \qquad l = 1, \ldots, n_x, j = 1, \ldots, n_w \tag{73}$$

Finally, each compound matrix can be written as

$$\begin{aligned} S^{(1)}_{x,k+1}(i) &= w^{(c)}_{d,i}[\boldsymbol{\chi}^{x}_{i,k+1|k} - \boldsymbol{\chi}^{x}_{i+n_x,k+1|k}], \quad i = 1, \ldots, (n_x + n_w) \\ S^{(2)}_{x,k+1}(i) &= w^{(c)}_{d,0}[\boldsymbol{\chi}^{x}_{i,k+1|k} - \boldsymbol{\chi}^{x}_{i+n_x,k+1|k} - 2\boldsymbol{\chi}^{x}_{0,k+1|k}], \quad i = 1, \ldots, (n_x + n_w) \end{aligned} \tag{74}$$

Similarly, the innovation covariance matrix given in equation (18) can be reformulated in a simplified compound matrix form and reduced to equation (19).