On-Orbit Gyro Calibration for Operationally Responsive Space Systems

Richard Linares,* John L. Crassidis† and Puneet Singla‡
University at Buffalo, State University of New York, Amherst, NY 14260-4400

This paper presents a method for accurate on-orbit calibration of rate-integrating three-axis gyroscopes. On-orbit calibration is required for operationally responsive space systems in order to deal with rapidly changing technology and support needs. However, a trade-off usually occurs between accuracy and speed of convergence of the to-be-determined calibration parameters. In order to achieve a fully autonomous and accurate calibration approach, a two-step process is proposed. The first step determines an initial estimate of the gyro calibration parameters using an angular velocity estimate determined from a star tracker. This angular velocity estimate is provided without the need to identify stars, thereby bypassing the associated adverse issues that typically arise in star tracker attitude determination systems. The second step uses the initial gyro calibration estimates in a generalized multiple-model adaptive estimation approach, which combines outputs from three filters to determine estimates that are better than each filter alone. Simulation results are shown to assess the performance of the proposed on-orbit calibration approach.

I. Introduction

Pre-built modular components that can be rearranged to meet diverse mission requirements have been suggested as a possible approach to achieve operationally responsive space (ORS) systems.[1] This "plug-and-play" concept is similar to that of the personal computer, wherein an open system architecture is used and components can be assembled in an application specific basis. Basically this approach uses spacecraft bus configurations defined on a mission categorization basis where mission components can be selected from an inventory of standard plug-and-play components or subsystems to meet mission objectives. The rapid assembly of components may mean that the system may lack precise sensor calibration. The guidance, navigation and control subsystem in particular requires precise calibration for proper operation. Typically calibration is usually performed prior to launch, but an accelerated launch schedule may require forgoing extensive calibration procedures. In terms of time required to prepare a mission it is necessary to perform calibration on-orbit, which is needed to accommodate fully responsive ORS systems.

This work focuses on filtering methods for on-orbit calibration for rate-integrating gyro sensors. Of all the attitude sensors, gyros are likely the most crucial sensor system. Accurate angular rate information is often required for precise attitude maneuvers and is also used for filtering of noisy measurements. On-orbit sequential filtering[2] approaches are preferred over batch[3] approaches because they provide realtime estimation of the calibration parameters. The traditional attitude on-orbit calibration approach involves using a 6-state extended Kalman filter (EKF) to estimate the current attitude and gyro biases simultaneously. This filter assumes unvarying alignment of the sensors involved in attitude estimation. However, sensor misalignment is inevitable and would contribute to unreliable attitude estimates.[4] More stringent attitude pointing accuracy requires misalignments to be estimated and implemented into the attitude estimator. It has been noted in various papers[5–7] of the importance of proper calibration for use in fault detection or rate derivation. Without on-orbit sequential calibration, the sensors have to be re-calibrated on the ground from time to time to improve residual characteristics.[6]

*Graduate Student, Department of Mechanical & Aerospace Engineering. Email: linares2@buffalo.edu. Student Member AIAA.

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

‡Assistant Professor, Department of Mechanical & Aerospace Engineering. Email: psingla@buffalo.edu, Senior Member AIAA.

Misalignment calibration is usually performed prior to launch. However, launch shock often makes this pre-launch calibration somewhat irrelevant. Thus on-orbit alignment calibration is needed to accommodate these unanticipated changes in alignment before the spacecraft can begin its nominal mission mode. The ill effects of attitude estimation using misaligned sensors are shown in Ref. 2. On-orbit realtime or sequential filtering is preferred since misalignment parameters could drift, either due to in-space disturbances (for example solar wind, aerodynamics, persistence thermal shock) or unattended changes in the attitude sensor's relative orientation, especially those with moving mechanisms. A sequential filter constantly outputs the best estimate of the calibration parameters and can tolerate parameter drifts in realtime with minimal ground crew intervention or mission disruption.

On-orbit calibration schemes would enhance the degree of autonomy of the attitude determination and control system to reduce ground crew intervention for an extended period of time. Specifically, autonomy refers to a mode of ground system operation in which manual human actions are not required to accomplish desired functions. Autonomous systems are self-acting, self-regulating systems on the spacecraft wherein functions are delegated to the spacecraft systems. In some cases it may be possible to perform sensor calibration without attitude knowledge, such as three-axis magnetometers, which is often preferred because attitude estimation is decoupled from the calibration. However, in other cases, such as typical gyro calibration, this is not possible.

Realtime gyro calibration is typically accomplished by appending filter states with calibration parameters. Good initialization of the parameters is often required to achieve proper convergence. The most basic approach involves estimating attitude and gyro bias calibration parameters,[8] which involves a 6-state filter. This filter typically works under low angular velocities but cannot correct for gyro scale factor and misalignment error effects during high angular velocities. Under such operating conditions, gyro scale factor and misalignment errors strongly degrade the 6-state filter performance.[9] Thus, higher-state filters may be required to estimate for scale factors and misalignments. However, the performance may also be degraded when using higher-state filters due to information dilution. Reference 10 shows that each time a calibration parameter state is added to the filter, the estimation of that parameter uses up a certain amount of the limited information available by the observations. Therefore, adding more states can decrease the amount of information which was previously available to each of the original states.

This paper presents an approach to provide good initialization of the gyro calibration parameters while mitigating the information dilution problem. The initialization is performed using an angular velocity estimate from only star tracker body observations.[11] This approach has several advantages over other existing star tracker determined angular velocity methods, including, it does not require stars to be identified to provide an angular velocity estimate and it is attitude independent. The information dilution problem is mitigated by employing a generalized multiple-model adaptive estimation (GMMAE) approach,[12] which combines three filters to provide overall calibration parameter estimates. In Ref. 9 a standard MMAE approach was employed using the three-filter concept. Standard MMAE approaches use only the current time measurement-minus-estimate residual in the algorithm. The GMMAE approach 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, thereby providing an increased convergence rate over the standard MMAE approach.[13] Taken together, the initialization procedure and GMMAE approach provide a robust approach for on-orbit gyro calibration.

The organization of this paper is as follows. First, the general gyro calibration procedure is discussed. Then, the gyro and star tracker measurement models are summarized, followed by the derivation of the angular velocity estimation approach directly from star body measurements. The initialization procedure for the gyro calibration parameters is then derived. Next, reviews of the extended Kalman filter for attitude estimation and the generalized multiple-model adaptive estimation approach are shown. Simulation results are then shown followed by a persistency of excitation analysis.

II. Gyro Calibration Procedure

The on-orbit gyro calibration approach in this work can be broken up into an initialization stage and an operational mode. Figure 1 provides the general framework for the on-orbit gyro calibration approach. It is well known that the observability of gyro parameters is an issue, which is especially crucial during the initialization stage when calibration parameters are not well known. To overcome this problem the notion of decoupling is applied for the estimation of the various calibration parameters as much as possible.

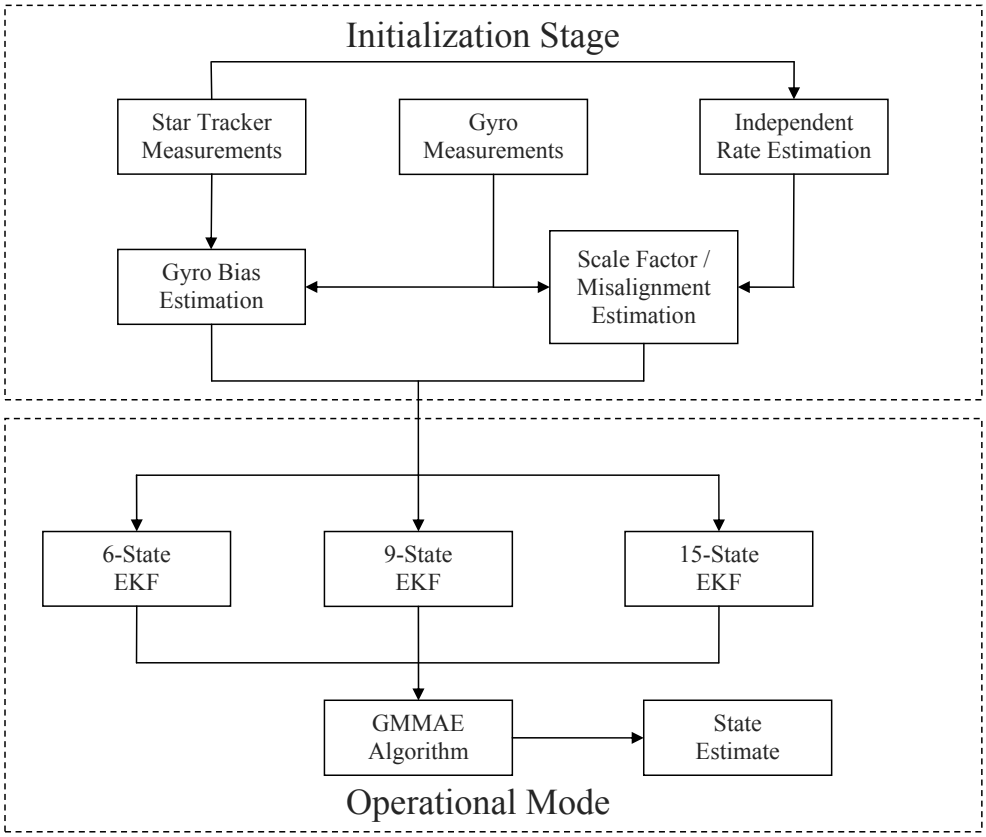


Figure 1. Gyro Calibration Process

The initialization stage can be decoupled into two steps: 1) scale factor / misalignment estimation, and 2) bias estimation. Using a difference of the gyro measurements at two times, a crude estimate of the scale factors and misalignment estimates is first determined. The bias estimate can be determined directly from a difference of star tracker body observations, providing an attitude-independent estimate that doesn't require star identification.

In the operational mode a multiple-model filtering approach is implemented. In particular, 6-state, 9-state and 15-state EKFs are employed to estimate the attitude, gyro biases, scale factors and misalignments. The 6-state filter estimates attitude and gyro biases, the 9-state filter estimates attitude, gyro biases and scale factors, and the 15-state filter estimates attitude, gyro biases, scale factors and misalignments. Multiple filters are used to overcome the problem of information dilution. The estimate of each filter is combined the GMMAE approach. The use of the different order filters effectively decouples the calibration parameter estimation process during the operational mode.

III. Gyro and Star Tracker Measurement Models

This section describes the gyro and star tracker measurement models. More details are provided in the references herein.

A. Gyro Measurement Model

A common sensor that measures the angular rate is a rate-integrating gyro. For this sensor, a widely used three-axis continuous-time model is given by[14]

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \boldsymbol{\beta} + \boldsymbol{\eta}_v \tag{1a}$$

$$\dot{\boldsymbol{\beta}} = \boldsymbol{\eta}_u \tag{1b}$$

where $\boldsymbol{\omega}$ is the true rate, $\tilde{\boldsymbol{\omega}}$ is the measured rate, $\boldsymbol{\beta}$ is the drift, and $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are independent zero-mean Gaussian white-noise processes with

$$E\left\{\boldsymbol{\eta}_v(t)\boldsymbol{\eta}_v^T(\tau)\right\} = \sigma_v^2\delta(t-\tau)I_{3\times3} \tag{2a}$$

$$E\left\{\boldsymbol{\eta}_u(t)\boldsymbol{\eta}_u^T(\tau)\right\} = \sigma_u^2\delta(t-\tau)I_{3\times3} \tag{2b}$$

where $E\{\cdot\}$ denotes expectation and $\delta(t-\tau)$ is the Dirac delta function. A more general gyro model includes scale factors and misalignments, which can also be estimated in realtime.[2,3] The general gyro model including scaling factors and misalignments is given by

$$\tilde{\boldsymbol{\omega}} = (I_{3\times3} + S)\,\boldsymbol{\omega} + \boldsymbol{\beta} + \boldsymbol{\eta}_v \tag{3}$$

with S denoting the matrix of scale factors $\mathbf{s} \equiv [s_1 \;\; s_2 \;\; s_3]^T$, and misalignments $\mathbf{k}_U \equiv [k_{U1} \;\; k_{U2} \;\; k_{U3}]^T$ and $\mathbf{k}_L \equiv [k_{L1} \;\; k_{L2} \;\; k_{L3}]^T$:

$$S \equiv \begin{bmatrix} s_1 & k_{U1} & k_{U2} \\ k_{L1} & s_2 & k_{U3} \\ k_{L2} & k_{L3} & s_3 \end{bmatrix} \tag{4}$$

Note that $\boldsymbol{\beta}$ and $\boldsymbol{\eta}_v$ are now expressed in the true (misaligned) gyro coordinate system. The models for the scale factors and misalignments are given by

$$\dot{\mathbf{s}} = \boldsymbol{\eta}_s, \quad E\left\{\boldsymbol{\eta}_s(t)\boldsymbol{\eta}_s^T(\tau)\right\} = \sigma_s^2\delta(t-\tau)I_{3\times3} \tag{5a}$$

$$\dot{\mathbf{k}}_U = \boldsymbol{\eta}_U, \quad E\left\{\boldsymbol{\eta}_U(t)\boldsymbol{\eta}_U^T(\tau)\right\} = \sigma_U^2\delta(t-\tau)I_{3\times3} \tag{5b}$$

$$\dot{\mathbf{k}}_L = \boldsymbol{\eta}_L, \quad E\left\{\boldsymbol{\eta}_L(t)\boldsymbol{\eta}_L^T(\tau)\right\} = \sigma_L^2\delta(t-\tau)I_{3\times3} \tag{5c}$$

For most spacecraft applications $\sigma_s = \sigma_U = \sigma_L = 0$.

For small errors and misalignments the true angular rate can be well approximated by using $(I_{3\times3} + S)^{-1} \approx I_{3\times3} - S$ so that

$$\boldsymbol{\omega} = \tilde{\boldsymbol{\omega}} - \boldsymbol{\beta} - \Omega\,\boldsymbol{\kappa} - (I_{3\times3} - S)\,\boldsymbol{\eta}_v \tag{6}$$

where $\boldsymbol{\kappa} \equiv [s_1 \;\; k_{U1} \;\; k_{U2} \;\; k_{L1} \;\; s_2 \;\; k_{U3} \;\; k_{L2} \;\; k_{L3} \;\; s_3]^T$, which is a vector made up of the rows of S. The matrix Ω is given by

$$\Omega \equiv \begin{bmatrix} (\tilde{\boldsymbol{\omega}} - \boldsymbol{\beta})^T & 0_{1\times3} & 0_{1\times3} \\ 0_{1\times3} & (\tilde{\boldsymbol{\omega}} - \boldsymbol{\beta})^T & 0_{1\times3} \\ 0_{1\times3} & 0_{1\times3} & (\tilde{\boldsymbol{\omega}} - \boldsymbol{\beta})^T \end{bmatrix} \tag{7}$$

Both Eq. (1) and Eq. (3) represent continuous-time gyro models but in practice discrete-time gyro measurements are employed and therefore discrete-time models are required. The discrete-time gyro measurement model corresponding to Eq. (1) is given by[15]

$$\tilde{\boldsymbol{\omega}}_{k+1} = \boldsymbol{\omega}_{k+1} + \frac{1}{2}(\boldsymbol{\beta}_{k+1} + \boldsymbol{\beta}_k) + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}\mathbf{N}_{v_k} \tag{8a}$$

$$\boldsymbol{\beta}_{k+1} = \boldsymbol{\beta}_k + \sigma_u\,\Delta t^{1/2}\mathbf{N}_{u_k} \tag{8b}$$

where the subscript k denotes the k^{th} time-step, Δt is the gyro sampling interval, and $\mathbf{N}_{v_k}$ and $\mathbf{N}_{u_k}$ are zero-mean Gaussian white-noise processes with covariance each given by the identity matrix. Replacing $\boldsymbol{\omega}_{k+1}$ with $(I_{3\times3} + S)\,\boldsymbol{\omega}_{k+1}$ in Eq. (8a) provides the discrete-time model for Eq. (3).

B. Star Tracker Measurement Model

The unit vector form for the j^{th} star tracker observation can be written as

$$\mathbf{b}_j = A\,\mathbf{r}_j \tag{9}$$

where $\mathbf{b}_j$ is star observation in body coordinates, $\mathbf{r}_j$ is the corresponding inertial unit vector and A is the attitude matrix which is a proper orthogonal matrix. When measurement noise is present, Ref. 16 has shown that nearly all the probability of the errors is concentrated on a very small area about the direction of $\mathbf{b}_j$, so the sphere containing that point can be approximated by a tangent plane, characterized by

$$\tilde{\mathbf{b}}_j = A\,\mathbf{r}_i + \boldsymbol{\upsilon}_j, \quad \boldsymbol{\upsilon}_j^T\mathbf{b}_j = 0 \tag{10}$$

where $\tilde{\mathbf{b}}_j$ denotes the measurement of $\mathbf{b}_j$ and the sensor error $\boldsymbol{\upsilon}_j$ is approximately Gaussian and white, and satisfies

$$E\left\{\boldsymbol{\upsilon}_j\right\} = \mathbf{0} \tag{11a}$$

$$E\left\{\boldsymbol{\upsilon}_j\boldsymbol{\upsilon}_j^T\right\} \equiv R_j = \sigma^2\left(I_{3\times 3} - \mathbf{b}_j\mathbf{b}_j^T\right) \tag{11b}$$

Equation (11b) is known as the *QUEST measurement model*. Note that Eq. (11b) is a also function of the unknown true values in $\mathbf{b}_j$. However, the advantage of using the QUEST measurement model is that the measurement covariance in a filter formulation can effectively be replaced by a nonsingular matrix, given by $\sigma^2 I_{3\times 3}$, which does not contain the unknown true values (see Ref. 17 for more details).

IV. Angular Velocity Estimation from Star Measurements

In this section a least-squares approach used for determining the angular velocity from star tracker body measurements alone is summarized.[11] Consider the following unit-vector measurement model at time t_k:

$$\tilde{\mathbf{b}}_{j_k} = A_k\,\mathbf{r}_j + \boldsymbol{\upsilon}_{j_k} \tag{12}$$

We note that the inertially-fixed star reference vector $\mathbf{r}_j$ is time-independent, neglecting effects such as proper motion and velocity aberration. Taking the difference between successive measurements of Eq. (12) gives

$$\tilde{\mathbf{b}}_{j_{k+1}} - \tilde{\mathbf{b}}_{j_k} = (A_{k+1} - A_k)\,\mathbf{r}_j + \boldsymbol{\upsilon}_{j_{k+1}} - \boldsymbol{\upsilon}_{j_k} \tag{13}$$

We assume that the body angular velocity $\boldsymbol{\omega}$ is constant between t_k and t_{k+1}, and ignore terms higher than first order in $\boldsymbol{\omega}\,\Delta t$. With these assumptions the following first-order approximation can be used:[18]

$$A_{k+1} \approx \left\{I_{3\times 3} - \Delta t\,[\boldsymbol{\omega}_k\times]\right\}\,A_k \tag{14}$$

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

$$[\mathbf{a}\times] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{15}$$

for any general 3×1 vector $\mathbf{a}$. Substituting Eq. (14) into Eq. (13) gives

$$\tilde{\mathbf{b}}_{j_{k+1}} - \tilde{\mathbf{b}}_{j_k} = -\Delta t\,[\boldsymbol{\omega}_k\times]\,A_k\,\mathbf{r}_j + \boldsymbol{\upsilon}_{j_{k+1}} - \boldsymbol{\upsilon}_{j_k} \tag{16}$$

Our goal is to determine an angular velocity estimate independent of attitude and the reference vectors. This is accomplished by solving Eq. (12) in terms of $A_k\,\mathbf{r}_i$ and substituting the resultant into Eq. (16), which yields

$$\frac{1}{\Delta t}(\tilde{\mathbf{b}}_{j_{k+1}} - \tilde{\mathbf{b}}_{j_k}) = [\tilde{\mathbf{b}}_{j_k}\times]\,\boldsymbol{\omega}_k + \mathbf{w}_{j_{k+1,k}} \tag{17}$$

where $\mathbf{w}_{j_{k+1,k}}$ is the new effective measurement noise vector given by

$$\mathbf{w}_{j_{k+1,k}} \equiv [\boldsymbol{\omega}_k\times]\,\boldsymbol{\upsilon}_{j_k} + \frac{1}{\Delta t}(\boldsymbol{\upsilon}_{j_{k+1}} - \boldsymbol{\upsilon}_{j_k}) \tag{18}$$

Note that Δt will have finite values, since discrete-time measurements are assumed.

The new measurement noise $\mathbf{w}_{j_{k+1},k}$ in Eq. (17) is now a function of the angular velocity vector. Assuming a stationary noise process for $\boldsymbol{\upsilon}_j$, the following covariance expression can be derived:

$$\begin{aligned} E\left\{\mathbf{w}_{j_{k+1},k}\mathbf{w}_{j_{k+1},k}^T\right\} \equiv \mathcal{R}_{j_{k+1},k} = \sigma_j^2\Big\{&[\boldsymbol{\omega}_k\times](I_{3\times3}-\mathbf{b}_{j_k}\mathbf{b}_{j_k}^T)[\boldsymbol{\omega}_k\times]^T \\ &-(1/\Delta t)[\boldsymbol{\omega}_k\times](I_{3\times3}-\mathbf{b}_{j_k}\mathbf{b}_{j_k}^T)+(1/\Delta t)(I_{3\times3}-\mathbf{b}_{j_k}\mathbf{b}_{j_k}^T)[\boldsymbol{\omega}_k\times] \\ &+(1/\Delta t^2)(2\,I_{3\times3}-\mathbf{b}_{j_{k+1}}\mathbf{b}_{j_{k+1}}^T-\mathbf{b}_{j_k}\mathbf{b}_{j_k}^T)\Big\} \end{aligned} \tag{19}$$

It can be shown that the terms that depend on $\boldsymbol{\omega}_k$ in Eq. (19) are second-order in nature and thus can be ignored.[11] Assuming that these terms can be ignored, the inverse of $\mathcal{R}_{j_{k+1},k}$ is given by[16]

$$\mathcal{R}_{j_{k+1},k}^{-1} = \frac{\Delta t^2}{2\sigma_j^2}\left\{I_{3\times3}+||\mathbf{b}_{j_{k+1}}\times\mathbf{b}_{j_k}||^{-2}\left[(\mathbf{b}_{j_{k+1}}^T\mathbf{b}_{j_k})(\mathbf{b}_{j_{k+1}}\mathbf{b}_{j_k}^T+\mathbf{b}_{j_k}\mathbf{b}_{j_{k+1}}^T)+\mathbf{b}_{j_{k+1}}\mathbf{b}_{j_{k+1}}^T+\mathbf{b}_{j_k}\mathbf{b}_{j_k}^T\right]\right\} \tag{20}$$

Note that the covariance is computed using the true values. However, using a similar analysis as the one shown in Ref. 19 indicates that the errors produced by replacing the true values with the measured ones are higher order in nature and thus are negligible.

Equation (17) can now be cast into a linear least-squares form for all measurement vectors, which leads to

$$\hat{\boldsymbol{\omega}}_k = \frac{1}{\Delta t}\left\{\sum_{j=1}^{n_k}[\tilde{\mathbf{b}}_{j_k}\times]^T\mathcal{R}_{j_{k+1},k}^{-1}[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1}\sum_{j=1}^{n_k}[\tilde{\mathbf{b}}_{j_k}\times]^T\mathcal{R}_{j_{k+1},k}^{-1}(\tilde{\mathbf{b}}_{j_{k+1}}-\tilde{\mathbf{b}}_{j_k}) \tag{21}$$

where $\hat{\boldsymbol{\omega}}_k$ is the estimate of $\boldsymbol{\omega}_k$ and n_k is the number of available body measurements at time t_k. Only knowledge of the body vector measurements, sampling interval and measurement covariance is required to derive an angular velocity estimate. Therefore, the attitude and star reference vectors are not required to be known. Hence, stars do not need to be identified to determine the angular velocity.

For small Δt the propagated true value of $\mathbf{b}_j$ can be given using Eq. (14):

$$\mathbf{b}_{j_{k+1}} \approx \{I_{3\times3}-\Delta t\,[\boldsymbol{\omega}_k\times]\}\,\mathbf{b}_{j_k} \tag{22}$$

Substituting Eq. (22) into Eq. (20), left multiplying by $[\mathbf{b}_{j_k}\times]^T$ and right multiplying by $[\mathbf{b}_{j_k}\times]$ gives

$$[\mathbf{b}_{j_k}\times]^T\mathcal{R}_{j_{k+1},k}^{-1}[\mathbf{b}_{j_k}\times] = \bar{\sigma}_j^{-2}[\mathbf{b}_{j_k}\times]^T[\mathbf{b}_{j_k}\times] \tag{23}$$

where $\bar{\sigma}_j^2 \equiv 2\sigma_j^2/\Delta t^2$. Also, since $\mathbf{b}_{j_{k+1}}^T\mathbf{b}_{j_k} \approx 1$, it is easy to show that $[\mathbf{b}_{j_k}\times]^T\mathcal{R}_{j_{k+1},k}^{-1}(\mathbf{b}_{j_{k+1}}-\mathbf{b}_{j_k}) \approx \bar{\sigma}_j^{-2}[\mathbf{b}_{j_k}\times]^T\mathbf{b}_{j_{k+1}}$. Therefore, Eq. (21) is well approximated by

$$\hat{\boldsymbol{\omega}}_k = \frac{1}{\Delta t}\left\{\sum_{j=1}^{n_k}\bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1}\sum_{j=1}^{n_k}\bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T\tilde{\mathbf{b}}_{j_{k+1}} \tag{24}$$

The predicted error-covariance (i.e., the best available estimate of the error-covariance) is given by

$$P_k \equiv E\left\{(\hat{\boldsymbol{\omega}}_k-\boldsymbol{\omega}_k)\,(\hat{\boldsymbol{\omega}}_k-\boldsymbol{\omega}_k)^T\right\} \approx \left\{\sum_{j=1}^{n_k}\bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T\,[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1} \tag{25}$$

At least two non-collinear body vectors are required for observability, i.e. for the inverse in Eq. (25) to exist. The error-covariance still depends on the sampling interval due to $\bar{\sigma}_j^2$. Also, if $\bar{\sigma}_j^2$ is replaced by σ_j^2 this yields the attitude error-covariance, which can be derived from maximum likelihood.[19]

Using a central difference the angular velocity estimate can be shown to be given by

$$\hat{\boldsymbol{\omega}}_k = \frac{1}{2\,\Delta t}\left\{\sum_{j=1}^{n_k}\bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1}\sum_{j=1}^{n_k}\bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T(\tilde{\mathbf{b}}_{j_{k+1}}-\tilde{\mathbf{b}}_{j_{k-1}}) \tag{26}$$

where $\bar{\sigma}_j^2 \equiv \sigma_j^2/\Delta t^2$. The error-covariance derivation leads to the same expression given in Eq. (25) with $\bar{\sigma}_j^2 = (1/\Delta t^2)\sigma_j^2$. Using a second-order difference the angular velocity estimate can be shown to be given by

$$\hat{\boldsymbol{\omega}}_k = \frac{1}{2\,\Delta t}\left\{\sum_{j=1}^{n_k} \bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1} \sum_{j=1}^{n_k} \bar{\sigma}_j^{-2}[\tilde{\mathbf{b}}_{j_k}\times]^T(4\,\tilde{\mathbf{b}}_{j_{k+1}} - \tilde{\mathbf{b}}_{j_{k+2}}) \tag{27}$$

where $\bar{\sigma}_j^2 \equiv [(13\sigma_j^2)/(2\,\Delta t^2)]$. The error-covariance derivation leads to the same expression given in Eq. (25) with $\bar{\sigma}_j^2 = [13/(2\,\Delta t^2)]\sigma_j^2$. More details can be found in Ref. 11.

V. Initialization

To allow the use of the GMMAE structure for calibration of a spacecraft gyro, the GMMAE algorithm must be provided with initial state estimates and estimated state error covariance. The GMMAE state vector consists of 15 elements: 3 attitude states, 3 gyro biases, 3 scaling factors and 6 misalignment parameters. Two initialization techniques shown here. The first is used to estimate for scaler factors and misalignments, and the second is used to estimate for the gyro biases.

A. Scale Factor/Misalignment Initialization

The scaling factors can be initialized independently of the gyro bias by first substituting Eq. (8b) into Eq. (8a), resulting in

$$\tilde{\boldsymbol{\omega}}_{k+1} = \mathcal{T}\boldsymbol{\omega}_{k+1} + \boldsymbol{\beta}_k + \frac{1}{2}\sigma_u\,\Delta t^{1/2}\mathbf{N}_{u_k} + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}\mathbf{N}_{v_k} \tag{28}$$

where $\mathcal{T} = (I_{3\times 3} + S)$. Taking this equation one time step forward gives

$$\tilde{\boldsymbol{\omega}}_{k+2} = \mathcal{T}\boldsymbol{\omega}_{k+2} + \boldsymbol{\beta}_{k+1} + \frac{1}{2}\sigma_u\,\Delta t^{1/2}\mathbf{N}_{u_{k+1}} + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}\mathbf{N}_{v_{k+1}} \tag{29}$$

Subtracting Eq. (28) from Eq. (29) yields the following expression:

$$\Delta\tilde{\boldsymbol{\omega}}_{k+2,k+1} = \mathcal{T}\Delta\boldsymbol{\omega}_{k+2,k+1} + \boldsymbol{\beta}_{k+1} - \boldsymbol{\beta}_k + \frac{1}{2}\sigma_u\,\Delta t^{1/2}(\mathbf{N}_{u_{k+1}} - \mathbf{N}_{u_k}) + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}(\mathbf{N}_{v_{k+1}} - \mathbf{N}_{v_k}) \tag{30}$$

where $\Delta\tilde{\boldsymbol{\omega}}_{k+2,k+1} = \tilde{\boldsymbol{\omega}}_{k+2} - \tilde{\boldsymbol{\omega}}_{k+1}$ and $\Delta\boldsymbol{\omega}_{k+2,k+1} = \boldsymbol{\omega}_{k+2} - \boldsymbol{\omega}_{k+1}$. Substituting $\boldsymbol{\beta}_{k+1} - \boldsymbol{\beta}_k = \sigma_u\,\Delta t^{1/2}\mathbf{N}_{u_k}$ from Eq. (8b) into Eq. (30) results in

$$\Delta\tilde{\boldsymbol{\omega}}_{k+2,k+1} = \mathcal{T}\Delta\boldsymbol{\omega}_{k+2,k+1} + \frac{1}{2}\sigma_u\,\Delta t^{1/2}(\mathbf{N}_{u_{k+1}} + \mathbf{N}_{u_k}) + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}(\mathbf{N}_{v_{k+1}} - \mathbf{N}_{v_k}) \tag{31}$$

Note that the difference is not a function of the gyro biases. Thus, Eq. (31) can be used to provide a crude estimate of the scale factors and misalignments, which can be written in the following form:

$$\Delta\tilde{\boldsymbol{\omega}}_{k+2,k+1} = \mathcal{T}\Delta\boldsymbol{\omega}_{k+2,k+1} + \Delta\mathbf{w}_{k+1,k} \tag{32}$$

where $\Delta\mathbf{w}_{k+1,k}$ is the effective noise term, which has zero mean and covariance given by

$$\begin{aligned} R_{\Delta w} \equiv E\{\Delta\mathbf{w}_{k+1,k}\Delta\mathbf{w}_{k+1,k}^T\} &= \left(\frac{1}{2}\sigma_u^2\,\Delta t\right)I_{3\times 3} + 2\left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)I_{3\times 3} \\ &= 2\left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{3}\sigma_u^2\,\Delta t\right)I_{3\times 3} \end{aligned} \tag{33}$$

The scale factors and misalignments estimation problem can now be written as a linear least squares problem with

$$\tilde{\mathbf{y}}_\kappa = H_\kappa\boldsymbol{\kappa} + \mathbf{v}_\kappa \tag{34}$$

where $\tilde{\mathbf{y}}_\kappa$ is the stacked measurement vector, $\mathbf{v}_\kappa$ is the stacked measurement noise vector and H_κ is the observation function matrix:

$$\tilde{\mathbf{y}}_\kappa = \begin{bmatrix} \Delta\tilde{\boldsymbol{\omega}}_{2,1} \\ \vdots \\ \Delta\tilde{\boldsymbol{\omega}}_{k+2,k+1} \end{bmatrix}, \quad \mathbf{v}_\kappa = \begin{bmatrix} \Delta\mathbf{w}_{0,1} \\ \vdots \\ \Delta\mathbf{w}_{k+1,k} \end{bmatrix}, \quad H_\kappa = \begin{bmatrix} \Delta\Omega_{2,1} \\ \vdots \\ \Delta\Omega_{k+2,k+1} \end{bmatrix} \tag{35}$$

where

$$\Delta\Omega_{k+2,k+1} = \begin{bmatrix} \Delta\boldsymbol{\omega}^T_{k+2,k+1} & 0_{1\times3} & 0_{1\times3} \\ 0_{1\times3} & \Delta\boldsymbol{\omega}^T_{k+2,k+1} & 0_{1\times3} \\ 0_{1\times3} & 0_{1\times3} & \Delta\boldsymbol{\omega}^T_{k+2,k+1} \end{bmatrix} \tag{36}$$

Then given all measurements for some value of κ the least square estimate is

$$\hat{\boldsymbol{\kappa}} = \left(H^T_\kappa R^{-1}_{\Delta w} H_\kappa\right)^{-1} H^T_\kappa R^{-1}_{\Delta w} \tilde{\mathbf{y}}_\kappa \tag{37}$$

where $\hat{\boldsymbol{\kappa}}$ denotes the estimate for $\boldsymbol{\kappa}$. The true angular velocity difference, $\Delta\boldsymbol{\omega}_{k+2,k+1}$, is not known in practice. This is replaced with the estimate given by the star tracker determined angular velocity, i.e. from Eq. (24). This unfortunately adds noise to the observation function matrix. For this case the optimal solution is given by total least squares. Since the errors in the estimated star tracker determined angular velocities are non-stationary, then an iterative approach is required to solve the total least squares problem.[20] However, the simple least squares solution given by Eq. (37) is adequate enough for the initialization of the filters. The main issue involved with this estimation approach is that noise is enhanced by the classic "root-two" factor in the velocity differences, plus the extra noise due to star-tracker derived angular velocity.

B. Bias Initialization

Once the scale factors and misalignments are estimated the next step involves estimating for the gyro biases. Solving Eq. (8b) for $\boldsymbol{\beta}_{k+1}$, then substituting the resultant into Eq. (8a) and taking one time-step back leads to

$$\boldsymbol{\omega}_k = \tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1} + \mathbf{z}_k \tag{38}$$

where

$$\mathbf{z}_k \equiv \frac{1}{2}\sigma_u\,\Delta t^{1/2}\mathbf{N}_{u_k} + \left(\frac{\sigma_v^2}{\Delta t} + \frac{1}{12}\sigma_u^2\,\Delta t\right)^{1/2}\mathbf{N}_{v_k} \tag{39}$$

Then

$$Q \equiv E\left\{\mathbf{z}_k\,\mathbf{z}_k^T\right\} = \left(\frac{1}{3}\sigma_u^2\,\Delta t + \frac{1}{12}\sigma_u^2\right) I_{3\times3} \tag{40}$$

Substituting Eq. (38) into Eq. (17) gives

$$\frac{1}{\Delta t}(\tilde{\mathbf{b}}_{j_{k+1}} - \tilde{\mathbf{b}}_{j_k}) + [\tilde{\boldsymbol{\omega}}_k\times]\tilde{\mathbf{b}}_{j_k} = -[\tilde{\mathbf{b}}_{j_k}\times]\,\boldsymbol{\beta}_{k-1} + \boldsymbol{\varpi}_{j_{k+1,k}} \tag{41}$$

where

$$\boldsymbol{\varpi}_{j_{k+1,k}} \equiv [(\tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1})\times]\boldsymbol{\upsilon}_{j_k} + [\tilde{\mathbf{b}}_{j_k}\times]\mathbf{z}_k + [\mathbf{z}_k\times]\boldsymbol{\upsilon}_{j_k} + \frac{1}{\Delta t}(\boldsymbol{\upsilon}_{j_{k+1}} - \boldsymbol{\upsilon}_{j_k}) \tag{42}$$

To compute the covariance of $\boldsymbol{\varpi}_{j_{k+1,k}}$ it is required to evaluate $E\{[\mathbf{z}_k\times]\boldsymbol{\upsilon}_{j_k}\boldsymbol{\upsilon}^T_{j_k}[\mathbf{z}_k\times]^T\}$. Consider two random and uncorrelated vectors $\mathbf{a}$ and $\mathbf{b}$ with zero mean and covariances given by A and B, respectively. Using the identities $\mathbf{b}\,\mathbf{b}^T = [\mathbf{b}\times]^2 + (\mathbf{b}^T\mathbf{b})I_{3\times3}$ and $[\mathbf{a}\times][\mathbf{b}\times] = \mathbf{b}\,\mathbf{a}^T - (\mathbf{a}^T\mathbf{b})I_{3\times3}$ leads to

$$[\mathbf{a}\times]\mathbf{b}\,\mathbf{b}^T[\mathbf{a}\times]^T = [(\mathbf{b}^T\mathbf{b})(\mathbf{a}^T\mathbf{a}) - (\mathbf{a}^T\mathbf{b})^2]I_{3\times3} + (\mathbf{a}^T\mathbf{b})(\mathbf{a}\,\mathbf{b}^T + \mathbf{b}\,\mathbf{a}^T) - (\mathbf{a}^T\mathbf{a})\mathbf{b}\,\mathbf{b}^T - (\mathbf{b}^T\mathbf{b})\mathbf{a}\,\mathbf{a}^T \tag{43}$$

The expectations of the various parts of Eq. (43) are given by

$$E\left\{(\mathbf{b}^T\mathbf{b})(\mathbf{a}^T\mathbf{a})\right\} = \mathrm{Tr}(A)\mathrm{Tr}(B) \tag{44a}$$

$$E\left\{(\mathbf{a}^T\mathbf{b})^2\right\} = \mathrm{Tr}(A\,B) \tag{44b}$$

$$E\left\{(\mathbf{a}^T\mathbf{b})(\mathbf{a}\,\mathbf{b}^T + \mathbf{b}\,\mathbf{a}^T)\right\} = A\,B + B\,A \tag{44c}$$

$$E\left\{(\mathbf{a}^T\mathbf{a})\mathbf{b}\,\mathbf{b}^T\right\} = \mathrm{Tr}(A)B \tag{44d}$$

$$E\left\{(\mathbf{b}^T\mathbf{b})\mathbf{a}\,\mathbf{a}^T\right\} = \mathrm{Tr}(B)A \tag{44e}$$

Then

$$E\left\{[\mathbf{a}\times]\mathbf{b}\,\mathbf{b}^T[\mathbf{a}\times]^T\right\} = [\mathrm{Tr}(A)\mathrm{Tr}(B) - \mathrm{Tr}(A\,B)]I_{3\times3} + A\,B + B\,A - \mathrm{Tr}(A)B - \mathrm{Tr}(B)A \tag{45}$$

Therefore the covariance of $\boldsymbol{\varpi}_{j_{k+1,k}}$ is given by

$$\begin{aligned}\mathcal{Z}_{j_{k+1,k}} \equiv E\left\{\boldsymbol{\varpi}_{j_{k+1,k}}\boldsymbol{\varpi}_{j_{k+1,k}}^T\right\} &= [(\tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1})\times]R_{j_k}[(\tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1})\times]^T \\ &- \frac{1}{\Delta t}\left\{[(\tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1})\times]R_{j_k} + R_{j_k}[(\tilde{\boldsymbol{\omega}}_k - \boldsymbol{\beta}_{k-1})\times]^T\right\} + [\tilde{\mathbf{b}}_{j_k}\times]Q\,[\tilde{\mathbf{b}}_{j_k}\times]^T \\ &+ \frac{1}{\Delta t^2}(R_{j_{k+1}} + R_{j_k}) + [\mathrm{Tr}(Q)\mathrm{Tr}(R_{j_k}) - \mathrm{Tr}(Q\,R_{j_k})]I_{3\times3} \\ &+ Q\,R_{j_k} + R_{j_k}Q - \mathrm{Tr}(Q)R_{j_k} - \mathrm{Tr}(R_{j_k})Q\end{aligned} \tag{46}$$

Using the same derivations provided in Ref. 11, if the sampling frequency is within Nyquist's limit then the terms that depend on $\tilde{\boldsymbol{\omega}}_k$ and $\boldsymbol{\beta}_{k-1}$ are second-order in nature and can be ignored. Thus, $\mathcal{Z}_{j_{k+1,k}}$ can be approximated by

$$\begin{aligned}\mathcal{Z}_{j_{k+1,k}} &\approx [\tilde{\mathbf{b}}_{j_k}\times]Q\,[\tilde{\mathbf{b}}_{j_k}\times]^T + \frac{1}{\Delta t^2}(R_{j_{k+1}} + R_{j_k}) \\ &+ [\mathrm{Tr}(Q)\mathrm{Tr}(R_{j_k}) - \mathrm{Tr}(Q\,R_{j_k})]I_{3\times3} + Q\,R_{j_k} + R_{j_k}Q - \mathrm{Tr}(Q)R_{j_k} - \mathrm{Tr}(R_{j_k})Q\end{aligned} \tag{47}$$

The least squares estimate for $\boldsymbol{\beta}_{k-1}$, denoted by $\hat{\boldsymbol{\beta}}_{k-1}$, is given by

$$\hat{\boldsymbol{\beta}}_{k-1} = -\left\{\sum_{j=1}^{n_k}[\tilde{\mathbf{b}}_{j_k}\times]^T\mathcal{Z}_{j_{k+1,k}}^{-1}[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1}\sum_{j=1}^{n_k}[\tilde{\mathbf{b}}_{j_k}\times]^T\mathcal{Z}_{j_{k+1,k}}^{-1}\left\{[\tilde{\boldsymbol{\omega}}_k\times]\tilde{\mathbf{b}}_{j_k} + \tilde{\mathbf{b}}_{j_{k+1}}/\Delta t\right\} \tag{48}$$

The covariance of $\hat{\boldsymbol{\beta}}_{k-1}$ is simply given by

$$E\left\{(\hat{\boldsymbol{\beta}}_{k-1} - \boldsymbol{\beta}_{k-1})(\hat{\boldsymbol{\beta}}_{k-1} - \boldsymbol{\beta}_{k-1})^T\right\} = \left\{\sum_{j=1}^{n_k}[\tilde{\mathbf{b}}_{j_k}\times]^T\mathcal{Z}_{j_{k+1,k}}^{-1}[\tilde{\mathbf{b}}_{j_k}\times]\right\}^{-1} \tag{49}$$

Note that, as with Eq. (25), at least two non-collinear body vectors are required for the inverse in Eq. (48) to exist.

VI. Kalman Filtering for Attitude Estimation

An EKF is now summarized for estimating attitude states and calibration parameters of a spacecraft (see Ref. 9 for more details). The system model is modeled by

$$\dot{\mathbf{x}} = \mathbf{f}\,(\mathbf{x}, t) + G\,(t)\,\mathbf{w}(t) \tag{50}$$

where $\mathbf{w}(t)$ is a Gaussian white-noise process term with correlation function $Q\delta(t_1 - t_2)$. The function $\mathbf{f}\,(\mathbf{x}, t)$ is a general nonlinear function. To solve the general nonlinear filtering problem the EKF linearizes the nonlinear equations to determine the mean and covariance propagation through the function $\mathbf{f}\,(\mathbf{x}, t)$. If the initial pdf $p(\mathbf{x}_o)$ that describes the associated state uncertainty is given, the solution for the time evolution of $p(\mathbf{x}, t)$ constitutes the nonlinear filtering problem. Given a system model with initial state and covariance values, the EKF propagates the state vector and the error covariance matrix recursively. At discrete observation times, the EKF updates the state and covariance matrix conditioned on the information gained by the measurements. The prediction phase is important for overall filter performance. In general, the discrete measurement equation can be expressed for the filter as

$$\tilde{\mathbf{y}}_k = \mathbf{h}\,(\mathbf{x}_k) + \mathbf{v}_k \tag{51}$$

where $\tilde{\mathbf{y}}_k$ is a measurement vector and $\mathbf{v}_k$ is the measurement noise, which is assumed to be a zero-mean Gaussian noise process with covariance R_k.

The 15-state EKF vector is made up of the quaternion, $\mathbf{q}$, gyro biases (drift), $\boldsymbol{\beta}$, scale factors, $\mathbf{s}$, and misalignments, $\mathbf{k}_U$ and $\mathbf{k}_L$. The error dynamics are given by

$$\Delta\dot{\tilde{\mathbf{x}}}(t) = F(t)\,\Delta\tilde{\mathbf{x}}(t) + G(t)\,\mathbf{w}(t) \tag{52}$$

where $F(t)$ is the Jacobian of $\mathbf{f}\,(\mathbf{x}, t)$ and

$$\Delta\tilde{\mathbf{x}}(t) \equiv \begin{bmatrix} \boldsymbol{\delta\alpha}^T(t) & \Delta\boldsymbol{\beta}^T(t) & \Delta\mathbf{s}(t) & \Delta\mathbf{k}_U(t) & \Delta\mathbf{k}_L(t) \end{bmatrix}^T \tag{53a}$$

$$\mathbf{w}(t) \equiv \begin{bmatrix} \boldsymbol{\eta}_v^T(t) & \boldsymbol{\eta}_u^T(t) & \boldsymbol{\eta}_s^T(t) & \boldsymbol{\eta}_U^T(t) & \boldsymbol{\eta}_L^T(t) \end{bmatrix}^T \tag{53b}$$

The term $\boldsymbol{\delta\alpha}$ is the vector of small attitude errors[8] and the rest of the terms in Eq. (53a) are the errors in the gyro biases, scale factors and misalignments. The matrices $F(t)$ and $G(t)$ matrices are now derived for the 15-state filter. A standard multiplicative error quaternion is used to derive the attitude errors:

$$\boldsymbol{\delta\mathbf{q}} = \mathbf{q} \otimes \hat{\mathbf{q}}^{-1} \approx \begin{bmatrix} \frac{1}{2}\boldsymbol{\delta\alpha} \\ 1 \end{bmatrix} \tag{54}$$

where $\boldsymbol{\delta\alpha}$ is the vector of small attitude (roll, pitch and yaw) attitude errors. The error-kinematics follow

$$\boldsymbol{\delta\dot{\alpha}} = -[\hat{\boldsymbol{\omega}}\times]\boldsymbol{\delta\alpha} + \boldsymbol{\delta\omega} \tag{55}$$

where $\boldsymbol{\delta\omega} = \boldsymbol{\omega} - \hat{\boldsymbol{\omega}}$. From Eq. (3) we have

$$\boldsymbol{\omega} = (I_{3\times3} - S)\tilde{\boldsymbol{\omega}} - (I_{3\times3} - S)\boldsymbol{\beta} - (I_{3\times3} - S)\boldsymbol{\eta}_v \tag{56a}$$

$$\hat{\boldsymbol{\omega}} = (I_{3\times3} - \hat{S})\tilde{\boldsymbol{\omega}} - (I_{3\times3} - \hat{S})\hat{\boldsymbol{\beta}} \tag{56b}$$

Then $\boldsymbol{\delta\omega}$ is given by

$$\boldsymbol{\delta\omega} = -\Delta S\,\tilde{\boldsymbol{\omega}} - \Delta\boldsymbol{\beta} + \left(\Delta S - \hat{S}\right)\left(\Delta\boldsymbol{\beta} + \hat{\boldsymbol{\beta}}\right) - \hat{S}\,\hat{\boldsymbol{\beta}} - \left(I_{3\times3} - \hat{S} - \Delta S\right)\boldsymbol{\eta}_v \tag{57}$$

where $\Delta S \equiv S - \hat{S}$ and $\Delta\boldsymbol{\beta} \equiv \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}$. Ignoring second-order terms leads to

$$\boldsymbol{\delta\omega} = -\left(I_{3\times3} - \hat{S}\right)\Delta\boldsymbol{\beta} - \mathrm{diag}\left(\tilde{\boldsymbol{\omega}} - \hat{\boldsymbol{\beta}}\right)\Delta\mathbf{s} - \hat{U}\Delta\mathbf{k}_U - \hat{L}\Delta\mathbf{k}_L - \left(I_{3\times3} - \hat{S}\right)\boldsymbol{\eta}_v \tag{58}$$

with

$$\hat{U} = \begin{bmatrix} \tilde{\omega}_2 - \hat{\beta}_2 & \tilde{\omega}_3 - \hat{\beta}_3 & 0 \\ 0 & 0 & \tilde{\omega}_3 - \hat{\beta}_3 \\ 0 & 0 & 0 \end{bmatrix} \tag{59a}$$

$$\hat{L} = \begin{bmatrix} 0 & 0 & 0 \\ \tilde{\omega}_1 - \hat{\beta}_1 & 0 & 0 \\ 0 & \tilde{\omega}_1 - \hat{\beta}_1 & \tilde{\omega}_2 - \hat{\beta}_2 \end{bmatrix} \tag{59b}$$

where diag denotes a diagonal matrix, $\Delta\mathbf{s}$ is a vector of the diagonal elements of ΔS, and $\Delta\mathbf{k}_U$ and $\Delta\mathbf{k}_L$ correspond to the upper and lower off-diagonal elements of ΔS, respectively. Hence, the matrices $F(t)$, $G(t)$

and $Q(t)$, which is the spectral density of $\mathbf{w}(t)$, are given by

$$F(t) = \begin{bmatrix} -[\hat{\boldsymbol{\omega}}(t)\times] & -\left(I_{3\times3} - \hat{S}\right) & -\text{diag}\left(\tilde{\boldsymbol{\omega}} - \hat{\boldsymbol{\beta}}\right) & -\hat{U} & -\hat{L} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{60a}$$

$$G(t) = \begin{bmatrix} -\left(I_{3\times3} - \hat{S}\right) & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{60b}$$

$$Q(t) = \text{blkdiag}\left(\left[\sigma_v^2\, I_{3\times3}\ \sigma_u^2\, I_{3\times3}\ \sigma_s^2\, I_{3\times3}\ \sigma_U^2\, I_{3\times3}\ \sigma_L^2\, I_{3\times3}\ \right]\right) \tag{60c}$$

where blkdiag denotes a block diagonal matrix of appropriate dimension.

The observation sensitivity matrix $H_k(\hat{\mathbf{x}}_k^-)$ used in the EKF is now derived. Multiple star tracker vector measurements can be concatenated to form

$$\tilde{\mathbf{y}}_k = \left.\begin{bmatrix} A(\mathbf{q})\mathbf{r}_1 \\ A(\mathbf{q})\mathbf{r}_2 \\ \vdots \\ A(\mathbf{q})\mathbf{r}_{n_k} \end{bmatrix}\right|_{t_k} + \left.\begin{bmatrix} \boldsymbol{\nu}_1 \\ \boldsymbol{\nu}_2 \\ \vdots \\ \boldsymbol{\nu}_{n_k} \end{bmatrix}\right|_{t_k} \equiv \mathbf{h}_k(\hat{\mathbf{x}}_k) + \mathbf{v}_k \tag{61a}$$

$$R = \text{blkdiag}\left[\sigma_1^2 I_{3\times3} \quad \sigma_2^2 I_{3\times3} \quad \cdots \quad \sigma_{n_k}^2 I_{3\times3}\right] \tag{61b}$$

where n_k is the number of stars available at times t_k. It is important to note that for each star measurement vector $\tilde{\mathbf{b}}_j$, the associated covariance matrix used in the EKF is given by $\sigma_j^2 I_{3\times3}$. Using this covariance model avoids singularity in the gain calculation[17] as discussed previously.

The true attitude matrix, $A(\mathbf{q})$, is related to the *a priori* attitude, $A(\hat{\mathbf{q}}^-)$, through

$$A(\mathbf{q}) = A(\boldsymbol{\delta}\mathbf{q})A(\hat{\mathbf{q}}^-) \tag{62}$$

The first-order approximation of the error-attitude matrix, $A(\boldsymbol{\delta}\mathbf{q})$, is given by

$$A(\boldsymbol{\delta}\mathbf{q}) \approx I_{3\times3} - [\boldsymbol{\delta}\boldsymbol{\alpha}\times] \tag{63}$$

For a single sensor the true and estimated body vectors are given by

$$\mathbf{b} = A(\mathbf{q})\mathbf{r} \tag{64a}$$

$$\hat{\mathbf{b}}^- = A(\hat{\mathbf{q}}^-)\mathbf{r} \tag{64b}$$

Substituting Eqs. (62) and (63) into Eq. (64) yields

$$\Delta\mathbf{b} = [A(\hat{\mathbf{q}}^-)\mathbf{r}\times]\boldsymbol{\delta}\boldsymbol{\alpha} \tag{65}$$

where $\Delta\mathbf{b} \equiv \mathbf{b} - \hat{\mathbf{b}}^-$. The sensitivity matrix for all measurement sets is therefore given by

$$H_k(\hat{\mathbf{x}}_k^-) = \left.\begin{bmatrix} [A(\hat{\mathbf{q}}^-)\mathbf{r}_1\times] & 0_{3\times12} \\ [A(\hat{\mathbf{q}}^-)\mathbf{r}_2\times] & 0_{3\times12} \\ \vdots & \vdots \\ [A(\hat{\mathbf{q}}^-)\mathbf{r}_{n_k}\times] & 0_{3\times12} \end{bmatrix}\right|_{t_k} \tag{66}$$

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

$$\Delta\hat{\tilde{\mathbf{x}}}_k^+ = K_k[\tilde{\mathbf{y}}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)] \tag{67}$$

where $\Delta\hat{\tilde{\mathbf{x}}}_k^+ \equiv \begin{bmatrix}\boldsymbol{\delta\hat{\alpha}}_k^{+T} & \Delta\hat{\boldsymbol{\beta}}_k^{+T} & \Delta\hat{\mathbf{s}}_k^{+T} & \Delta\hat{\mathbf{k}}_{U_k}^{+T} & \Delta\hat{\mathbf{k}}_{L_k}^{+T}\end{bmatrix}^T$ and $\mathbf{h}_k(\hat{\mathbf{x}}_k^-)$ is the estimate output, given by

$$\mathbf{h}_k(\hat{\mathbf{x}}_k^-) = \left.\begin{bmatrix} A(\hat{\mathbf{q}}^-)\mathbf{r}_1 \\ A(\hat{\mathbf{q}}^-)\mathbf{r}_2 \\ \vdots \\ A(\hat{\mathbf{q}}^-)\mathbf{r}_{n_k}\end{bmatrix}\right|_{t_k} \tag{68}$$

The gyro bias and other calibration parameters updates are simply given by

$$\hat{\boldsymbol{\beta}}_k^+ = \hat{\boldsymbol{\beta}}_k^- + \Delta\hat{\boldsymbol{\beta}}_k^+ \tag{69a}$$

$$\hat{\mathbf{s}}_k^+ = \hat{\mathbf{s}}_k^- + \Delta\hat{\mathbf{s}}_k^+ \tag{69b}$$

$$\hat{\mathbf{k}}_{U_k}^+ = \hat{\mathbf{k}}_{U_k}^- + \Delta\hat{\mathbf{k}}_{U_k}^+ \tag{69c}$$

$$\hat{\mathbf{k}}_{L_k}^+ = \hat{\mathbf{k}}_{L_k}^- + \Delta\hat{\mathbf{k}}_{L_k}^+ \tag{69d}$$

To within first-order the quaternion update is given by

$$\hat{\mathbf{q}}_k^+ = \begin{bmatrix}\frac{1}{2}\boldsymbol{\delta\hat{\alpha}}_k^+ \\ 1\end{bmatrix} \otimes \hat{\mathbf{q}}_k^- \tag{70}$$

Using quaternion multiplication gives[8]

$$\hat{\mathbf{q}}_k^+ = \hat{\mathbf{q}}_k^- + \frac{1}{2}\Xi(\hat{\mathbf{q}}_k^-)\boldsymbol{\delta\hat{\alpha}}_k^+ \tag{71}$$

where

$$\Xi(\mathbf{q}) \equiv \begin{bmatrix} q_4 I_{3\times 3} + [\boldsymbol{\varrho}\times] \\ -\boldsymbol{\varrho}^T \end{bmatrix} \tag{72}$$

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 (the bias initial condition is usually assumed zero) and error-covariance matrix. The first three diagonal elements of the error-covariance matrix correspond to attitude errors. Then, the Kalman gain is computed using the measurement-error covariance R and sensitivity matrix in Eq. (66). The state error-covariance follows the standard EKF update, while the error-state update is computed using Eq. (67). The gyro bias, scale factors and misalignments updates are given by Eq. (69) and the quaternion update is given by Eq. (71). Also, the updated quaternion is re-normalized by brute force. Finally, the estimated angular velocity is used to propagate the quaternion kinematics model in and standard error-covariance in the EKF. The 6-state and 9-state filters are simply truncated versions of the 15-state filter shown here.

A discrete propagation can be used for the covariance matrix in order to reduce the computational load:

$$P_{k+1}^- = \Phi_k P_k^+ \Phi_k^T + \mathcal{Q}_k \tag{73}$$

where Φ_k is the state transition matrix and $\mathcal{Q}_k$ is the discrete-time covariance matrix. The first step is to set up the following $2n$ by $2n$ matrix:[21]

$$\mathcal{A} = \begin{bmatrix} -F & GQG^T \\ 0_{n\times n} & F^T \end{bmatrix}\Delta t \tag{74}$$

The matrix exponential is then calculated:

$$\mathcal{B} = e^{\mathcal{A}} = \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{12} \\ 0_{n\times n} & \mathcal{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{B}_{11} & \Phi_k^{-1}\mathcal{Q}_k \\ 0_{n\times n} & \Phi_k^T \end{bmatrix} \tag{75}$$

Table 1. Extended Kalman Filter for Attitude Estimation

Initialize	$\hat{\mathbf{q}}(t_0) = \mathbf{q}_0, \quad \hat{\boldsymbol{\beta}}(t_0) = \boldsymbol{\beta}_0, \quad \hat{\mathbf{s}}(t_0) = \mathbf{s}_0$ $\hat{\mathbf{k}}_U(t_0) = \mathbf{k}_{U_0}, \quad \hat{\mathbf{k}}_L(t_0) = \mathbf{k}_{L_0}, \quad P(t_0) = P_0$
Gain	$K_k = P_k^- H_k^T(\hat{\mathbf{x}}_k^-)[H_k(\hat{\mathbf{x}}_k^-)P_k^- H_k^T(\hat{\mathbf{x}}_k^-) + R]^{-1}$ $H_k(\hat{\mathbf{x}}_k^-) = \begin{bmatrix} [A(\hat{\mathbf{q}}^-)\mathbf{r}_1\times] & 0_{3\times 12} \\ \vdots & \vdots \\ [A(\hat{\mathbf{q}}^-)\mathbf{r}_{n_k}\times] & 0_{3\times 12} \end{bmatrix}\Bigg\|_{t_k}$
Update	$P_k^+ = [I - K_k H_k(\hat{\mathbf{x}}_k^-)]P_k^-$ $\Delta\hat{\tilde{\mathbf{x}}}_k^+ = K_k[\tilde{\mathbf{y}}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)]$ $\Delta\hat{\tilde{\mathbf{x}}}_k^+ \equiv \begin{bmatrix} \boldsymbol{\delta\hat{\alpha}}_k^{+T} & \Delta\hat{\boldsymbol{\beta}}_k^{+T} & \Delta\hat{\mathbf{s}}_k^{+T} & \Delta\hat{\mathbf{k}}_{U_k}^{+T} & \Delta\hat{\mathbf{k}}_{L_k}^{+T} \end{bmatrix}^T$ $\mathbf{h}_k(\hat{\mathbf{x}}_k^-) = \begin{bmatrix} A(\hat{\mathbf{q}}^-)\mathbf{r}_1 \\ A(\hat{\mathbf{q}}^-)\mathbf{r}_2 \\ \vdots \\ A(\hat{\mathbf{q}}^-)\mathbf{r}_{n_k} \end{bmatrix}\Bigg\|_{t_k}$ $\hat{\mathbf{q}}_k^+ = \hat{\mathbf{q}}_k^- + \frac{1}{2}\Xi(\hat{\mathbf{q}}_k^-)\boldsymbol{\delta\hat{\alpha}}_k^+, \quad$ re-normalize quaternion $\hat{\boldsymbol{\beta}}_k^+ = \hat{\boldsymbol{\beta}}_k^- + \Delta\hat{\boldsymbol{\beta}}_k^+$ $\hat{\mathbf{s}}_k^+ = \hat{\mathbf{s}}_k^- + \Delta\hat{\mathbf{s}}_k^+$ $\hat{\mathbf{k}}_{U_k}^+ = \hat{\mathbf{k}}_{U_k}^- + \Delta\hat{\mathbf{k}}_{U_k}^+$ $\hat{\mathbf{k}}_{L_k}^+ = \hat{\mathbf{k}}_{L_k}^- + \Delta\hat{\mathbf{k}}_{L_k}^+$
Propagation	$\hat{\boldsymbol{\omega}}(t) = [I_{3\times 3} - \hat{S}(t)][\tilde{\boldsymbol{\omega}}(t) - \hat{\boldsymbol{\beta}}(t)]$ $\dot{\hat{\mathbf{q}}}(t) = \frac{1}{2}\Xi(\hat{\mathbf{q}}(t))\,\hat{\boldsymbol{\omega}}(t)$ $\dot{P}(t) = F(t)\,P(t) + P(t)\,F^T(t) + G(t)\,Q(t)\,G^T(t)$

The state transition and covariance matrices are then given as

$$\Phi_k = \mathcal{B}_{22}^T \tag{76a}$$

$$\mathcal{Q}_k = \Phi_k \mathcal{B}_{12} \tag{76b}$$

Note that Eq. (74) is only valid for constant system and covariance matrices. However, for small Δt it is a good approximation of the actual discrete-time matrices that are time varying.

VII. Generalized Multiple-Model Adaptive Estimation

In this section a review of GMMAE is shown. More details can be found in Ref. 11. Multiple-model adaptive estimation is a recursive estimator that uses a bank of M filters that depend on some unknown parameters, denoted by the vector $\mathbf{p}$, which is assumed to be constant (at least throughout the interval of adaptation). Note that the case shown in this paper there is no parameter $\mathbf{p}$ but rather different models themselves. Still, the theory shown here is valid in this case. The MMAE process is shown in Figure 2.

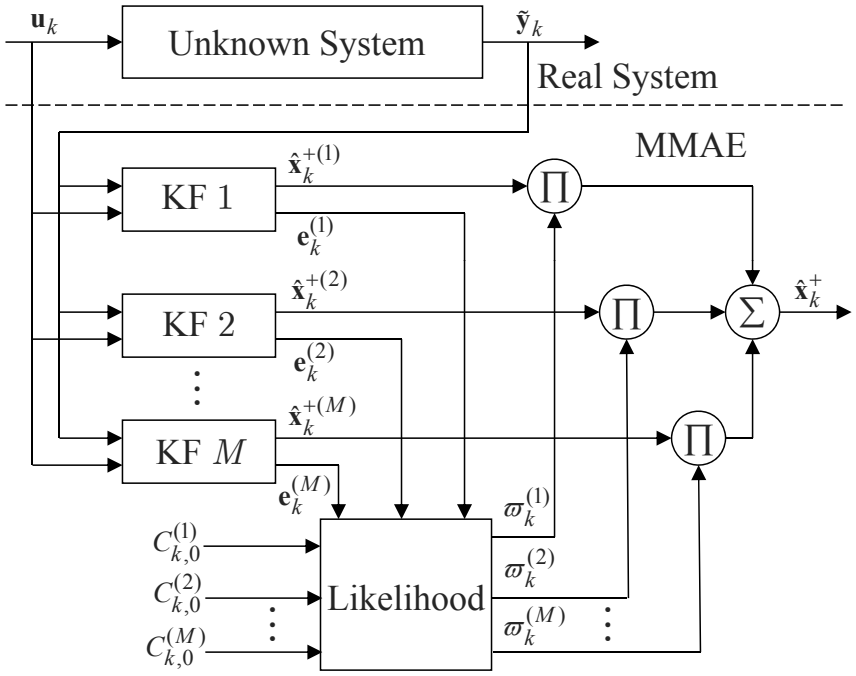


Figure 2. MMAE Process

Note that we do not 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 elements is generated for each of the M filters 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 all the measurements. Application of Bayes' rule yields

$$p(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_k) = \frac{p(\tilde{\mathbf{Y}}_k|\mathbf{p}^{(\ell)})\, p(\mathbf{p}^{(\ell)})}{\sum_{j=1}^{M} p(\tilde{\mathbf{Y}}_k|\mathbf{p}^{(j)})\, p(\mathbf{p}^{(j)})} \tag{77}$$

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[22]

$$\begin{aligned} p(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_k) &= \frac{p(\tilde{\mathbf{y}}_k,\, \mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_{k-1})}{p(\tilde{\mathbf{y}}_k|\tilde{\mathbf{Y}}_{k-1})} \\ &= \frac{p(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)})\, p(\mathbf{p}^{(\ell)}|\tilde{\mathbf{Y}}_{k-1})}{\sum_{j=1}^{M} \left[p(\tilde{\mathbf{Y}}_k|\hat{\mathbf{x}}_k^{-(j)})\, p(\mathbf{p}^{(j)}|\tilde{\mathbf{Y}}_{k-1}) \right]} \end{aligned} \tag{78}$$

with $p(\tilde{\mathbf{y}}_k, |\tilde{\mathbf{Y}}_{k-1},\, \mathbf{p}^{(\ell)})$ given by $p(\tilde{\mathbf{y}}_k|\hat{\mathbf{x}}_k^{-(\ell)})$ in the Kalman recursion, which is computed as a probability density function of the Kalman filter residual $\mathbf{e}_k^{(\ell)}$. Suppose that $\mathbf{e}_k^{(\ell)}$ is a zero-mean Gaussian random vector with covariance $C_{k,0}^{(\ell)}$. Then $p(\tilde{\mathbf{y}}_k, |\tilde{\mathbf{Y}}_{k-1},\, \mathbf{p}^{(\ell)}) = 1/\sqrt{\det[2\pi\, C_{k,0}^{(\ell)}]} \exp\left[-\mathbf{e}_k^T \left(C_{k,0}^{(\ell)} \right)^{-1} \mathbf{e}_k / 2 \right]$. Note that the denominator of (78) 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

$$\begin{aligned} \varpi_k^{(\ell)} &= \varpi_{k-1}^{(\ell)} p\left(\tilde{\mathbf{y}}_k | \hat{\mathbf{x}}_k^{-(\ell)}\right) \\ \varpi_k^{(\ell)} &\leftarrow \frac{\varpi_k^{(\ell)}}{\sum_{j=1}^{M} \varpi_k^{(j)}} \end{aligned} \tag{79}$$

where $\varpi_k^{(\ell)} \equiv p\left(\mathbf{p}^{(\ell)} | \tilde{\mathbf{y}}_k\right)$. Note that only the current time measurement $\tilde{\mathbf{y}}_k$ is needed to update the weights. 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. 22, 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.[23]

The model is assumed to be linear, but time varying which retains the EKF structure, with

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \Gamma_k \mathbf{u}_k + \Upsilon_k \mathbf{w}_k \tag{80a}$$

$$\tilde{\mathbf{y}}_k = H_k \mathbf{x}_k + \mathbf{v}_k \tag{80b}$$

where Υ_k is the discrete-time process noise distribution matrix. Note that the to-be-estimated parameter vector $\mathbf{p}$ can be function of any unknown model quantity in Eq. (80). The discrete-time residual is given by

$$\begin{aligned} \mathbf{e}_k &\equiv \tilde{\mathbf{y}}_k - H_k \hat{\mathbf{x}}_k^- \\ &= -H_k \tilde{\mathbf{x}}_k^- + \mathbf{v}_k \end{aligned} \tag{81}$$

where $\tilde{\mathbf{x}}_k^- \equiv \hat{\mathbf{x}}_k^- - \mathbf{x}_k$ and $\hat{\mathbf{x}}_k^- = \Phi_{k-1}(I - K_{k-1}H_{k-1})\,\hat{\mathbf{x}}_{k-1}^- + \Phi_{k-1}K_{k-1}\tilde{\mathbf{y}}_{k-1}$ with K_{k-1} being an arbitrary gain. Note that K_{k-1} is not necessarily the optimal Kalman gain. In the standard MMAE $\mathbf{e}_k$ is used to develop the update law. The GMMAE generalizes the MMAE by using a set of residuals going back i steps in time, given by

$$\boldsymbol{\epsilon}_{k,i} \equiv \begin{bmatrix} \mathbf{e}_k \\ \mathbf{e}_{k-1} \\ \vdots \\ \mathbf{e}_{k-i} \end{bmatrix} \tag{82}$$

The likelihood function associated with $\boldsymbol{\epsilon}_{k,i}$ is given by

$$L_{k,i} = \frac{1}{[\det(2\pi\,\mathcal{C}_{k,i})]^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{\epsilon}_{k,i}^T \mathcal{C}_{k,i}^{-1} \boldsymbol{\epsilon}_{k,i}\right) \tag{83}$$

where $\mathcal{C}_{k,i} \equiv E\{\boldsymbol{\epsilon}_{k,i}\boldsymbol{\epsilon}_{k,i}^T\}$ is given by

$$\mathcal{C}_{k,i} = \begin{bmatrix} C_{k,\,0} & C_{k,\,1} & C_{k,\,2} & \cdots & C_{k,\,i} \\ C_{k,\,1}^T & C_{k-1,\,0} & C_{k-1,\,1} & \cdots & C_{k-1,\,i-1} \\ C_{k,\,2}^T & C_{k-1,\,1}^T & C_{k-2,\,0} & \cdots & C_{k-2,\,i-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{k,\,i}^T & C_{k-1,\,i-1}^T & C_{k-2,\,i-2}^T & \cdots & C_{k-i,\,0} \end{bmatrix} \tag{84}$$

When $i = 0$ the likelihood function reduces down to

$$L_{k,0} = \frac{1}{\left\{\det[2\pi\,(H_k P_k^- H_k^T + R_k)]\right\}^{1/2}} \exp\left[-\frac{1}{2}\mathbf{e}_k^T (H_k P_k^- H_k^T + R_k)^{-1} \mathbf{e}_k\right] \tag{85}$$

which is the likelihood used in the standard MMAE.

The GMMAE adaptive law is a modification of the MMAE adaptive law, given by:

$$\begin{aligned} \varpi_k^{(\ell)} &= \varpi_{k-1}^{(\ell)} L_{k,i}^{(\ell)} \\ \varpi_k^{(\ell)} &\leftarrow \frac{\varpi_k^{(\ell)}}{\sum_{j=1}^{M} \varpi_k^{(j)}} \end{aligned} \tag{86}$$

with

$$L_{k,i}^{(\ell)} = \frac{1}{\left[\det\left(2\pi\,\mathcal{C}_{k,i}^{(\ell)}\right)\right]^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{\epsilon}_i^{(\ell)T}\left(\mathcal{C}_{k,i}^{(\ell)}\right)^{-1}\boldsymbol{\epsilon}_i^{(\ell)}\right] \tag{87}$$

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 estimates for $C_{k,\,i}^{(\ell)}$, which are sub-blocks of $\mathcal{C}_{k,i}^{(\ell)}$, are given by

$$\hat{C}_{k,\,i}^{(\ell)} = \begin{cases} H_k(\hat{\mathbf{x}}_k^{-(\ell)})P_k^{-(\ell)}H_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + R_k & i = 0 \\ H_k(\hat{\mathbf{x}}_k^{-(\ell)})\Phi_{k-1}(\hat{\mathbf{x}}_{k-1}^{-(\ell)})\left[P_{k-1}^{-(\ell)}H_{k-1}^T(\hat{\mathbf{x}}_{k-1}^{-(\ell)}) - \hat{K}_{k-1}C_{k-1,\,0}^{(\ell)}\right] & i = 1 \\ H_k(\hat{\mathbf{x}}_k^{-(\ell)})\left\{\prod_{j=1}^{i-1}\Phi_{k-j}(\hat{\mathbf{x}}_{k-j}^{-(\ell)})\left[I - \hat{K}_{k-j}\,H_{k-j}(\hat{\mathbf{x}}_{k-j}^{-(\ell)})\right]\right\}\Phi_{k-i}(\hat{\mathbf{x}}_{k-i}^{-(\ell)}) & \\ \times\left[P_{k-i}^{-(\ell)}\,H_{k-i}^T(\hat{\mathbf{x}}_{k-1}^{-(\ell)}) - \hat{K}_{k-i}\,C_{k-i,\,0}^{(\ell)}\right] & i > 1 \end{cases} \tag{88}$$

where

$$C_{k-i,\,0}^{(\ell)} \equiv H_{k-i}(\hat{\mathbf{x}}_{k-i}^{-(\ell)})\,P_{k-i}^{-(\ell)}\,H_{k-i}^T(\hat{\mathbf{x}}_{k-i}^{-(\ell)}) + R_{k-i} \tag{89}$$

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

$$P_{k+1}^{-(\ell)} = \Phi_k(\hat{\mathbf{x}}_k^{-(\ell)})P_k^{+(\ell)}\Phi_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + Q^{(\ell)} \tag{90a}$$

$$P_k^{+(\ell)} = \left[I - K_k^{(\ell)}H_k(\hat{\mathbf{x}}_k^{-(\ell)})\right]P_k^{-(\ell)} \tag{90b}$$

$$K_k^{(\ell)} = P_k^{-(\ell)}H_k^T\left[H_k(\hat{\mathbf{x}}_k^{-(\ell)})P_k^{-(\ell)}H_k^T(\hat{\mathbf{x}}_k^{-(\ell)}) + R_k\right]^{-1} \tag{90c}$$

The estimate of the optimal gain is computed using

$$\hat{K}_k = \hat{P}_k^- H_k^T(\hat{\mathbf{x}}_k^-)\left[H_k(\hat{\mathbf{x}}_k^-)\hat{P}_k^- H_k^T(\hat{\mathbf{x}}_k^-) + R_k\right]^{-1} \tag{91}$$

with

$$\hat{P}_{k+1}^- = \Phi_k(\hat{\mathbf{x}}_k^-)\hat{P}_k^+\Phi_k^T(\hat{\mathbf{x}}_k^-) + \hat{Q}_k \tag{92a}$$

$$\hat{P}_k^+ = \left[I - \hat{K}_k H_k(\hat{\mathbf{x}}_k^-)\right]\hat{P}_k^- \tag{92b}$$

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 is executed. For each filter the state estimates, $\hat{\mathbf{x}}_k^{-(\ell)}$, and measurements are used to form the residual, $\boldsymbol{\epsilon}_{k,i}^{(\ell)}$, 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 $\hat{\mathcal{C}}_{k,i}^{(\ell)}$. Note that at each new measurement time, all elements of $\hat{\mathcal{C}}_{k,i}^{(\ell)}$ 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 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 $L_{k,i}^{(\ell)}$. These functions are used to update the weights. A proof of convergence of the GMMAE process is given in Ref. 13, which shows that the GMMAE always provides better convergence properties than the standard MMAE.

VIII. Persistency of Excitation

The calibration observability is a function of the attitude motion, therefore a requirement is often placed on the spacecraft calibration maneuver to ensure observability. The exponential convergence rate of $\boldsymbol{\kappa}$ in Eq. (6) is proportional to the lower bound of the observability Grammian of the pair $[F(t),\, H]$, where $F(t)$

for $\boldsymbol{\kappa}$ is a matrix of zeros and therefore $\exp[F(t)\Delta t] = I_{9\times 9}$. Given that $\boldsymbol{\kappa}$ is constant the observability Grammian can be written as

$$W(t) = \int_0^T \Omega^T \Omega \, dt \tag{93}$$

where T is the period over which data is collected and Ω is given by Eq. (7). The lower bound of the observability Grammian for $[F(t),\ H]$ is proportional to the lower bound of the observability Grammian of the equivalent system $[F(t) - K(t)H,\ H]$ where $K(t)$ is the Kalman gain. Therefore the minimum eigenvalue of the observability Grammian gives the exponential convergence rate. In other words the larger the eigenvalue for a given T, the faster the estimate converges over that period. The amplitude and frequency of the desired angular velocity can be set based on mission or vehicle constraints, such as available torque authority.

IX. Simulation Results

In this section simulation results are shown for both scale factor and misalignment parameters as well as gyro bias parameter estimates. It is assumed that the first set of data is collected for 10 minutes during the initialization mode to initialize all calibration parameters. Two examples are considered to show the proposed calibration process. In the first example, errors exist in all calibration parameters, while in the second example there are only bias and scale factor errors but no misalignment errors. For the simulation examples a spacecraft is considered that has both a star tracker and a three-axis gyro onboard. Also it is assumed that at least two stars always exist in the field-of-view of the star tracker during the entire simulation time period. The spacecraft angular velocity is given by $\boldsymbol{\omega} = 0.0087 \times [\sin(0.01t) \ \ \sin(0.0085t) \ \ \cos(0.0085t)]^T$ rad/sec.

A. Initialization

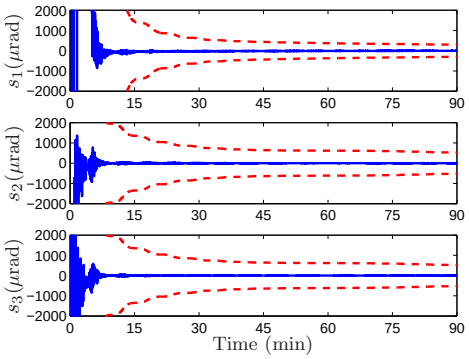
The calibration parameters used in this simulation are:

- Initial gyro bias: 0.1 deg/hr for each axis
- Gyro symmetric scale factors: $s_1 = 1,500$ μrad, $s_2 = 1,000$ μrad, $s_3 = 1,500$ μrad
- Gyro misalignments: $k_{U1} = 1,000$ μrad, $k_{U2} = 1,500$ μrad, $k_{U3} = 2,000$ μrad; $k_{L1} = 500$ μrad, $k_{L2} = 1,000$ μrad, $k_{L3} = 1,500$ μrad
- Star tracker standard derivation: 6 arc-sec
- Gyro noise parameters: $\sigma_u = \sqrt{10} \times 10^{-10}$ rad/sec$^{3/2}$, $\sigma_v = \sqrt{10} \times 10^{-7}$ rad/sec$^{1/2}$, $\sigma_s = \sigma_U = \sigma_L = 0$
- Initial quaternion: $\mathbf{q}_0 = [1/\sqrt{2} \ \ 0 \ \ 0 \ \ 1/\sqrt{2}]^T$

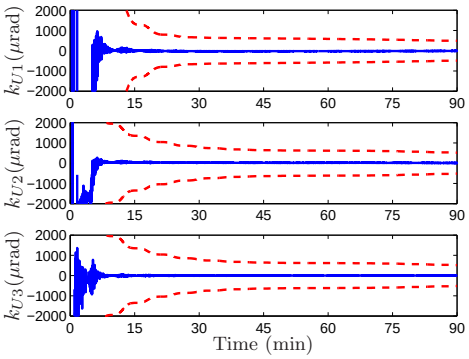
First Eq. (24) is used to determine the angular velocity independently of gyro measurements. Then following this the angular velocity in Eq. (37) is used to determine the scale factors independently of the gyro biases. Once the scale factors have been determined the gyro biases can be determined from Eq. (48). Although the solutions for the angular velocity, scale factors and misalignment parameters, as well as the gyro biases estimates are written in least squares form, the estimates in this simulation are determined recursively using a sequential least squares process. This is done to show how the estimates converge over time, which can be used to assess the convergence of the initialization procedure for proper handoff to the operational mode. Figure 3 shows the results for the initialization of the scale factors and misalignment parameters. All parameters show good convergence properties. The gyro bias initialization estimates are show in Figure 4 and show a similar convergence rate as the scale factors and misalignment parameters.

B. GMMAE Approach: Example One

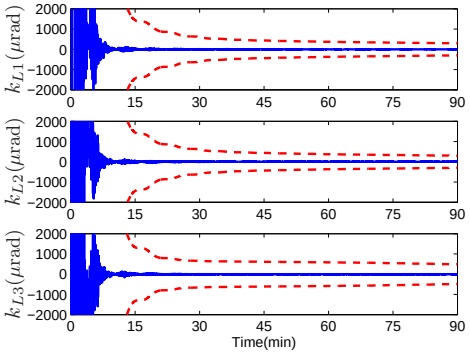
In this simulation scenario the GMMAE approach is used to calibrate the gyro sensor given that both misalignment and scale factor errors exist along with gyro bias errors. The true gyro calibration parameters are the same as the initialization simulations in the previous scenario.



(a) Scaling Factor Errors



(b) Upper Misalignment Errors



(c) Lower Misalignment Errors

Figure 3. Scaling and Misalignment Initialization

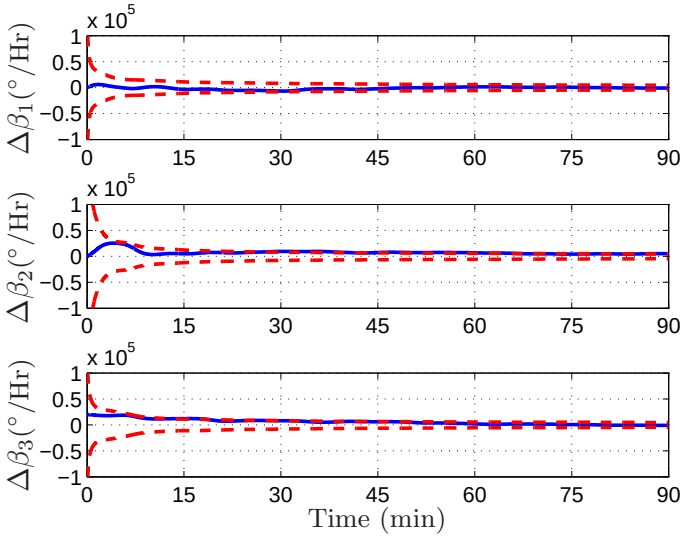


Figure 4. Bias Initialization Errors

The initial covariances given to each filters is the same. The initial attitude error standard deviation is $40/3$ μrad. The gyro bias initial standard deviation is 20 deg/hr for each component, and the scale factors and misalignment initial standard deviation is $2000/3$ μrad for each component. The initial estimates are drawn from a Gaussian distribution centered at the true values for the calibration parameters and attitude

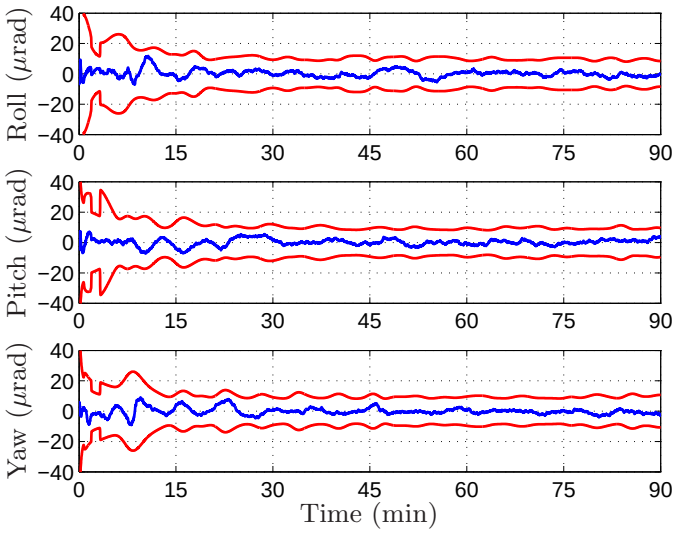


Figure 5. GMMAE Attitude Estimate Errors – Example 1

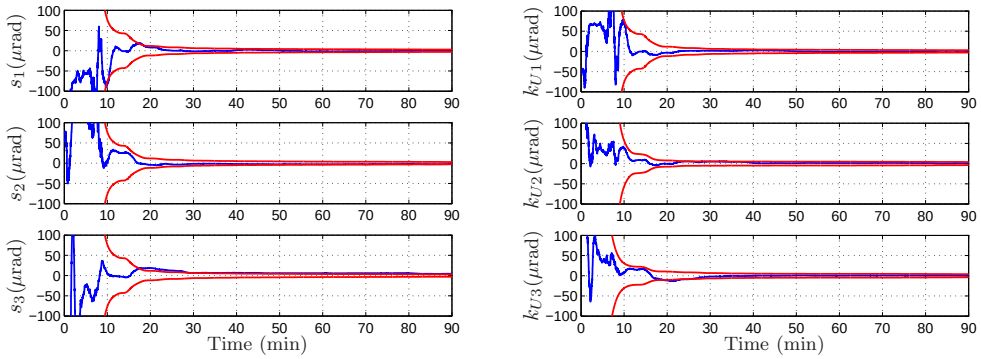
states with aforementioned standard deviations. Both the star tracker and gyro measurements are sampled at 1 Hz. Both the GMMAE and the MMAE (the MMAE is equivalent to GMMAE with one time-step back) run a combination of three filters to estimate the attitude and gyro calibration parameters for the spacecraft considered. The initial probabilities for all three filters in both the MMAE and GMMAE are set to 1/3, so each model is weighted equally at the beginning of the estimation process. For this simulation example the GMMAE is given 20 time-steps back.

The attitude estimation results for the GMMAE approach are given in Figure 5. It is clear from this figure that accurate attitude estimates are achieved in the presence of gyro misalignments, scale factor and bias errors. The calibration parameters estimates are shown in Figure 6 and good convergence is achieved. The GMMAE approach uses probabilistic weighting of the three filters. Since there are large errors in all calibration parameters the 15-state filter, which includes all the parameters, performs the best. It can be seen from Figure 7 that both the MMAE and GMMAE approach give the 15-state filter as the most probable model. From Figure 7 it is also noted that the GMMAE converges faster then the MMAE. Figure 8 shows the root-mean-squared (RMS) error for the 6-state, 9-state, and 15-state filters as well as the GMMAE mixture. From this figure it is seen that the 15-state filter constantly outperforms the other two filters and therefore is the filter selected by the GMMAE approach.

C. GMMAE Approach: Example Two

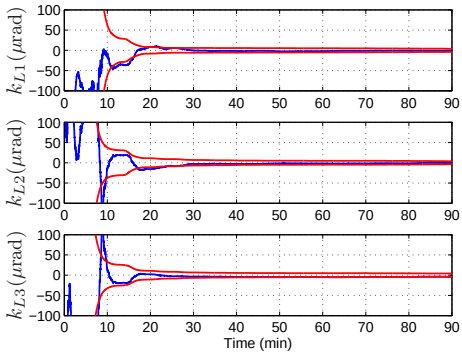
In this simulation scenario the GMMAE approach is used to calibrate a gyro sensor which has scale factor and bias errors but no misalignment errors. Also all simulation parameters are the same as is used in the previous GMMAE simulation except that the true misalignment errors are zero. The 15-state filter has the same initial condition as the previous example except that now the misalignment initial estimate is set to zero.

The RMS errors for each filter are plotted in Figure 9. From this figure it can be seen that initial errors for all filters are the same but as the 9-state filter converges it becomes the filter with the best performance. Since the 6-state filter does not include scale factors its performance is far less than the 9-state filter. The 15-state filter includes the scale factor parameters in its state vector but also estimates for the misalignment parameters which do not exist in the measurements. Therefore the 15-state filter has additional error due to information dilution which degrades the attitude state estimates. The attitude estimate errors determined from the GMMMAE approach are given in Figure 10. From Figure 11 it can be seen that the probability of both the MMAE and GMMAE converge to the 9-state filter but the GMMAE convergence is faster. The GMMAE takes about fifteen minutes to fully converge to the 9-state filter while the MMAE doesn't fully converge until well after that time.

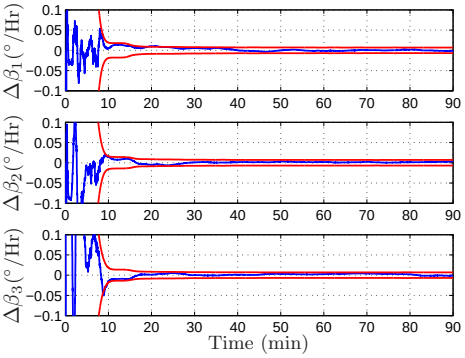


(a) Scaling Factor Errors

(b) Upper Misalignment Errors

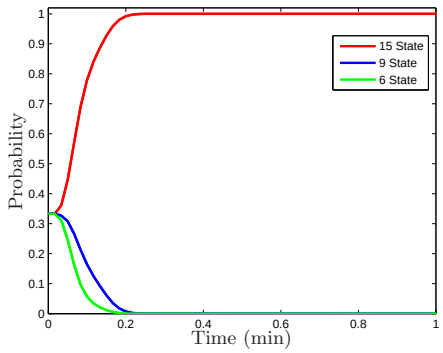


(c) Lower Misalignment Errors

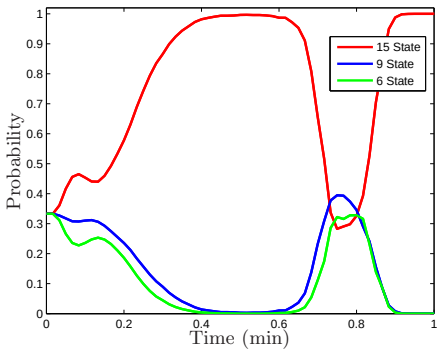


(d) Bias Errors

Figure 6. GMMAE Calibration Estimate Errors – Example 1



(a) GMMAE



(b) MMAE

Figure 7. Comparison of GMMAE and MMAE – Example 1

X. Conclusions

Operational responsive space systems require fast and accurate on-orbit calibration procedures. This paper presented a two-step process for robust calibration of gyros. The initialization procedure provides deterministic solutions for the gyro biases, scale factors and misalignments so that the issue of filter divergence due to large errors in the assumed initial parameters is mitigated. These solutions are then passed to a generalized multiple-model adaptive estimation approach, which combines three filters to overcome the issue of information dilution. The generalized multiple-model adaptive estimation is useful in that it automatically

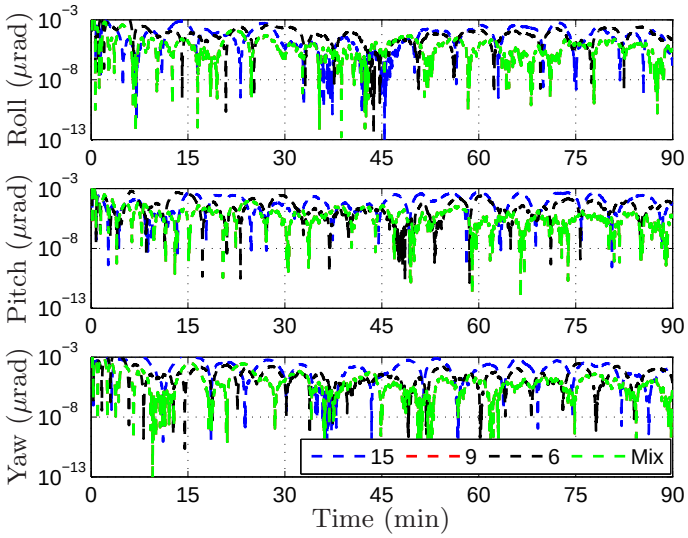


Figure 8. RMS Error for 6-, 9-, 15-State Filters and GMMAE Results – Example 1

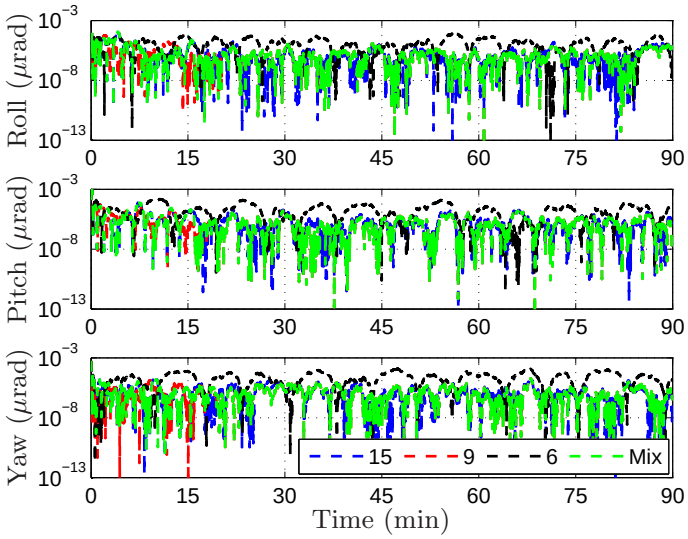


Figure 9. RMS Error for 6-, 9-, 15-State Filters and GMMAE Results – Example 2

chooses the model that produces the best estimates from a probabilistic point of view. The simulation results showed that the proposed two-step calibration approach effectively provides accurate estimates for the gyro calibration parameters.

XI. Acknowledgement

This work was sponsored by the Air Force Research Laboratory's Space Vehicles Directorate grant number FA9453-09-C-0356 under the supervision of Captain Chester Douglas McFarland. The authors greatly appreciate the support. The authors also wish to thank Dr. Quang Lam from Orbital Sciences Corporation for many help discussions and comments.

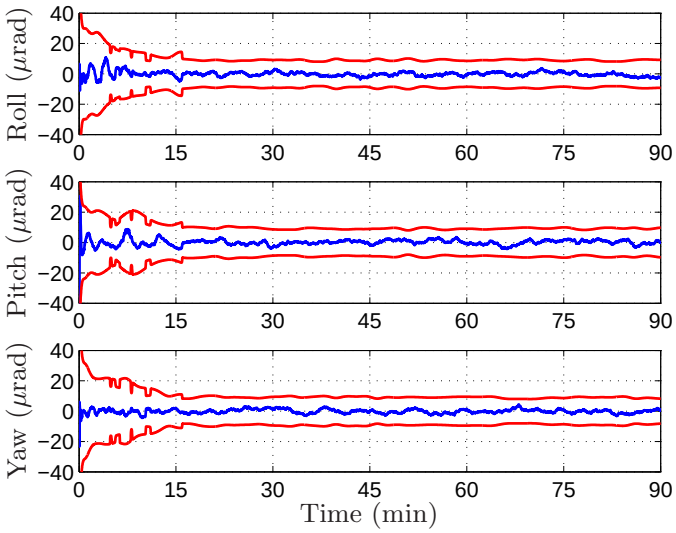
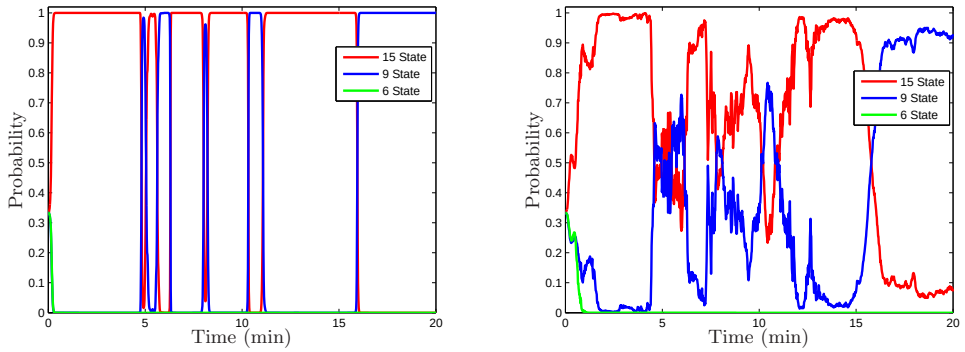


Figure 10. GMMAE Attitude Estimate Errors – Example 2



(a) GMMAE (b) MMAE

Figure 11. Comparison of GMMAE and MMAE – Example 2

References

[1]Hansen, L. J., Graven, P., Fogle, D., and Lyke, J., "The Feasibility of Applying Plug-and-Play Concepts to Spacecraft Guidance, Navigation, and Control Systems to Meet the Challenges of Future Responsive Missions," *7th International ESA Conference on Guidance, Navigation & Control Systems*, Tralee, County Kerry, Ireland, June 2008.

[2]Pittelkau, M. E., "Kalman Filtering for Spacecraft System Alignment Calibration," *Journal of Guidance, Control, and Dynamics*, Vol. 24, No. 6, 2001.

[3]Pandiyan, R., Solaiappan, A., and Malik, N., "A One Step Batch Filter for Estimating Gyroscope Calibration Parameters Using Star Vectors," *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Providence, RI, Aug. 2004, AIAA-04-4858.

[4]Shuster, M. D. and Pitone, D. S., "Batch Estimation of Spacecraft Sensor Alignments II. Absolute Alignment Estimation," *Journal of the Astronautical Sciences*, Vol. 39, No. 4, Oct.-Dec. 1991.

[5]Ward, D. K., Davis, G. T., and O'Donnell, J. R., "The Microwave Anisotropy Probe Guidance, Navigation and Control Hardware Suite," *AIAA Guidance, Navigation and Control Conference*, Monterey, CA, Aug. 2002, AIAA 2002-4579.

[6]Andrews, S. F. and Bilanow, S., "Recent Flight Results of the TRMM Kalman Filter," *AIAA Guidance, Navigation and Control Conference*, Monterey, CA, Aug. 2002, AIAA 2002-5047.

[7]Andrews, S. F. and Morgenstern, W. M., "Design, Implementation, Testing, and Flight Results of the TRMM Kalman Filter," *AIAA Guidance, Navigation and Control Conference*, Boston, MA, Aug. 1998, AIAA 1998-4509.

[8]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.

[9]Lam, Q. M. and Crassidis, J. L., "Evaluation of a Multiple Model Adaptive Estimation for Space Vehicle's Enhanced Navigation Solution," *AIAA Guidance, Navigation, and Control Conference*, Hilton Head, SC, Aug. 2007, AIAA 2007-6816.

[10]Rapoport, I. and Bar-Itzhack, I. Y., "On the Information Dilution Theorem and its Application to Attitude Determination," *Journal of the Astronautical Sciences*, Vol. 49, No. 2, July-Sept. 2001, pp. 489–508.

[11]Crassidis, J. L., "Angular Velocity Determination Directly from Star Tracker Measurements," *Journal of Guidance, Control and Dynamics*, Vol. 25, No. 6, Nov.-Dec. 2002, pp. 1165–1168.

[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]Alsuwaidan, B. and Crassidis, J. L., "Convergence Properties of Autocorrelation-Based Generalized Multiple-Model Adaptive Estimation," *AIAA Guidance, Navigation, and Control Conference*, Honolulu, HI, Aug. 2008, AIAA 2008-6476.

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

[15]Crassidis, J. L., "Sigma-Point Kalman Filtering for Integrated GPS and Inertial Navigation," *AIAA Guidance, Navigation, and Control Conference*, San Francisco, CA, Aug. 2005, AIAA-05-6052.

[16]Shuster, M. D. and Oh, S. D., "Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, Jan.-Feb. 1981, pp. 70–77.

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

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

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

[20]Schuermans, M., Markovsky, I., Wentzell, P. D., and Van Huffel, S., "On the Equivalence Between Total Least Squares and Maximum Likelihood PCA," *Analytica Chimica Acta*, Vol. 544, No. 1-2, 2005, pp. 254–267.

[21]van Loan, C. F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 3, June 1978, pp. 396–404.

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

[23]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.