Attitude Estimation Without Rate Gyros Using Generalized Multiple Model Adaptive Estimation

Jeremy M. Marschke,* John L. Crassidis†

University at Buffalo, State University of New York, Amherst, NY 14260-4400

Quang M. Lam‡

Orbital Sciences Corporation, Dulles, VA 20166

A generalized multiple model adaptive estimation (GMMAE) scheme is derived to determine the attitude of a spacecraft without the use of rate information provided by gyros. Multiple model adaptive estimation (MMAE) uses several extended Kalman filters (EFKs) running in parallel, each representing a hypothesis of the actual system, to generate enhanced state and parameter estimates. The estimates of each parallel filter are combined based on the likelihood that each hypothesis is correct, which is determined from measurement residuals. GMMAE is an extension of this approach wherein a window of previous-time data is used via the autocorrelation matrix to perform the adaptive update. This approach shows significant improvement over both the standard EKF and standard MMAE approaches in both accuracy and convergence of the estimates. Each of the multiple models within the filter makes a separate hypothesis as to the process noise covariance of the system, and the combination of these separate filters allows it to be adaptively estimated. The filter formulation is based on the standard equations of attitude dynamics, with global attitude parameterization given by a quaternion. A multiplicative quaternion-error approach is used to guarantee that quaternion normalization is maintained in the filters. A Markov model approach for modeling the angular rates, which requires no torque input, is considered. Simulation results are provided to show the effectiveness of the GMMAE scheme.

I. Introduction

Since the launch of Explorer 1 in 1958, the satellite has become an integral part of both the scientific researcher's arsenal and the world's communication infrastructure. As such, necessity for an increase in satellite functional capability, life, precision pointing, attitude estimation accuracy, and maneuverability abounded. One manner in which functional improvements can be realized is through the preparation for all possible eventualities within the useful life of the satellite. The deterioration or failure of rate sensors known as gyros is one such eventuality.

A typical spacecraft attitude determination and control system (ADCS) contains sensors for both the attitude of the vehicle (i.e. star trackers, Sun sensors, etc.) and its body rates (i.e. gyros). However, it is possible that at some point during its mission, spacecraft ADCS may face the situation that no rate information is available. Whether due to a temporary gyro outage or a permanent gyro damage, this situation can lead to serious spacecraft instability and even potentially terminate the mission. This motivates the modern ADCS designer to search for new and innovative ways to make the system robust to any such failures, and thus meet the needs of the mission more completely.

Filtering algorithms, such as the extended Kalman filter (EKF),[1] the Unscented filter (UF)[2] and Particle filters (PFs),[3,4] are commonly used to both estimate unmeasurable states and filter noisy measurements. The

*Graduate Student, Department of Mechanical & Aerospace Engineering. Email: j253m@eng.buffalo.edu. Student Member AIAA.

†Professor, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Associate Fellow AIAA.

‡Chief Scientist, Orbital Sciences Corporation. Email: Lam.Quang@orbital.com. Senior Member AIAA.

EKF and UF assume that the process noise and measurement noise are represented by zero-mean Gaussian white-noise processes. Even if this is true, both filters only provide approximate solutions when the state and/or measurement models are nonlinear, since the posterior density function is most often non-Gaussian. The EKF typically works well only in the region where the first-order Taylor-series linearization adequately approximates the non-Gaussian probability density function (pdf). The Unscented filter works on the premise that with a fixed number of parameters it should be easier to approximate a Gaussian distribution than to approximate an arbitrary nonlinear function. This in essence can provide higher-order moments for the computation of the posterior function without the need to calculate jacobian matrices as required in the EKF. Still, the standard form of the EKF has remained the most popular method for nonlinear estimation to this day, and other designs are investigated only when the performance of this standard form is not sufficient.

Like other approximate approaches to optimal filtering, the ultimate objective of a PF is to construct the posterior pdf of the state vector, or the pdf of the state vector conditioned on all the available measurements. However, the approximation of a PF is vastly different from that of conventional nonlinear filters. The central idea of the PF approximation is to represent a continuous distribution of interest by a finite (but large) number of weighted random samples of the state vector, or particles. Particle filters do not assume the posterior distribution of the state vector to be a Gaussian distribution or any other distribution of known form. In principle, they can estimate probability distributions of arbitrary form and solve any nonlinear and/or non-Gaussian system.

Even if the process noise and/or measurement noise are Gaussian, all standard forms of the EKF, UF and PFs require knowledge of their characteristics, such as the mean and covariance for a Gaussian process. The covariance and mean of the measurement noise can be inferred from statistical inferences and calibration procedures of the hardware sensing devices. The calibration procedures can also be used to determine the nature of the measurement process distribution. The kurtosis characterizes the relative compactness of the distribution around the mean, relative to a Gaussian distribution. A common kurtosis, called the "Pearson kurtosis," divides the fourth moment by the second moment.[5] Positive kurtosis indicates a relatively peaked distribution, while negative kurtosis indicates a relatively flat distribution. However, the process noise is extremely difficult to characterize because it is usually used to represent modeling errors. The covariance is usually determined by *ad hoc* or heuristic approaches, which leads to the classical "tuning of the filter" problem. Fortunately, there are tools available to aid the filter designer. For example, several tests can be applied to check the *consistency* of the filter from the desired characteristics of the measurement residuals. These include the normalized error square test, the autocorrelation test and the normalized mean error test.[6] These tests can, at the very least, provide mechanisms to show that a filter is not performing in an optimal or desired fashion.

In practice the tuning of a filter can be arduous and time consuming. A classic approach to overcome this difficulty is to use adaptive filters. Adaptive filtering can be divided into four general categories: Bayesian, maximum likelihood, covariance matching, and correlation approaches.[7] Bayesian and maximum likelihood methods may be well suited to a multi-model approaches, but sometimes require large computational loads. Covariance matching is the computation of the covariances from the residuals of the state estimation problem, but have been shown to give biased estimates of the true covariances. A widely used correlation-based approach for a linear Kalman filter with stationary/Gaussian process and measurement noise is based on "residual whitening".[8] In particular, the autocorrelation matrix, which can be computed from the measurement-minus-estimate residuals, is used with the system state matrices to provide a least-squares estimate of the Kalman filter error covariance times the measurement output matrix. If the number of unknowns in the process noise covariance is equal to or less than the number of states times the number of outputs, then the error-covariance/output-matrix estimate can be used to find an estimate of the process noise covariance by solving for a set of linear equations. These equations are not linearly independent and one has to choose a linearly independent subset of these equations.[8]

Adaptive filtering for nonlinear systems has recently gained attention. Parlos et al. shows a neural net to constructively approximate the state equations.[9] The proposed algorithms in their paper make minimal assumptions regarding the underlying nonlinear dynamics and their noise statistics. Nonadaptive and adaptive state filtering algorithms are presented with both off-line and on-line learning stages. Good performance is shown for a number of test cases. Lho and Painter show an adaptive filter using fuzzy membership functions, where the fuzzy processing is driven by an inaccurate online estimate of signal-to-noise ratio for the signal being tracked.[10] Good results are shown for a simple tracking problem. Lee and

Alfriend show an adaptive scheme that can be used to estimate the process noise covariance for both the UF and the first-order divided difference filter.[11] The standard update approach requires proper selection of a window size to control the level of the variance update. The innovation of their paper is a procedure that automatically calculates the window size using a derivative-free numerical optimization technique. Good results are shown for satellite orbit determination applications.

A new approach is derived in Ref. 12 for adaptive filtering based on generalizing the standard multiple model adaptive estimation (MMAE) algorithm.[13] A MMAE algorithm uses a parallel bank of filters to provide multiple estimates, where each filter corresponds with a dependence on some unknowns, which can be the process noise covariance elements if desired. The state estimate is provided through a sum of each filter's estimate weighted by the likelihood of the unknown elements conditioned on the measurement sequence. The likelihood function gives the associated hypothesis that each filter is the correct one. Standard MMAE algorithms use only the current time measurement-minus-estimate residual to test the hypothesis. The approach in Ref. 12 is a generalization of Ref. 14, which uses the time correlation of the filter residuals to assign the likelihood for each of the modeled hypotheses. In particular, the spectral content of the residuals is used and only scalar measurements are assumed in Ref. 14. The authors also state that if multiple measurements are available, then a diagonal matrix can be used with elements given by the spectral content of each measurement residual, but this assumes that the cross-correlation terms are negligible. Also, the focus of their paper is on the detection of actuator failures with known control-input frequency content.

The new approach, called generalized multiple model adaptive estimation (GMMAE), is based on calculating the time-domain autocorrelation function, which is used to form the covariance of a generalized residual involving any number of backward time steps. This covariance matrix also includes the time correlated terms, thus providing a more rigorous approach. The unknown elements in our design are the parameters of the process noise covariance. Process noise covariance elements can be drawn from any sample distribution as long as the resulting covariance matrix remains positive semi-definite. A Hammersley quasi-random sequence[15] is chosen due to its well distributed pattern. The covariance elements are estimated using a weighted sum of the quasi-random elements, similar to the approach used for state estimation in PFs. An expression for the error-covariance of the estimates is also provided, which gives a bound on the process noise parameter estimates. In this paper the GMMAE approach is applied to the gyroless attitude estimation problem.

The organization of the remainder of this paper proceeds as follows. First, the standard equations of attitude kinematics are summarized, as they are the basis for the filter design. Then, the standard EKF equations are summarized, since this filter will be used in the simulations. Next, a review of the standard MMAE algorithm is given and the new adaptive approach is shown. Finally, simulation results are provided to demonstrate the effectiveness of GMMAE application to gyroless attitude estimation.

II. Attitude Kinematics

In this section a review of the basic kinematic equations for a three-axis stabilized spacecraft is provided. The attitude parameterization used is the quaternion,[16] which is a four-dimensional vector, defined as

$$\mathbf{q} \equiv \begin{bmatrix} \boldsymbol{\varrho} \\ q_4 \end{bmatrix} \tag{1}$$

with

$$\boldsymbol{\varrho} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \hat{\mathbf{e}} \sin(\vartheta/2) \tag{2a}$$

$$q_4 = \cos(\vartheta/2) \tag{2b}$$

where $\hat{\mathbf{e}}$ is the axis of rotation and ϑ is the angle of rotation.

The quaternion kinematics equation is given by

$$\dot{\mathbf{q}} = \frac{1}{2}\Xi(\mathbf{q})\boldsymbol{\omega} = \frac{1}{2}\Omega(\boldsymbol{\omega})\mathbf{q} \tag{3}$$

with

$$\Xi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_{3\times 3} + [\boldsymbol{\varrho}\times] \\ -\boldsymbol{\varrho}^T \end{bmatrix} \tag{4a}$$

$$\Omega(\boldsymbol{\omega}) \equiv \begin{bmatrix} -[\boldsymbol{\omega}\times] & \boldsymbol{\omega} \\ \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix} \tag{4b}$$

where $[\boldsymbol{a}\times]$ is the cross product matrix, defined by

$$[\boldsymbol{a}\times] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{5}$$

Since a four-dimensional vector is used to describe three dimensions, the quaternion components cannot be independent of each other. The quaternion satisfies a single constraint given by $\mathbf{q}^T\mathbf{q} = 1$, which is analogous to requiring that $\hat{\mathbf{e}}$ be a unit vector in the Euler axis/angle parameterization.[17]

A major advantage of using quaternions is that the kinematics equation is linear in the quaternion and is also free of singularities. Another advantage of quaternions is that successive rotations can be accomplished using quaternion multiplication. Here we adopt the convention of Lefferts, Markley, and Shuster[18] who multiply the quaternions in the same order as the attitude matrix multiplication (in contrast to the usual convention established by Hamilton[16]). Suppose we wish to perform a successive rotation. This can be written using

$$A(\mathbf{q}')A(\mathbf{q}) = A(\mathbf{q}' \otimes \mathbf{q}) \tag{6}$$

The composition of the quaternions is bilinear, with

$$\mathbf{q}' \otimes \mathbf{q} = \begin{bmatrix} \Psi(\mathbf{q}') & \mathbf{q}' \end{bmatrix} \mathbf{q} = \begin{bmatrix} \Xi(\mathbf{q}) & \mathbf{q} \end{bmatrix} \mathbf{q}' \tag{7}$$

Also, the inverse quaternion is defined by

$$\mathbf{q}^{-1} \equiv \begin{bmatrix} -\boldsymbol{\varrho} \\ q_4 \end{bmatrix} \tag{8}$$

Note that $\mathbf{q} \otimes \mathbf{q}^{-1} = [0\ \ 0\ \ 0\ \ 1]^T$, which is the identity quaternion. A computationally efficient algorithm to extract the quaternion from the attitude matrix is given in Ref. 19. A more thorough review of the attitude representations shown in this section, as well as others, can be found in the excellent survey paper by Shuster[17] and in the book by Kuipers.[20]

III. Extended Kalman Filter Equations

In this section the implementation equations for the EKF are shown. The truth equations are given by

$$\dot{\mathbf{q}} = \frac{1}{2}\Xi(\mathbf{q})\boldsymbol{\omega} \tag{9a}$$

$$\boldsymbol{\omega} = \boldsymbol{\omega}(t) \tag{9b}$$

where $\boldsymbol{\omega}(t)$ is the rate trajectory for the specific mission. In the numerical simulation to follow, brownian motion is used in order to provide an exact process noise covariance for comparison. The MMAE scheme uses several filters in parallel, but each of these filters uses the same dynamic model given by:

$$\dot{\hat{\mathbf{q}}} = \frac{1}{2}\Xi(\hat{\mathbf{q}})\hat{\boldsymbol{\omega}} \tag{10a}$$

$$\dot{\hat{\boldsymbol{\omega}}} = 0 \tag{10b}$$

We now derive the error equations, which are used in the EKF covariance propagation. The quaternion is linearized using a multiplicative approach.[18] First, an error quaternion is defined by

$$\boldsymbol{\delta}\mathbf{q} = \mathbf{q} \otimes \hat{\mathbf{q}}^{-1} \tag{11}$$

with $\boldsymbol{\delta}\mathbf{q} \equiv [\boldsymbol{\delta\varrho}^T \;\; \delta q_4]^T$, where the quaternion multiplication is defined by Eq. (7).

If the error quaternion is "small" then to within first order we have $\boldsymbol{\delta\varrho} \approx \boldsymbol{\delta\alpha}/2$ and $\delta q_4 \approx 1$, where $\boldsymbol{\delta\alpha}$ is a small error-angle rotation vector. Also, the quaternion inverse is defined by Eq. (8). The linearized model error-kinematics follow[18]

$$\boldsymbol{\delta\dot{\varrho}} = -[\hat{\boldsymbol{\omega}}\times]\boldsymbol{\delta\varrho} + \frac{1}{2}\boldsymbol{\delta\omega} \tag{12a}$$

$$\delta\dot{q}_4 = 0 \tag{12b}$$

where $\boldsymbol{\delta\omega} \equiv \boldsymbol{\omega} - \hat{\boldsymbol{\omega}}$. Note that the fourth error-quaternion component is constant. The first-order approximation, which assumes that the true quaternion is "close" to the estimated quaternion, gives $\delta q_4 \approx 1$. This allows us to reduce the order of the system in the EKF by one state. The linearization using Eq. (11) maintains quaternion normalization to within first-order if the estimated quaternion is "close" to the true quaternion, which is within the first-order approximation in the EKF.

Equation (12) uses the vector part of the quaternion. It is more convenient to use the attitude-error vector, $\boldsymbol{\delta\alpha} = 2\boldsymbol{\delta\varrho}$, which leads to

$$\boldsymbol{\delta\dot{\alpha}} = -[\hat{\boldsymbol{\omega}}\times]\boldsymbol{\delta\alpha} + \boldsymbol{\delta\omega} \tag{13}$$

Thus we will choose the state $\mathbf{x}$ and state-error vector $\Delta\mathbf{x}$ to be defined as

$$\mathbf{x} \equiv \begin{bmatrix} \mathbf{q} \\ \boldsymbol{\omega} \end{bmatrix}, \quad \Delta\mathbf{x} \equiv \begin{bmatrix} \boldsymbol{\delta\alpha} \\ \boldsymbol{\delta\omega} \end{bmatrix} \tag{14}$$

These being so defined, the error-dynamics used in the EKF propagation are given by

$$\Delta\dot{\mathbf{x}} = F\Delta\mathbf{x} + G\mathbf{w} \tag{15}$$

where

$$F = \begin{bmatrix} -[\hat{\boldsymbol{\omega}}(t)\times] & I_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{16}$$

$$G = \begin{bmatrix} 0_{3\times3} \\ I_{3\times3} \end{bmatrix} \tag{17}$$

while the process noise vector and covariance are given by

$$\mathbf{w} \equiv \boldsymbol{\eta}_\omega \tag{18}$$

and

$$Q = \sigma_\omega^2 I_{3\times3} \tag{19}$$

The goal of the multiple model approach is to determine the value of the tuning parameter σ_ω^2.

It is assumed that the quaternion serves as the direct measurement, which may be obtained from a quaternion-out star tracker along with its associated error covariance. Our next step involves the determination of the sensitivity matrix $H_k(\hat{\mathbf{x}}_k^-)$ used in the EKF. Discrete-time attitude observations for a single sensor are given by

$$\tilde{\mathbf{y}}_k = \mathbf{h}_k(\hat{\mathbf{x}}_k) + \mathbf{v}_k \tag{20a}$$

where the covariance of $\mathbf{v}_k$, denoted by R_k, is given by covariance of the attitude errors from the determine quaternion. The sensitivity matrix is given by

$$H_k(\hat{\mathbf{x}}_k^-) = \begin{bmatrix} I_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{21}$$

Note that the number of columns of $H_k(\hat{\mathbf{x}}_k^-)$ is six, which is the dimension of the reduced-order state.

The final part in the EKF involves the quaternion and bias updates. The error-state update follows

$$\Delta\hat{\tilde{\mathbf{x}}}_k^+ = K_k[\tilde{\mathbf{y}}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)] \tag{22}$$

Table 1. **Extended Kalman Filter for Attitude Estimation**

Initialize	$\hat{\mathbf{q}}(t_0) = \hat{\mathbf{q}}_0, \quad \boldsymbol{\omega}(t_0) = \boldsymbol{\omega}_0$ $P(t_0) = P_0$
Gain	$K_k = P_k^- H_k^T(\hat{\mathbf{x}}_k^-)[H_k(\hat{\mathbf{x}}_k^-)P_k^- H_k^T(\hat{\mathbf{x}}_k^-) + R_k]^{-1}$ $H_k(\hat{\mathbf{x}}_k^-) = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \end{bmatrix}$
Update	$P_k^+ = [I - K_k H_k(\hat{\mathbf{x}}_k^-)]P_k^-$ $\Delta\hat{\tilde{\mathbf{x}}}_k^+ = K_k[\tilde{\mathbf{y}}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)]$ $\Delta\hat{\tilde{\mathbf{x}}}_k^+ \equiv \begin{bmatrix} \boldsymbol{\delta}\hat{\boldsymbol{\alpha}}_k^{+T} & \boldsymbol{\delta}\hat{\boldsymbol{\omega}}_k^{+T} \end{bmatrix}^T$ $\mathbf{h}_k(\hat{\mathbf{x}}_k^-) = \hat{\mathbf{q}}_k^-$ $\hat{\mathbf{q}}_k^+ = \hat{\mathbf{q}}_k^- + \frac{1}{2}\Xi(\hat{\mathbf{q}}_k^-)\boldsymbol{\delta}\hat{\boldsymbol{\alpha}}_k^+$, re-normalize quaternion $\hat{\boldsymbol{\omega}}_k^+ = \hat{\boldsymbol{\omega}}_k^- + \boldsymbol{\delta}\hat{\boldsymbol{\omega}}_k^+$
Propagation	$\dot{\hat{\mathbf{q}}}(t) = \frac{1}{2}\Xi\left(\hat{\mathbf{q}}(t)\right)\hat{\boldsymbol{\omega}}(t)$ $\dot{\hat{\boldsymbol{\omega}}}(t) = 0$ $\dot{P}(t) = F(\hat{\mathbf{x}}(t),\, t)\, P(t) + P(t)\, F^T(\hat{\mathbf{x}}(t),\, t) + G(t)\, Q(t)\, G^T(t)$ $F(\hat{\mathbf{x}}(t),\, t) = \begin{bmatrix} -[\hat{\boldsymbol{\omega}}(t)\times] & I_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix}, \quad G(t) = \begin{bmatrix} 0_{3\times 3} \\ I_{3\times 3} \end{bmatrix}$

where $\Delta\hat{\tilde{\mathbf{x}}}_k^+ \equiv \begin{bmatrix} \boldsymbol{\delta}\hat{\boldsymbol{\alpha}}_k^{+T} & \boldsymbol{\delta}\hat{\boldsymbol{\omega}}_k^{+T} \end{bmatrix}^T$, $\tilde{\mathbf{y}}_k$ is the measurement output, and $\mathbf{h}_k(\hat{\mathbf{x}}_k^-)$ is the estimate output, given by

$$\mathbf{h}_k(\hat{\mathbf{x}}_k^-) = \hat{\mathbf{q}}_k^- \tag{23}$$

The rate update is simply given by

$$\hat{\boldsymbol{\omega}}_k^+ = \hat{\boldsymbol{\omega}}_k^- + \boldsymbol{\delta}\hat{\boldsymbol{\omega}}_k^+ \tag{24}$$

The quaternion update is more complicated. As previously mentioned the fourth component of $\boldsymbol{\delta}\mathbf{q}$ is nearly one. Therefore, to within first-order the quaternion update is given by

$$\hat{\mathbf{q}}_k^+ = \begin{bmatrix} \frac{1}{2}\boldsymbol{\delta}\hat{\boldsymbol{\alpha}}_k^+ \\ 1 \end{bmatrix} \otimes \hat{\mathbf{q}}_k^- \tag{25}$$

Note that the small angle approximation has been used to define the vector part of the error-quaternion. Using the quaternion multiplication rule of eqn. (7) in eqn. (25) gives

$$\hat{\mathbf{q}}_k^+ = \hat{\mathbf{q}}_k^- + \frac{1}{2}\Xi(\hat{\mathbf{q}}_k^-)\boldsymbol{\delta}\hat{\boldsymbol{\alpha}}_k^+ \tag{26}$$

This updated quaternion is a unit vector to within first-order; however, a brute-force normalization should be performed to insure $\hat{\mathbf{q}}_k^{+T}\hat{\mathbf{q}}_k^+ = 1$.

The attitude estimation algorithm is summarized in Table 1. The filter is first initialized with a known state and error-covariance matrix. Then, the Kalman gain is computed using the measurement-error covariance R and sensitivity matrix in eqn. (21). The state error-covariance follows the standard EKF update, while the error-state update is computed using eqn. (22). The rate and quaternion updates are now given by eqns. (24) and (26). Also, the updated quaternion is re-normalized by brute force. Finally, the estimated angular velocity and quaternion are used to propagate the model in eqn. (10) and standard error-covariance in the EKF.

IV. Multiple Model Adaptive Estimation

In this section a review of MMAE is shown. More details can be found in Refs. 21 and 22. Multiple model adaptive estimation is a recursive estimator that uses a bank of filters that depend on some unknown parameters. In our case these parameters are the process noise covariance, denoted by the vector $\mathbf{p}$, which is assumed to be constant (at least throughout the interval of adaptation). Note that we do not necessarily need to make the stationary assumption for the state and/or output processes though, i.e. time varying state and output matrices can be used. A set of distributed elements is generated from some known pdf of $\mathbf{p}$, denoted by $p(\mathbf{p})$, to give $\{\mathbf{p}^{(\ell)};\, \ell = 1, \ldots, M\}$. The goal of the estimation process is to determine the conditional pdf of the ℓ^{th} element $\mathbf{p}^{(\ell)}$ given the current-time measurement $\tilde{\mathbf{y}}_k$.

Application of Bayes' rule yields

$$p\,(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_k) = \frac{p\,(\tilde{\mathbf{Y}}_k|\mathbf{p}^{(\ell)})\,p\,(\mathbf{p}^{(\ell)})}{\sum_{j=1}^{M} p\,(\tilde{\mathbf{Y}}_k|\mathbf{p}^{(j)})\,p\,(\mathbf{p}^{(j)})} \tag{27}$$

where $\tilde{\mathbf{Y}}_k$ denotes the sequence $\{\tilde{\mathbf{y}}_0, \tilde{\mathbf{y}}_1, \ldots, \tilde{\mathbf{y}}_k\}$. The *a posteriori* probabilities can be computed through[23]

$$\begin{aligned} p\,(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_k) &= \frac{p\,(\tilde{\mathbf{y}}_k,\, \mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_{k-1})}{p\,(\tilde{\mathbf{y}}_k|\tilde{\mathbf{Y}}_{k-1})} \\ &= \frac{p\,(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)})\,p\,(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_{k-1})}{\sum_{j=1}^{M} \left[p\,(\tilde{\mathbf{Y}}_k|\hat{\mathbf{x}}_k^{-(j)})\,p\,(\mathbf{p}^{(j)}|\tilde{\mathbf{Y}}_{k-1})\right]} \end{aligned} \tag{28}$$

since $p\,(\tilde{\mathbf{y}}_k, |\tilde{\mathbf{Y}}_{k-1},\, \mathbf{p}^{(\ell)})$ is given by $p\,(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)})$ in the Kalman recursion. Note that the denominator of Eq. (28) is just a normalizing factor to ensure that $p\,(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_k)$ is a pdf. The recursion formula can now be cast into a set of defined weights $\varpi_k^{(\ell)}$, so that

$$\begin{aligned} \varpi_k^{(\ell)} &= \varpi_{k-1}^{(\ell)} p\,(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)}) \\ \varpi_k^{(\ell)} &\leftarrow \frac{\varpi_k^{(\ell)}}{\sum_{j=1}^{M} \varpi_k^{(j)}} \end{aligned} \tag{29}$$

where $\varpi_k^{(\ell)} \equiv p\,(\mathbf{p}^{(\ell)}|\tilde{\mathbf{y}}_k)$. The weights at time t_0 are initialized to $\varpi_0^{(\ell)} = 1/M$ for $\ell = 1, 2, \ldots, M$. The convergence properties of MMAE are shown in Ref. 23, which assumes ergodicity in the proof. The ergodicity assumptions can be relaxed to asymptotic stationarity and other assumptions are even possible for non-stationary situations.[24]

The conditional mean estimate is the weighted sum of the parallel filter estimates:

$$\hat{\mathbf{x}}_k^- = \sum_{j=1}^{M} \varpi_k^{(j)} \hat{\mathbf{x}}_k^{-(j)} \tag{30}$$

Also, the error covariance of the state estimate can be computed using

$$P_k^- = \sum_{j=1}^{M} \varpi_k^{(j)} \left[\left(\hat{\mathbf{x}}_k^{-(j)} - \hat{\mathbf{x}}_k^-\right) \left(\hat{\mathbf{x}}_k^{-(j)} - \hat{\mathbf{x}}_k^-\right)^T + P_k^{-(\ell)} \right] \tag{31}$$

The specific estimate for $\mathbf{p}$ at time t_k, denoted by $\hat{\mathbf{p}}_k$, and error covariance, denoted by $\mathcal{P}_k$, are given by

$$\hat{\mathbf{p}}_k = \sum_{j=1}^{M} \varpi_k^{(j)} \mathbf{p}^{(j)} \tag{32a}$$

$$\mathcal{P}_k = \sum_{j=1}^{M} \varpi_k^{(j)} \left(\mathbf{p}^{(j)} - \hat{\mathbf{p}}_k\right) \left(\mathbf{p}^{(j)} - \hat{\mathbf{p}}_k\right)^T \tag{32b}$$

Equation (32b) can be used to define 3σ bounds on the estimate $\hat{\mathbf{p}}_k$.

The likelihood function associated with the output residual is given by

$$p\left(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)}\right) = \frac{1}{\left\{\det[2\pi\left(H_k P_k^- H_k^T + R_k\right)]\right\}^{1/2}} \exp\left[-\frac{1}{2}\mathbf{e}_k^{(\ell)T}(H_k P_k^- H_k^T + R_k)^{-1}\mathbf{e}_k^{(\ell)}\right] \tag{33}$$

where $\mathbf{e}_k^{(\ell)} \equiv \tilde{\mathbf{y}}_k - \hat{\mathbf{y}}_k^{-(\ell)}$.

V. Generalized Multiple Model Adaptive Estimation

In this section a review of generalized multiple model adaptive estimation (GMMAE)is provided. GMMAE is an extension of the MMAE approach wherein a window of previous-time data is used via the autocorrelation matrix to perform the adaptive update. Further details and the complete derivation can be found in Ref. 12.

First, a residual vector is defined to concatenate residuals from the current and i previous time steps:

$$\boldsymbol{\epsilon}_i \equiv \begin{bmatrix} \mathbf{e}_k \\ \mathbf{e}_{k-1} \\ \vdots \\ \mathbf{e}_{k-i} \end{bmatrix} \tag{34}$$

The likelihood function associated with $\boldsymbol{\epsilon}_i$ is given by

$$L_i = \frac{1}{[\det(2\pi\,\mathcal{C}_i)]^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_i^T \mathcal{C}_i^{-1} \boldsymbol{\epsilon}_i\right) \tag{35}$$

where $\mathcal{C}_i \equiv E\{\boldsymbol{\epsilon}_i \boldsymbol{\epsilon}_i^T\}$ is given by

$$\mathcal{C}_i = \begin{bmatrix} C_{k,\,0} & C_{k,\,1} & C_{k,\,2} & \cdots & C_{k,\,i} \\ C_{k,\,1}^T & C_{k-1,\,0} & C_{k-1,\,1} & \cdots & C_{k-1,\,i-1} \\ C_{k,\,2}^T & C_{k-1,\,2}^T & C_{k-2,\,0} & \cdots & C_{k-2,\,i-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{k,\,i}^T & C_{k-1,\,i-1}^T & C_{k-2,\,i-2}^T & \cdots & C_{k-i,\,0} \end{bmatrix} \tag{36}$$

with

$$C_{k,\,i} = \begin{cases} H_k P_k^- H_k^T + R_k & i = 0 \\ H_k \Phi_{k-1}(P_{k-1}^- H_{k-1}^T - K_{k-1} C_{k-1,\,0}) & i = 1 \\ H_k \left[\prod_{j=1}^{i-1} \Phi_{k-j}\left(I - K_{k-j}\, H_{k-j}\right)\right] & \\ \times \Phi_{k-i}\left(P_{k-i}^-\, H_{k-i}^T - K_{k-i}\, C_{k-i,\,0}\right) & i > 1 \end{cases} \tag{37}$$

where

$$C_{k-i,\,0} \equiv H_{k-i}\, P_{k-i}^-\, H_{k-i}^T + R_{k-i} \tag{38}$$

When $i = 0$ the likelihood function reduces down to that shown in Eq. (33). This likelihood is widely used in MMAE algorithms,[13, 21] but ignores correlations between different measurement times. However, it is simpler to evaluate than the general likelihood function in Eq. (35) since no storage of data or system matrices is required.

The GMMAE update law, though similar to that of the MMAE algorithm, is based on carrying Eq. (28) i steps back:

$$\begin{aligned} \varpi_k^{(\ell)} &= \varpi_{k-1}^{(\ell)} L_i^{(\ell)} \\ \varpi_k^{(\ell)} &\leftarrow \frac{\varpi_k^{(\ell)}}{\sum_{j=1}^{M} \varpi_k^{(j)}} \end{aligned} \tag{39}$$

with

$$L_i^{(\ell)} = \frac{1}{\left[\det\left(2\pi\,\mathcal{C}_i^{(\ell)}\right)\right]^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{\epsilon}_i^{(\ell)T}\left(\mathcal{C}_i^{(\ell)}\right)^{-1}\boldsymbol{\epsilon}_i^{(\ell)}\right] \tag{40}$$

where $\boldsymbol{\epsilon}_i^{(\ell)}$ is defined as $\boldsymbol{\epsilon}_i^{(\ell)} \equiv [\mathbf{e}_k^{(\ell)T} \;\; \mathbf{e}_{k-1}^{(\ell)T} \;\; \cdots \;\; \mathbf{e}_{k-i}^{(\ell)T}]^T$. The matrix $\mathcal{C}_i^{(\ell)}$ is given by Eqs. (37) and (36) evaluated at the ℓ^{th} covariance and the *optimal* Kalman gain. Unfortunately, the optimal gain is a function of the actual covariance Q_k, which is not known. Specifically, if K_k is substituted into Eq. (37), then for $i \geq 1$ the correlated terms $C_{k,\,i}$ will always be zero. However, assuming that the measurement noise is small compared to the signal so that the Gaussian nature of the measurement residual is maintained, estimates for $C_{k,\,i}^{(\ell)}$ can be given by

$$\hat{C}_{k,\,i}^{(\ell)} = \begin{cases} H_k(\hat{\mathbf{x}}_k^{-(\ell)}) P_k^{-(\ell)} H_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + R_k & i = 0 \\[2ex] H_k(\hat{\mathbf{x}}_k^{-(\ell)}) \Phi_{k-1}(\hat{\mathbf{x}}_{k-1}^{-(\ell)}) & \\ \times [P_{k-1}^{-(\ell)} H_{k-1}^T(\hat{\mathbf{x}}_{k-1}^{-(\ell)}) - \hat{K}_{k-1} C_{k-1,\,0}^{(\ell)}] & i = 1 \\[2ex] H_k(\hat{\mathbf{x}}_k^{-(\ell)}) \{\prod_{j=1}^{i-1} \Phi_{k-j}(\hat{\mathbf{x}}_{k-j}^{-(\ell)}) & \\ \times [I - \hat{K}_{k-j}\, H_{k-j}(\hat{\mathbf{x}}_{k-j}^{-(\ell)})]\} & \\ \times \Phi_{k-i}(\hat{\mathbf{x}}_{k-i}^{-(\ell)})\, [P_{k-i}^{-(\ell)}\, H_{k-i}^T(\hat{\mathbf{x}}_{k-i}^{-(\ell)}) & \\ - \hat{K}_{k-i}\, C_{k-i,\,0}^{(\ell)}] & i > 1 \end{cases} \tag{41}$$

where

$$C_{k-i,\,0}^{(\ell)} \equiv H_{k-i}(\hat{\mathbf{x}}_{k-i}^{-(\ell)})\, P_{k-i}^{-(\ell)}\, H_{k-i}^T(\hat{\mathbf{x}}_{k-i}^{-(\ell)}) + R_{k-i} \tag{42}$$

The covariance matrix $P_k^{-(\ell)}$ is computed using

$$P_{k+1}^{-(\ell)} = \Phi_k(\hat{\mathbf{x}}_k^{-(\ell)}) P_k^{+(\ell)} \Phi_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + Q^{(\ell)} \tag{43a}$$

$$P_k^{+(\ell)} = \left[I - K_k^{(\ell)} H_k(\hat{\mathbf{x}}_k^{-(\ell)})\right] P_k^{-(\ell)} \tag{43b}$$

$$K_k^{(\ell)} = P_k^{-(\ell)} H_k^T \left[H_k(\hat{\mathbf{x}}_k^{-(\ell)}) P_k^{-(\ell)} H_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + R_k\right]^{-1} \tag{43c}$$

where $Q^{(\ell)}$ is computed using $\mathbf{p}^{(\ell)}$. The estimate of the optimal gain is computed using

$$\hat{K}_k = \hat{P}_k^- H_k^T(\hat{\mathbf{x}}_k^-) \left[H_k(\hat{\mathbf{x}}_k^-) \hat{P}_k^- H_k^T(\hat{\mathbf{x}}_k^-) + R_k\right]^{-1} \tag{44}$$

with

$$\hat{P}_{k+1}^- = \Phi_k(\hat{\mathbf{x}}_k^-) \hat{P}_k^+ \Phi_k^T(\hat{\mathbf{x}}_k^-) + \hat{Q}_k \tag{45a}$$

$$\hat{P}_k^+ = \left[I - \hat{K}_k H_k(\hat{\mathbf{x}}_k^-)\right] \hat{P}_k^- \tag{45b}$$

where $\hat{Q}_k$ is computed using $\hat{\mathbf{p}}_k$.

Using the current measurement, $\tilde{\mathbf{y}}_k$, along with the ℓ^{th} element, $\mathbf{p}^{(\ell)}$, $1 \le \ell \le M$, a bank of filters are executed. For each filter the state estimates, $\hat{\mathbf{x}}_k^{-(\ell)}$, and measurements are used to form the residual, $\boldsymbol{\epsilon}_i^{(\ell)}$, going back i steps. The filter error covariance, $P_k^{-(\ell)}$, and state matrices, $\Phi_k^{-(\ell)}$ and $H_k^{-(\ell)}$, evaluated at the current estimates are used to update the estimate of the autocorrelation, denoted by $\hat{\mathcal{C}}_i^{(\ell)}$. Note that at each new measurement time, all elements of $\hat{\mathcal{C}}_i^{(\ell)}$ need to be recalculated since a new estimate $\hat{\mathbf{p}}_k$ is provided, which is used to compute an estimate of the optimal gain. Unfortunately, this can significantly increase the computational costs. The diagonal elements do not need to be recomputed though, since they are not a function of the optimal gain. The residuals and autocorrelations are then used to evaluate the likelihood functions $L_i^{(\ell)}$. These functions are used to update the weights, which gives the estimate $\hat{\mathbf{p}}_k$ using Eq. (32a).

There are many possibilities for the chosen distribution of the process noise covariance parameters. A simple approach is to a assume a uniform distribution. We instead choose a Hammersley quasi-random sequence[15] due to its well distributed pattern. In low dimensions, the multidimensional Hammersley sequence quickly "fills up" the space in a well-distributed pattern. However, for very high dimensions, the initial elements of the Hammersley sequence can be very poorly distributed. Only when the number of sequence elements is large enough relative to the spatial dimension is the sequence properly behaved. This isn't much of a concern for the process noise covariance adaption problem since the dimension of the elements will be much larger than the dimension of the unknown process noise parameters. Remedies for this problem are given in Ref. 25 if needed.

VI. Simulation Results

To verify the effectiveness of the GMMAE scheme, a simulation is developed. A combined quaternion from two trackers is assumed for the measurement. In order to generate synthetic measurements the following model is used:

$$\tilde{\mathbf{q}} = \begin{bmatrix} 0.5\mathbf{v} \\ 1 \end{bmatrix} \otimes \mathbf{q}$$

where $\tilde{\mathbf{q}}$ is the quaternion measurement, $\mathbf{q}$ is the truth, and $\mathbf{v}$ is the measurement noise, which is assumed to be a zero-mean Gaussian noise process with covariance given by $0.001 I_{3\times 3}$ deg^2. Note, the measured quaternion is normalized to within first-order, but a brute-force normalization is still taken to ensure a normalized measurement. The quaternion measurements are sampled at 0.1 Hz.

Both MMAE and GMMAE simulations are performed in order to provide a benchmark for the performance of the latter. Time plots of the adaptively estimated process noise parameter are shown for both filters in Figure 1. It is apparent that, though both filters converge to the correct value of $\sigma_\omega^2 = 2 \times 10^{-7}$, the GMMAE scheme does so far more rapidly. Plots of the state errors with 3σ bounds shown in Figure 2 demonstrate that the two schemes provide comparably accurate state estimates. However, within the larger framework of MMAE application to other problems, fault detection for example,[26] it desirable to have faster convergence of the process noise covariance estimates. This is the main benefit of the GMMAE solution. Considering these results, it is clear that the GMMAE scheme is an attractive solution for gyroless attitude estimation due to its excellent filter performance, and improved convergence speed over the MMAE approach.

VII. Conclusions

In this paper the GMMAE algorithm was applied to the problem of spacecraft attitude estimation without rate gyros. Simulation results indicate that the GMMAE approach presents a clear improvement over the traditional MMAE in the speed of convergence towards accurate process noise covariance parameters. Both algorithms offer a clear advantage over the traditional EKF in that the process noise covariance is determined by the filter itself. As such, these approaches will clearly outperform the EKF unless one happens to be fortunate enough to determine an accurate process noise covariance exactly.

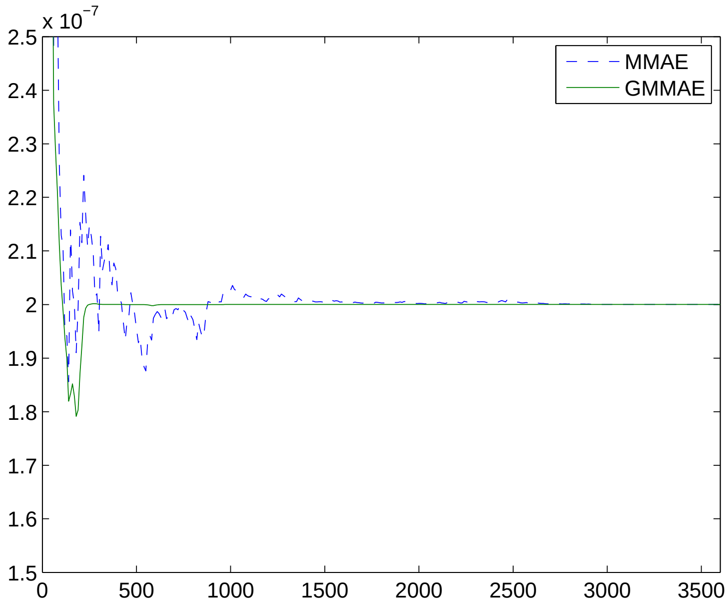
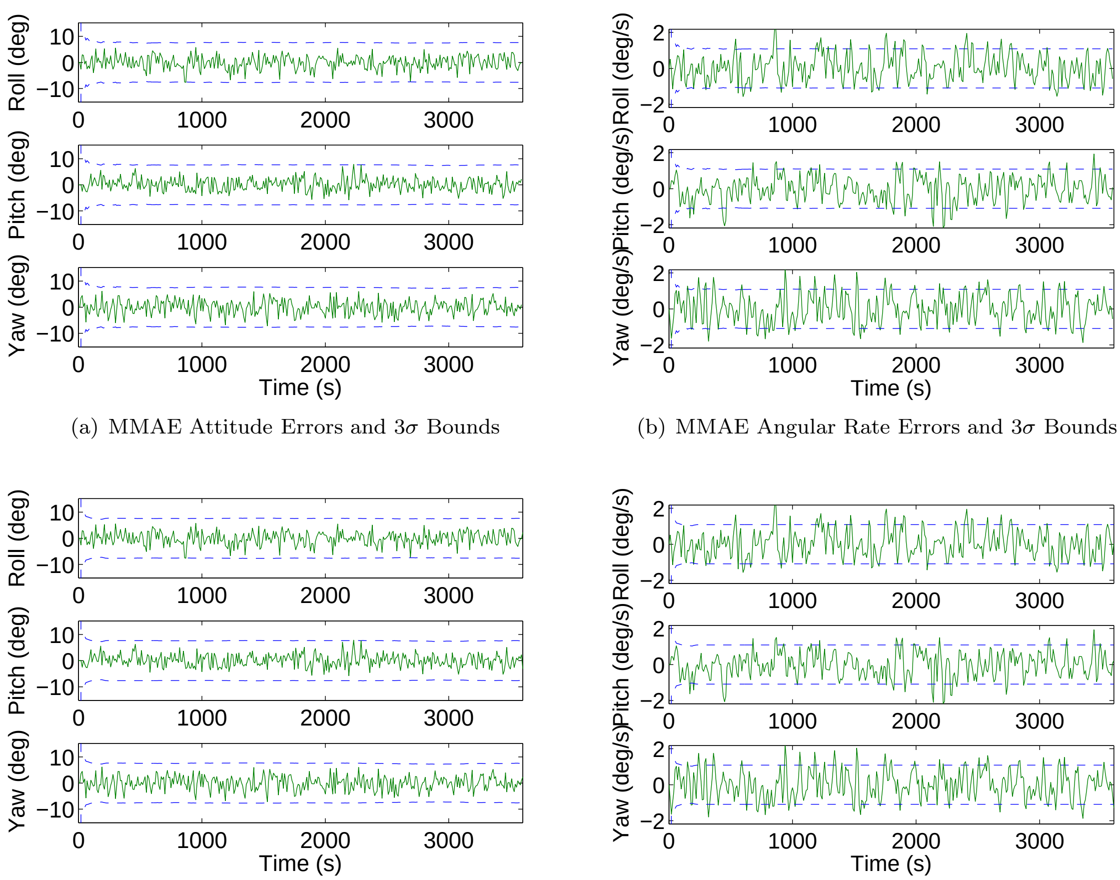


Figure 1. Parameter Estimation Comparison



(a) MMAE Attitude Errors and 3σ Bounds

(b) MMAE Angular Rate Errors and 3σ Bounds

(c) GMMAE Attitude Errors and 3σ Bounds

(d) GMMAE Angular Rate Errors and 3σ Bounds

Figure 2. MMAE vs GMMAE Filter Results

References

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

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

[3]Gordon, N. J., Salmond, D. J., and Smith, A. F. M., "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE Proceedings, Part F - Communications, Radar and Signal Processing*, Vol. 140, No. 2, Seattle, WA, April 1993, pp. 107–113.

[4]Arulampalam, M. S., Maskell, S., Gordon, N., 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. 174–185.

[5]Pearson, K., "Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder," *Biometrika*, Vol. 4, 1905, pp. 169–212.

[6]Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, New York, NY, 2001, pp. 232–245.

[7]Odelson, B. J., Rajamani, M. R., and Rawlings, J. B., "A New Autocovariance Least-Squares Method for Estimating Noise Covariances," Tech. Rep. 2003-04, Texas-Wisconsin Modeling and Control Consortium, Sept. 2003.

[8]Mehra, R. K., "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control*, Vol. AC-15, No. 2, April 1970, pp. 175–184.

[9]Parlos, A. G., Menon, S. K., and Atiya, A. F., "An Algorithmic Approach to Adaptive State Filtering Using Recurrent Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 12, No. 6, Nov. 2001, pp. 1411–1432.

[10]Lho, Y. H. and Painter, J. H., "A Fuzzy-Tuned Adaptive Kalman Filter," *Proceedings of the Third International IEEE Conference on Industrial Fuzzy Control and Intelligent Systems*, Houston, TX, Dec. 1993, pp. 144–148.

[11]Lee, D.-J. and Alfriend, K. T., "Adaptive Sigma Point Filtering for State and Parameter Estimation," *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Providence, RI, Aug. 2004, AIAA-2004-5101.

[12]Crassidis, J. L. and Cheng, Y., "Generalized Multiple-Model Adaptive Estimation Using an Autocorrelation Approach," *Proceedings of the 9th International Conference on Information Fusion*, Florence, Italy, July 2006, paper 223.

[13]Magill, D. T., "Optimal Adaptive Estimation of Sampled Stochastic Processes," *IEEE Transactions on Automatic Control*, Vol. AC-10, No. 4, Oct. 1965, pp. 434–439.

[14]Hanlon, P. D. and Maybeck, P. S., "Multiple-Model Adaptive Estimation Using a Residual Correlation Kalman Filter Bank," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-36, No. 2, April 2000, pp. 393–406.

[15]Hammersley, J. M., "Monte Carlo Methods for Solving Multivariable Problems," *Annals of the New York Academy of Sciences*, Vol. 86, 1960, pp. 844–874.

[16]Hamilton, W. R., *Elements of Quaternions*, Longmans, Green and Co., London, England, 1866.

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

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

[19]Shepperd, S. W., "Quaternion from Rotation Matrix," *Journal of Guidance and Control*, Vol. 1, No. 3, May-June 1978, pp. 223–224.

[20]Kuipers, J. B., *Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace, and Virtual Reality*, Princeton University Press, Princeton, NJ, 1999.

[21]Brown, R. G. and Hwang, P. Y. C., *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley & Sons, New York, NY, 3rd ed., 1997, pp. 353–361.

[22]Stengel, R. F., *Optimal Control and Estimation*, Dover Publications, New York, NY, 1994, pp. 402–407.

[23]Anderson, B. D. O. and Moore, J. B., *Optimal Filtering*, chap. 10.1, Dover Publications, Mineola, NY, 2005.

[24]Anderson, B. D. O., Moore, J. B., and Hawkes, R. M., "Model Approximations via Prediction Error Identification," *Automatica*, Vol. 14, No. 6, Nov. 1978, pp. 615–622.

[25]Kocis, L. and Whiten, W. J., "Computational Investigations of Low-Discrepancy Sequences," *ACM Transactions on Mathematical Software*, Vol. 23, No. 2, June 1997, pp. 266–294.

[26]Hoffmann, G. M., Gorinevsky, D., Mah, R. W., Tomlin, C. J., and Mitchellk, J. D., "Fault Tolerant Relative Navigation Using Inertial and Relative Sensors," *AIAA Guidance, Navigation and Control Conference*, Hilton Head, SC, Aug. 2001, AIAA 2007-6789.