Attitude Estimation Employing Common Frame Error Representations

Michael S. Andrle* and John L. Crassidis†

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

This work studies vector representation in the attitude filtering problem. Since vector coordinates must be estimated in lieu of the abstract vector itself, the choice of representation is important. A challenge arises in attitude estimation when the coordinates of a vector are sought with respect to the attitude state coordinate frame. While a wealth of research examines filtering with attitude state constraints or discontinuities, this work focuses on vector state estimation with respect to the attitude state frame. Specifically, an alternative vector state error is defined using common coordinates over all vector error realizations. A modified extended Kalman Filter is developed with this error realization as its foundation. Numerical simulations are conducted, and results obtained using the new filter are compared with those generated by employing a standard estimate error definition.

I. Introduction

Estimating the attitude of a rigid body has applications to many systems, including spacecraft [1], inertial navigation of aerial vehicles (both piloted and uninhabited) [2], underwater vehicles [3] and robotic systems [4]. In addition to the usual difficulties of nonlinear filtering, the attitude estimation problem poses additional challenges. Challenges are encountered either due to the non-global nature of minimal attitude parameterizations or because of the constraints inherent to non-minimal attitude state vectors.

The obvious advantage of employing a minimal attitude parameterization is that the need for enforcing constraints is circumvented. However, the difficulty is then that the attitude trajectory is not globally defined. This problem can either be overcome through a reference reset or by the operating mode of the satellite itself. For example, Euler angles can be directly employed in an extended Kalman filter (EKF) provided that the coordinate axes to which the vehicle attitude is referenced are appropriately adjusted to avoid a singularity [5]. Alternatively, discontinuities may be avoided altogether if the satellite operates well outside the neighborhood of any singularity. The operating mode dictates an appropriate implementation, such as the use of Modified Rodrigues Parameters EKF [6].

General satellite motion requires the use of globally defined attitude parameter sets. The challenge is that such attitude parameterizations must respect inherent constraints. The quaternion is a popular parameter choice since it contains the minimal number of parameters necessary for a globally defined attitude. Further, its kinematics do not involve trigonometric quantities and the constraint is easily restored by normalization [7].

Relaxing the norm constraint permits an immediate implementation using an EKF, however violating the quaternion group properties is undesirable. A form of the EKF constructed upon a multiplicative error addresses this concern [1]. This algorithm is referred to as the multiplicative extended Kalman filter (MEKF) [7]. Its development is sometimes interpreted as using a quaternion to parameterize the probability density function conditional mean while the associated error covariance is represented in a minimal parameter space [8].

*Technical Staff, MIT Lincoln Laboratory. Email: msandrle@buffalo.edu. Member AIAA.

†CUBRC Professor in Space Situational Awareness, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Fellow AIAA.

Filters that propagate and update a discrete set of sigma points rather than approximating a nonlinear function of the probability density function, such as the Unscented Kalman filter (UKF) [9], also overcome the above difficulties in attitude filtering. The UnScented QUaternion Estimator (USQUE) addresses these challenges [10]. The USQUE maintains the quaternion unit norm by construction since minimally parameterized attitude error sigma points are exactly transformed into global quaternion samples for the propagation stage. The update itself is performed directly in error space. Therefore the limitation is only that the attitude error covariance cannot approach a singularity.

While the aforementioned filtering methods, as well as several not mentioned, address attitude parameter singularities or constraints, much less attention has been given to the problem of estimating vector state coordinates with respect to the attitude state coordinate frame. A judicious treatment of constraints lead to alternative attitude error definitions, such as multiplicative quaternion error in the MEKF. It is demonstrated here that an alternative vector error definition is required for certain vector states as well. This work explores a new error definition in which vector error quantities are defined using elements expressed in a common frame. Developments concern specifically with the EKF since it is a popular algorithm for attitude estimation, and also since it serves well to illustrate the aforementioned issue.

This paper is organized as follows. First, the system equations are provided, and notation is established in Section II. The entirety of Section III is devoted to a detailed discussion of the ideas underlying this work. A modified EKF is developed in Section IV. This algorithm is referred to as the Geometric Extended Kalman Filter (GEKF), which explicitly accounts for state vector component coordinate bases through use of a new error definition. Section V conducts simulations and illustrates results for a realistic scenario. Finally, concluding remarks are made in Section VI.

II. Attitude System Model

Before proceeding with presenting new developments, notation is established by providing a terse explanation of the dynamic model employed for attitude estimation. To describe the attitude, two coordinate systems are defined: one on the vehicle body and one on the reference frame. The attitude matrix, A, maps from the reference frame to the vehicle body frame. according to $A\,\mathbf{r}$, where $\mathbf{r}$ is a component vector given with respect to the reference frame.

The attitude is subsequently parameterized by the quaternion $\mathbf{q}$. The quaternion is a four-dimensional vector, defined as

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

with

$$\boldsymbol{\varrho} \triangleq [q_1 \;\; q_2 \;\; q_3]^T = \mathbf{e}\sin(\vartheta/2) \tag{2a}$$

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

where $\mathbf{e}$ is the unit Euler axis and ϑ is the rotation angle [11]. A quaternion parameterizing an attitude satisfies a single constraint given by $\|\mathbf{q}\| = 1$. In terms of the quaternion, its associated attitude matrix is given by

$$A(\mathbf{q}) = \left(q_4^2 - ||\boldsymbol{\varrho}||^2\right) I_{3\times 3} + 2\boldsymbol{\varrho}\,\boldsymbol{\varrho}^T - 2q_4[\boldsymbol{\varrho}\times] \tag{3}$$

where $I_{3\times 3}$ is a 3×3 identity matrix. The matrix $[\boldsymbol{\varrho}\times]$ is the standard cross product matrix with

$$[\boldsymbol{\varrho}\times] \triangleq \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \tag{4}$$

With attitude parameterized by the quaternion $\mathbf{q}$, the physical model is then the quaternion kinematics, given by

$$\dot{\mathbf{q}} = \frac{1}{2}\Xi(\mathbf{q})\boldsymbol{\omega} = \frac{1}{2}\Omega(\boldsymbol{\omega})\mathbf{q} \tag{5}$$

where $\boldsymbol{\omega}$ is the angular rate vector and

$$\Xi(\mathbf{q}) \triangleq \begin{bmatrix} q_4 I_{3\times 3} + [\boldsymbol{\varrho}\times] \\ -\boldsymbol{\varrho}^T \end{bmatrix} \tag{6a}$$

$$\Omega(\boldsymbol{\omega}) \triangleq \begin{bmatrix} -[\boldsymbol{\omega}\times] & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix} \tag{6b}$$

Attitude estimation typically consists of combining the physical model with sensor measurements in order to calculate an attitude trajectory that is, in some sense, stochastically optimal. In addition to attitude sensing hardware, a rate-integrating gyro is employed to obtain angular rate information. The gyro output $\tilde{\boldsymbol{\omega}}$ is governed by

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \mathbf{b} + \boldsymbol{\eta}_v \tag{7a}$$

$$\dot{\mathbf{b}} = \boldsymbol{\eta}_u \tag{7b}$$

where the vector $\mathbf{b}$ is the gyro bias, and the vectors $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are assumed to be zero-mean, Gaussian white-noise processes with spectral densities given by $\sigma_v^2 I_{3\times 3}$ and $\sigma_u^2 I_{3\times 3}$, respectively. The covariance of

$$\mathbf{w} \triangleq \begin{bmatrix} \boldsymbol{\eta}_v \\ \boldsymbol{\eta}_u \end{bmatrix} \tag{8}$$

is given by

$$E\{\mathbf{w}(t)\mathbf{w}^T(\tau)\} = Q(t)\delta(t-\tau) \tag{9}$$

where the spectral density $Q(t)$ is then given by

$$Q(t) = \begin{bmatrix} \sigma_v^2 I_{3\times 3} & 0_{3\times 3} \\ 0_{3\times 3} & \sigma_u^2 I_{3\times 3} \end{bmatrix} \tag{10}$$

where $0_{3\times 3}$ denotes a 3×3 matrix of zeros. The model in Eq. (7) developed by Farrenkopf accounts for the bias drift associated with several types of gyros [12].

It is customary to substitute $\boldsymbol{\omega}$ from Eq. (7a) directly into Eq. (5) by which the complete state dynamics become

$$\dot{\mathbf{q}} = \frac{1}{2}\Xi(\mathbf{q})(\tilde{\boldsymbol{\omega}} - \mathbf{b} - \boldsymbol{\eta}_v) \tag{11a}$$

$$\dot{\mathbf{b}} = \boldsymbol{\eta}_u \tag{11b}$$

with $\tilde{\boldsymbol{\omega}}$ interpreted as a known input. The state vector $\mathbf{x}$ is then given by

$$\mathbf{x} = \begin{bmatrix} \mathbf{q} \\ \mathbf{b} \end{bmatrix} \tag{12}$$

This practice circumvents the need to employ Euler's equation which can have adverse consequences when torque disturbances are not well characterized or the inertia tensor is poorly known. The mean angular rate $\hat{\boldsymbol{\omega}}$ is found from Eq. (7a) to be

$$\hat{\boldsymbol{\omega}} = \tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}} \tag{13}$$

which is associated as an angular rate estimate, though it is important to emphasize that it is not a state and therefore not filtered by sequential estimation algorithms.

III. Constraints, Frames and Attitude Filtering

State estimation algorithms, including the EKF, are based on the definition of the state error. It has been argued and widely agreed upon in the literature that a quaternion error $\Delta\mathbf{q}$ given by

$$\Delta\mathbf{q} \triangleq \mathbf{q} - \hat{\mathbf{q}} \tag{14}$$

is not appropriate. This approach ignores the quaternion unit norm constraint. The MEKF addresses this problem by constructing an EKF based on a multiplicative update. As a result, the constraint is maintained to within first-order. The MEKF has subsequently become the workhorse of attitude filtering [1]. In this respect, the unconstrained nature of the EKF (and other filtering techniques) has been deeply investigated and this problem is well understood.

It is argued here that the bias error

$$\Delta\mathbf{b} \triangleq \mathbf{b} - \hat{\mathbf{b}} \tag{15}$$

also may not be appropriate. Consider the calculation of the bias estimate error covariance according to the MEKF (as well as according to the UKF and other manifestations of approximate nonlinear filtering). Under the assumption that the attitude error $\mathbf{dq} = \mathbf{q} \otimes \hat{\mathbf{q}}^{-1}$ is small, it can be approximated by the three-parameter small rotation vector $\mathbf{d}\boldsymbol{\alpha}$, which will be mathematically shown later, lying in the plane tangent to the quaternion unit hypersphere. For this case, the error covariance for the state defined in Eq. (12) is written as

$$E\{(\mathbf{e} - E\{\mathbf{e}\})(\mathbf{e} - E\{\mathbf{e}\})^T\} \triangleq \int_{\mathbb{R}^6} p_{\mathbf{e}}(\boldsymbol{e})\,(\boldsymbol{e} - \hat{\mathbf{e}})\,(\boldsymbol{e} - \hat{\mathbf{e}})^T d\boldsymbol{e} \tag{16}$$

where $\mathbf{e}$ is defined as

$$\mathbf{e} = \begin{bmatrix} \mathbf{d}\boldsymbol{\alpha} \\ \mathbf{b} - \hat{\mathbf{b}} \end{bmatrix} \tag{17}$$

and $\boldsymbol{e}$ is a realization of $\mathbf{e}$. Also, $p_{\mathbf{e}}(\boldsymbol{e})$ is the probability density function. However, for each realization $\boldsymbol{x}$ of the system state $\mathbf{x}$ given in Eq. (12), the inherent hypothesis is that $\boldsymbol{b}$, which is a realization of the bias, is a vector of components given with respect to the coordinate frame parameterized by $\boldsymbol{q}$, which is a realization of the quaternion. Therefore, for a given pair of bias vector realizations, their coordinates are not with respect to a common frame. This observation is also true when considering the true bias $\mathbf{b}$ and its estimate $\hat{\mathbf{b}}$ since these are two particular elements in $\boldsymbol{b}$ from the sample space. The covariance given in Eq. (16) appears to violate this fact since the the difference between $\boldsymbol{e}$ and $\hat{\mathbf{e}}$ appears though their elements are not expressed with respect to a common frame. Further, this expression is integrated over a domain for which each sample $\boldsymbol{e}$ is not, in general, expressed in common coordinates. Each sample $\boldsymbol{e}$ in Eq. (16) should first be rotated into mean coordinates. Likewise, since the true and estimate biases $\mathbf{b}$ and $\hat{\mathbf{b}}$ are column vectors of coordinates given in the true and estimate body frames, respectively, it is reasoned here that the bias error be defined as

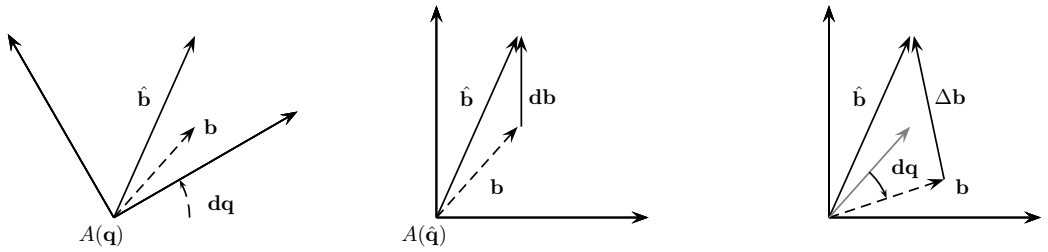
$$\mathbf{db} \triangleq A^T(\mathbf{dq})\mathbf{b} - \hat{\mathbf{b}} \tag{18}$$

where $A(\mathbf{dq}) = A(\mathbf{q})A^T(\hat{\mathbf{q}})$. Every realization of $\mathbf{db}$ written as in Eq. (18) is expressed with respect to the same coordinate basis. The vector error in Eq. (18) as well as integrated quantities such as the covariance of $\mathbf{db}$ in Eq. (16) are well-defined in this respect.

The above rationalization is now supplemented by sketches provided in Figure 1. The coordinate axes of the estimate body frame parameterized by $\hat{\mathbf{q}}$ are indicated by $A(\hat{\mathbf{q}})$ at their origin, and the basis vectors of the true body frame parameterized by $\mathbf{q}$ are marked at their origin by $A(\mathbf{q})$. Figures 1(a) and 1(b) show respectively the true and estimate body frame coordinate axes. The estimate body frame is misaligned from the true body frame by the attitude estimate error $\mathbf{dq}$. Expressing $\mathbf{b}$ and $\hat{\mathbf{b}}$ in common coordinates, either with respect to the true body frame or the estimate body frame, leads to the bias vector error $\mathbf{db}$. Treating the elements $\mathbf{b}$ and $\hat{\mathbf{b}}$ as already having a common basis leads to the bias error defined in Eq. (15). This error is correlated with the attitude error $\mathbf{dq}$ as illustrated in Figure 1(c). The consequences are unintuitive since $\|\Delta\mathbf{b}\|$ implicitly depends on the attitude error $\mathbf{dq}$ while $\|\mathbf{db}\|$ is intrinsically independent of $\mathbf{dq}$.

It is important to emphasize that a vector is not estimated. Rather, components of that vector are estimated. The above issue is then understood to manifest itself when estimating the components of a vector with respect to a frame parameterized by a random variable. In other words, there is an inherent error by associating the true bias vector components with an incorrect coordinate basis. This is exactly what is depicted in Figure 1(c), where components are not associated with a particular frame. This vector error defined in Eq. (15) is referred to as being "unframed."

The objective of this work is to develop a new variant of the EKF that is constructed upon the bias error defined in Eq. (18) in addition to the multiplicative quaternion error employed by the MEKF. The desired result is to overcome both the unconstrained and unframed limitations of the EKF.



(a) **True Body Frame** (b) **Estimate Body Frame** (c) **Unframed**

Figure 1. Illustration of Different Bias Error Representations

IV. Geometric Extended Kalman Filter

In subsequent developments, a Kalman filter is derived for the attitude system presented in Section II. What is fundamentally different is that not only is the filter constructed upon a multiplicative quaternion error as in the MEKF, but it is also developed to employ the bias error defined in Eq. (18). The approach begins with a classical EKF, and then transforms its propagation and update equations to operate on the proposed error definition. Toward this end, the EKF equations are first developed, followed by an appropriate transformation.

The standard EKF approximates the local dynamics with the truncated Taylor series expansion, given by

$$\dot{\mathbf{x}} \approx \mathbf{f}(\hat{\mathbf{x}}) + F_a \Delta\mathbf{x} + G_a \mathbf{w} \tag{19}$$

which is expanded about the approximate conditional mean $\hat{\mathbf{x}}$ (and about the mean process noise vector, $\hat{\mathbf{w}} = \mathbf{0}$). In Eq. (19), the standard EKF error $\Delta\mathbf{x}$ is defined by the concatenation of Eqs. (14) and (15) and $\mathbf{f}(\hat{\mathbf{x}})$ represents the concatenation of the state dynamics of Eq. (11). From the attitude dynamics in Eq. (11), the linearized system matrices of Eq. (19) are given by

$$F_a(t) = \begin{bmatrix} \frac{1}{2}\Omega(\tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}}) & -\frac{1}{2}\Xi(\hat{\mathbf{q}}) \\ 0_{3\times4} & 0_{3\times3} \end{bmatrix} \tag{20}$$

and

$$G_a(t) = \begin{bmatrix} -\frac{1}{2}\Xi(\hat{\mathbf{q}}) & 0_{4\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{21}$$

The algorithm resulting from direct implementation of the propagation in Eqs. (19)−(21) with a classically derived Kalman update is referred to as the *Additive* EKF in attitude estimation. This is because in Eq. (19), the unit norm quaternion constraint is relaxed and quaternion error is defined additively.

Ideally, the appropriate filter would employ state errors defined by

$$\mathbf{dq} \triangleq \mathbf{q} \otimes \hat{\mathbf{q}}^{-1} \equiv \begin{bmatrix} \mathbf{d}\boldsymbol{\varrho}^T & \mathrm{d}q_4 \end{bmatrix}^T \tag{22a}$$

$$\mathbf{db} \triangleq A^T(\mathbf{dq})\mathbf{b} - \hat{\mathbf{b}} \tag{22b}$$

where $\otimes$ indicates quaternion multiplication [1], and all realizations of $\mathbf{db}$ are expressed within the mean coordinate frame. When the attitude error $\mathbf{dq}$ is small, the rotation matrix in Eq. (22b) can be approximated by [13]

$$A(\mathbf{d}\boldsymbol{\alpha}) \approx I_{3\times3} - [\mathbf{d}\boldsymbol{\alpha}\times] \tag{23}$$

with the small angle vector $\mathbf{d}\boldsymbol{\alpha} \approx 2\mathbf{d}\boldsymbol{\varrho}$.

A relationship between the error defined in Eq. (22) and $\Delta\mathbf{x}$ can be found. Because the series expansions deriving the EKF are truncated to first-order, a linearized approximation is adequate since higher order terms will ultimately be discarded. Thus, the error definitions are related according to

$$\mathbf{q} - \hat{\mathbf{q}} \approx \frac{1}{2}\Xi(\hat{\mathbf{q}})\mathbf{d}\boldsymbol{\alpha} \tag{24a}$$

$$\mathbf{b} - \hat{\mathbf{b}} \approx [\hat{\mathbf{b}}\times]\mathbf{d}\boldsymbol{\alpha} + \mathbf{d}\mathbf{b} \tag{24b}$$

where the approximation to the attitude matrix in Eq. (23) is used in obtaining Eq. (24b). Equation (24a) is recognized as the mapping employed by the "reduced covariance" approach to deriving the MEKF [1]. Assembling the components of Eq. (24) leads to

$$\Delta\mathbf{x} \approx C(t)\mathbf{d}\mathbf{x} \tag{25}$$

where the "error map" C is defined according to

$$C(t) \triangleq \begin{bmatrix} \frac{1}{2}\Xi(\hat{\mathbf{q}}) & 0_{4\times 3} \\ [\hat{\mathbf{b}}\times] & I_{3\times 3} \end{bmatrix} \tag{26}$$

and the reduced order, or geometric, error is given by

$$\mathbf{d}\mathbf{x} \triangleq \begin{bmatrix} \mathbf{d}\boldsymbol{\alpha} \\ \mathbf{d}\mathbf{b} \end{bmatrix} \tag{27}$$

In the following subsections, the error map C is employed to transform the EKF propagation and update equations.

A. Prediction

To develop the GEKF prediction equations, return first to the linearized dynamics of Eq. (19). The standard EKF error definition $\Delta\mathbf{x}$ can be replaced using the relationship of Eq. (25). The state dynamics become

$$\dot{\mathbf{x}} = \mathbf{f}(\hat{\mathbf{x}}) + F_a C\mathbf{d}\mathbf{x} + G_a\mathbf{w} \tag{28}$$

Following classical developments of the EKF, the dynamics in Eq. (28) lead to

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}) \tag{29}$$

This is accomplished first by integration then taking the expectation of the result. Interchanging the integral and expectation operators, and differentiating [14] results in Eq. (29). Equation (29) describes the conditional mean trajectory $\hat{\mathbf{x}}(t)$ to within first-order. As with any EKF, it must be emphasized that $\hat{\mathbf{x}}$ *approximates* the conditional mean. This must not be confused with the linear filtering problem where this notation denotes the exact conditional mean.

Differentiating Eq. (25) produces a relationship between the dynamics of $\Delta\mathbf{x}$ and those of $\mathbf{d}\mathbf{x}$, which is given by

$$\dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} = \dot{C}\mathbf{d}\mathbf{x} + C\mathbf{d}\dot{\mathbf{x}} \tag{30}$$

Substituting Eqs. (28) and (29) into Eq. (30) leads to

$$C\mathbf{d}\dot{\mathbf{x}} = (F_a C - \dot{C})\mathbf{d}\mathbf{x} + G_a\mathbf{w} \tag{31}$$

which is a first-order differential equation for the geometric error $\mathbf{d}\mathbf{x}$. The system in Eq. (31) is overdetermined and therefore does not, in general have a solution. If it does, pre-multiplying by C^T and assuming that the inverse of $C^T C$ exists leads to

$$\mathbf{d}\dot{\mathbf{x}} = (C^T C)^{-1}C^T(F_a C - \dot{C})\mathbf{d}\mathbf{x} + (C^T C)^{-1}C^T G_a\mathbf{w} \tag{32}$$

While $\mathbf{d\dot{x}}$ satisfies the resulting system in Eq. (32), it does not necessarily satisfy the original system in Eq. (31). This must be verified before proceeding further. Substituting Eq. (32) into Eq. (31) and gathering terms leads to

$$\left[I_{7\times 7} - C(C^TC)^{-1}C^T\right]\left[(F_aC - \dot{C})\mathbf{dx} + G_a\mathbf{w}\right] = \mathbf{0}_{7\times 1} \tag{33}$$

where $\mathbf{0}_{7\times 1}$ is a 7×1 vector of zeros. In other words, defining

$$S \triangleq \left[I_{7\times 7} - C(C^TC)^{-1}C^T\right] \tag{34a}$$

$$\mathbf{z} \triangleq (F_aC - \dot{C})\mathbf{dx} + G_a\mathbf{w} \tag{34b}$$

it is seen that Eq. (33) states that $\mathbf{z}$ must belong to the null space of S. This is verified in what follows.

First, using the identity [15]

$$\frac{d}{dt}\Xi(\mathbf{q}) = \frac{1}{2}\Omega(\boldsymbol{\omega})\Xi(\mathbf{q}) + \Xi(\mathbf{q})[\boldsymbol{\omega}\times] \tag{35}$$

and the fact that $\dot{\hat{\mathbf{b}}} = \mathbf{0}$ from Eq. (29) leads to

$$\dot{C} = \begin{bmatrix} \frac{1}{4}\Omega(\hat{\boldsymbol{\omega}})\Xi(\hat{\mathbf{q}}) + \frac{1}{2}\Xi(\hat{\mathbf{q}})[\hat{\boldsymbol{\omega}}\times] & 0_{4\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix} \tag{36}$$

It is also verified by direct calculation that

$$(C^TC)^{-1} = 4\begin{bmatrix} I_{3\times 3} & -[\hat{\mathbf{b}}\times]^T \\ -[\hat{\mathbf{b}}\times] & \left(\frac{1}{4}I_{3\times 3} + [\hat{\mathbf{b}}\times][\hat{\mathbf{b}}\times]^T\right) \end{bmatrix} \tag{37}$$

As shown by Eq. (37), the matrix C^TC is invertible for all $\hat{\mathbf{x}}$, and therefore the appearance of its inverse in Eqs. (32) and (34) is justified. Substituting Eqs. (20), (21), (26), (36) and (37) into Eq. (34) it is found that

$$S = \begin{bmatrix} \mathbf{q}\mathbf{q}^T & 0_{4\times 3} \\ 0_{3\times 4} & 0_{3\times 3} \end{bmatrix} \tag{38a}$$

$$\mathbf{z} = \begin{bmatrix} \frac{1}{2}\Xi(\hat{\mathbf{q}})\big(\mathbf{db} - [\tilde{\boldsymbol{\omega}}\times]\mathbf{d}\boldsymbol{\alpha} - \boldsymbol{\eta}_v\big) \\ \boldsymbol{\eta}_u \end{bmatrix} \tag{38b}$$

As a result, $\mathbf{z}$ belonging to the null space of S translates to $\mathbf{z}$ being orthogonal to $[\hat{\mathbf{q}}^T \ \ \mathbf{0}_{1\times 3}]$. This is easily verified to be the case since $\Xi^T(\hat{\mathbf{q}})\hat{\mathbf{q}} = \mathbf{0}_{3\times 1}$. Therefore, the solution of Eq. (32) does indeed also satisfy Eq. (31). This justifies the pseudoinverse appearing in Eq. (32). The trajectory obtained by projecting the four-component attitude error satisfying Eq. (31) onto the unit hypersphere tangent plane evolves according to Eq. (32).

To complete the covariance propagation derivation, the coefficient matrices in Eq. (32) are simplified. Substituting Eqs. (20), (21), (36) and (37) into Eq. (32) while using the definition of C from Eq. (26) yields the error dynamics

$$\mathbf{d\dot{x}} = F_g\mathbf{dx} + G_g\mathbf{w} \tag{39}$$

The system matrices in Eq. (39) are defined according to

$$F_g(t) \triangleq \begin{bmatrix} -[\tilde{\boldsymbol{\omega}}\times] & -I_{3\times 3} \\ [\hat{\mathbf{b}}\times][\tilde{\boldsymbol{\omega}}\times] & [\hat{\mathbf{b}}\times] \end{bmatrix} \tag{40}$$

and

$$G_g(t) \triangleq \begin{bmatrix} -I_{3\times 3} & 0_{3\times 3} \\ [\hat{\mathbf{b}}\times] & I_{3\times 3} \end{bmatrix} \tag{41}$$

Equation (39) has the same form as the $\Delta\mathbf{x}$ dynamics of the standard EKF. The same developments therefore show that the geometric error covariance trajectory $\mathcal{P}(t)$ is governed by the differential equation [14]

$$\dot{\mathcal{P}} = F_g\mathcal{P} + \mathcal{P}F_g^T + G_gQG_g^T \tag{42}$$

Equations (29) and (42) comprise the prediction stage of the GEKF algorithm.

There are two important points to make concerning the above filter propagation equations. First, if the bias vector error were instead chosen to be expressed in true coordinates according to

$$\mathbf{db} = \mathbf{b} - A(\mathbf{dq})\hat{\mathbf{b}} \tag{43}$$

it can be verified that the same equations result. This is only true to first order when all higher order terms are discarded. Second, the covariance propagation can be derived using the multiplicative approach followed in developing the MEKF [1]. This is demonstrated in the Appendix. The same result is obtained without the pseudoinverse that appears in Eq. (32), however a truncation is applied to the state error vector. In this respect, the truncation performs the same projection that the pseudoinverse accomplishes. This truncation in the multiplicative derivation is a special case of the above derivation obtained by projection.

B. Update

The update stage for the GEKF is developed in the same spirit as is its prediction stage in the previous subsection. Transforming the EKF update equations yields a modified procedure constructed upon the error definition $\mathbf{dx}$ defined in Eq. (27). It is important to acknowledge both the *a priori* and *a posteriori* error definition maps. In other words, the relationship in Eq. (24) has a discontinuity at t_k when an update occurs.

Consider a vector output model given by

$$\mathbf{h}_k(\mathbf{x}_k) = A(\mathbf{q}_k)\mathbf{r}_k \tag{44}$$

where $\mathbf{r}_k$ is the k^{th} available reference vector. Employing an additive update like the EKF, the unbiased state update is found to be [14]

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + K_k\big[\mathbf{y}_k - \hat{\mathbf{h}}_k(\mathbf{x}_k)\big] \tag{45}$$

where the available measurement $\mathbf{y}_k$ is

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k \tag{46}$$

with zero-mean, Gaussian noise $\mathbf{v}_k$. Expanding $\mathbf{h}_k(\mathbf{x}_k)$ in a Taylor series about the *a priori* state estimate $\hat{\mathbf{x}}_k^-$ and truncating to first-order leads to

$$\mathbf{h}_k(\mathbf{x}_k) \approx \mathbf{h}_k(\hat{\mathbf{x}}_k^-) + H_k C_k^- \mathbf{dx}_k^- \tag{47}$$

where $C_k^- \triangleq C(t_k^-)$ and $\mathbf{dx}_k^-$ denote the *a priori* error map and state error, respectively.

The measurement residual does not follow the same logic as that employed to define Eq. (22b). The bias vector is a random variable and as argued in Section III, each realization from the sample space in general has components given with respect to different coordinate axes as a result of the attitude also being distributed. On the other hand, both the predicted output and vector measurement are deterministic. The residual is non-zero because of the noise $\mathbf{v}_k$ and more importantly because of the implicit hypothesis that the estimated body frame is the true body frame.

An observation of the form in Eq. (44) is treated as an abstract vector with coordinates measured in the true body frame. Error is therefore angular and manifests itself by associating the component vector with the incorrect coordinate basis. Further, because the true frame is not known, a vector with components expressed with respect to the true basis could not be correctly expressed with respect to another coordinate frame.

Using the identity [15]

$$\frac{\partial}{\partial \mathbf{q}} A(\mathbf{q})\mathbf{r} = 2[A(\mathbf{q})\mathbf{r}\times]\Xi^T(\mathbf{q}) \tag{48}$$

the measurement sensitivity matrix H_k in Eq. (47) is determined to be

$$H_k = \left.\frac{\partial \mathbf{h}_k}{\partial \mathbf{x}_k}\right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k^-} = \Big[2[A(\hat{\mathbf{q}}_k^-)\mathbf{r}\times]\Xi^T(\hat{\mathbf{q}}_k^-) \quad 0_{3\times 3}\Big] \tag{49}$$

Substituting Eqs. (46) and (47) into Eq. (45) leads to

$$\hat{\mathbf{x}}_k^+ = \hat{\mathbf{x}}_k^- + K_k H_k C_k^- \mathbf{dx}_k^- + K_k \mathbf{v}_k \tag{50}$$

Then, both sides of Eq. (50) are subtracted from $\mathbf{x}_k$ to obtain

$$C_k^+ \mathbf{dx}_k^+ = [I_{7\times 7} - K_k H_k] C_k^- \mathbf{dx}_k^- - K_k \mathbf{v}_k \tag{51}$$

with $C_k^+ \triangleq C(t_k^+)$. Performing the outer product of Eq. (51) with itself, and calculating the expectation of the result gives

$$C_k^+ \mathcal{P}_k^+ [C_k^+]^T = [I_{7\times 7} - K_k H_k] C_k^- \mathcal{P}_k^- [C_k^-]^T [I_{7\times 7} - K_k H_k]^T + K_k R_k K_k^T \tag{52}$$

where $\mathcal{P}$ denotes the state error covariance and $R_k \triangleq E\{\mathbf{v}_k \mathbf{v}_k^T\}$ is the measurement error covariance. When the observation in Eq. (44) is of unit length, the QUEST Measurement Model is often employed, which leads to a singular covariance matrix. This however can effectively be replaced with a nonsingular covariance matrix [16].

To determine the gain K_k, the objective function

$$J(K_k) \triangleq \mathrm{Tr}\big(C_k^+ \mathcal{P}_k^+ [C_k^+]^T\big) \tag{53}$$

is minimized. Ideally, as in the EKF, it would be preferred to minimize $\mathrm{Tr}\big(\mathcal{P}_k^+\big)$. This corresponds to minimizing the mean square error. Since the expression for $\mathcal{P}_k^+$ depends on the as of yet unknown updated state vector $\hat{\mathbf{x}}_k^+$, the objective function is chosen out of convenience.

An analysis on the differences between minimizing the trace of $C_k^+ \mathcal{P}_k^+ [C_k^+]^T$ and the trace of $\mathcal{P}_k^+$ is now shown. For this analysis the subscript k and superscript $+$ are not needed, and thus removed. Partition the matrix $\mathcal{P}$ into its 3×3 sub-matrices:

$$\mathcal{P} \triangleq \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{12}^T & \mathcal{P}_{22} \end{bmatrix} \tag{54}$$

where $\mathcal{P}_{11}$ corresponds to the attitude error portions, and $\mathcal{P}_{22}$ corresponds to the gyro-bias error portions. The matrix $C\,\mathcal{P}\,C^T$ is given by

$$C\,\mathcal{P}\,C^T \triangleq \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^T & Z_{22} \end{bmatrix} \tag{55}$$

where

$$Z_{11} = \frac{1}{4} \Xi(\mathbf{q})\, \mathcal{P}_{11}\, \Xi^T(\mathbf{q}) \tag{56a}$$

$$Z_{12} = \frac{1}{2} \Xi(\mathbf{q}) \left(\mathcal{P}_{11}\, [\mathbf{b}\times]^T + \mathcal{P}_{12}\right) \tag{56b}$$

$$Z_{22} = [\mathbf{b}\times]\, \mathcal{P}_{11}\, [\mathbf{b}\times]^T + \mathcal{P}_{12}^T\, [\mathbf{b}\times]^T + [\mathbf{b}\times]\, \mathcal{P}_{12} + \mathcal{P}_{22} \tag{56c}$$

The trace of $C\,\mathcal{P}\,C^T$ is given by the sum of the eigenvalues of the matrices Z_{11} and Z_{22}. The sum of the eigenvalues of Z_{11} is equivalent to within a constant scalar multiple of the sum of the eigenvalues of $\mathcal{P}_{11}$, so the operation $C\,\mathcal{P}\,C^T$ does not affect the attitude error portion of $\mathcal{P}$. Specifically, minimizing the trace of Z_{11} is exactly equivalent to minimizing the trace of $\mathcal{P}_{11}$.

This, however, is not true for the gyro-bias portions associated with Z_{22}. Here, some cases are shown to assess when this is true for the gyro-bias portions. This is accomplished using Farrenkopf's analysis, which involves single-axis attitude estimation [12]. Consider the case of a gyro with the following spectral densities: $\sigma_u = \sqrt{10} \times 10^{-10}\,\mathrm{rad/s^{3/2}}$ and $\sigma_v = \sqrt{10} \times 10^{-7}\,\mathrm{rad/s^{1/2}}$. Assuming a sampling interval of 10 s and a 1 deg standard deviation for the attitude error measurement gives the following covariance from Farrenkopf's analysis:

$$\mathcal{P} = \begin{bmatrix} 3.2638 \times 10^{-7} & -1.7444 \times 10^{-11} \\ -1.7444 \times 10^{-11} & 1.8705 \times 10^{-15} \end{bmatrix} \tag{57}$$

Table 1. Single-Axis Covariance Analysis for Various Biases

Bias b (deg/h)	$b^2\mathcal{P}_{11} + 2b\mathcal{P}_{12}$	$\mathcal{P}_{22}$
0.01	-1.6925×10^{-18}	1.8705×10^{-15}
0.1	-1.6855×10^{-17}	1.8705×10^{-15}
1	-1.6164×10^{-16}	1.8705×10^{-15}
5	-6.5462×10^{-16}	1.8705×10^{-15}

Table 1 compares the scalar version of the first three elements of Z_{22}, given by $b^2\mathcal{P}_{11} + 2b\mathcal{P}_{12}$, to $\mathcal{P}_{22}$ for various values of the scalar bias, denoted by b. The first three elements of Z_{22} are about the same order of magnitude as $\mathcal{P}_{22}$ when a bias of around 5 deg/h is given. This is the limit where minimizing the trace of Z_{22} is effectively equivalent to minimizing the trace of $\mathcal{P}_{22}$ in a scalar sense. However, in all these scenarios the attitude error variance is at least seven orders of magnitude greater than the gyro-bias error variance. Hence, the attitude error portions dominate the trace of both $C\,\mathcal{P}\,C^T$ and $\mathcal{P}$. This analysis strongly indicates that minimizing the trace of $C\,\mathcal{P}\,C^T$ is effectively equivalent to minimizing the trace of $\mathcal{P}$.

Minimizing the objective in Eq. (53) yields a gain matrix of

$$K_k = C_k^- \mathcal{P}_k^- [C_k^-]^T H_k^T \left[H_k C_k^- \mathcal{P}_k^- [C_k^-]^T H_k^T + R_k \right]^{-1} \tag{58}$$

It is to be noted that

$$\nabla_K^2 J(K_k) = 2 \left[R_k + H_k^- C_k^- \mathcal{P}_k^- [C_k^-]^T [H_k^-]^T \right] \tag{59}$$

which is positive definite for $R_k > 0$ and therefore Eq. (53) is indeed minimized. Finally, letting $\bar{H}_k \triangleq H_k C_k^-$ and $\bar{K}_k \triangleq \mathcal{P}_k^- \bar{H}_k^T [\bar{H}_k \mathcal{P}_k^- \bar{H}_k + R_k]^{-1}$ allows Eq. (52) to be written as

$$\mathcal{P}_k^+ = \bar{M}_k \left\{ [I_{6\times 6} - \bar{K}_k \bar{H}_k] \mathcal{P}_k^- [I_{6\times 6} - \bar{K}_k \bar{H}_k]^T + \bar{K}_k R_k \bar{K}_k^T \right\} \bar{M}_k^T \tag{60}$$

where the transformation $\bar{M}_k$ is given by

$$\bar{M}_k \triangleq \left([C_k^+]^T C_k^+ \right)^{-1} [C_k^+]^T C_k^- = \begin{bmatrix} \Xi^T(\hat{\mathbf{q}}_k^+) \Xi(\hat{\mathbf{q}}_k^-) & 0_{3\times 3} \\ [\hat{\mathbf{b}}_k^- \times] - [\hat{\mathbf{b}}_k^+ \times] \Xi^T(\hat{\mathbf{q}}_k^+) \Xi(\hat{\mathbf{q}}_k^-) & I_{3\times 3} \end{bmatrix} \tag{61}$$

Note that as $\hat{\mathbf{x}}_k^-$ approaches $\hat{\mathbf{x}}_k^+$, the transformation $\bar{M}_k$ approaches identity. Equations Eqs. (45) and (60) define the GEKF update stage. The GEKF for attitude estimation is summarized in Table 2, where $\tilde{\mathbf{y}}_k$ is vector that contains all of the concatenated body-frame observations.

C. Discussion

The new developments have several advantages. With regards to the attitude state $\mathbf{q}$, two points must be emphasized. Firstly, the predicted attitude estimate is identified with the EKF approximate conditional mean. An expectation is taken in order to determine the conditional mean dynamics in Eq. (29). Alternate derivations describe $\hat{\mathbf{q}}$ as a "reference attitude" from which an unconstrained, minimally parameterized attitude error is drawn [8,15]. The reference attitude interpretation deviates from the theoretical derivation of the EKF. As it turns out, this is of little importance in an EKF since to within first-order, both approaches yield the same approximate conditional mean dynamics. The advantage is therefore only academic.

Secondly, both the state and state error covariance update equations acknowledge the non-uniform map from the multiplicative error $\mathbf{dq}$ to the additive error $\Delta\mathbf{q}$ employed in the filter by construction. The state estimate update in Table 2 is seen to be composed of two steps. First, the gain $\bar{K}_k$ multiplies the measurement residual producing a multiplicative error approximated to first order in the hyperplane tangent to $\hat{\mathbf{q}}_k^-$. This error is then transformed from $\mathbf{d}\boldsymbol{\alpha}$ space to $\Delta\mathbf{q}$ space, through multiplication by C_k^-, where then the additive update is performed to obtain $\hat{\mathbf{q}}_k^+$.

Unlike the MEKF accepted in practice [1,17], the GEKF accounts for the non-uniformity of the multiplicative-additive quaternion error map over the unit hypersphere. The fact that mapping quaternion multiplicative

Table 2. Geometric Extended Kalman Filter for Attitude Estimation

Model	$\dot{\mathbf{q}} = \frac{1}{2}\Xi(\mathbf{q})(\tilde{\boldsymbol{\omega}} - \mathbf{b} - \boldsymbol{\eta}_v), \qquad E\{\boldsymbol{\eta}_v(t)\boldsymbol{\eta}_v^T(\tau)\} = \sigma_v^2 I_{3\times3}\delta(t-\tau)$ $\dot{\mathbf{b}} = \boldsymbol{\eta}_u, \qquad E\{\boldsymbol{\eta}_u(t)\boldsymbol{\eta}_u^T(\tau)\} = \sigma_u^2 I_{3\times3}\delta(t-\tau)$ $\tilde{\mathbf{y}}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \qquad \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, R_k)$
Initialize	$\hat{\mathbf{q}}(t_0) = \hat{\mathbf{q}}_0, \quad \hat{\mathbf{b}}(t_0) = \hat{\mathbf{b}}_0$ $\mathcal{P}(t_0) = E\{\mathbf{d}\mathbf{x}_0\mathbf{d}\mathbf{x}_0^T\}$
Gain	$\bar{K}_k = \mathcal{P}_k^- \bar{H}_k^T[\bar{H}_k\mathcal{P}_k^-\bar{H}_k^T + R_k]^{-1}$ $\bar{H}_k = H_k C_k^-$ $H_k = \begin{bmatrix} 2\left[A(\hat{\mathbf{q}}_k^-)\mathbf{r}_{k1}\times\right]\Xi^T(\hat{\mathbf{q}}_k^-) & 0_{3\times3} \\ \vdots & \vdots \\ 2\left[A(\hat{\mathbf{q}}_k^-)\mathbf{r}_{kn}\times\right]\Xi^T(\hat{\mathbf{q}}_k^-) & 0_{3\times3} \end{bmatrix}$ $C_k^- = \begin{bmatrix} \frac{1}{2}\Xi(\hat{\mathbf{q}}_k^-) & 0_{4\times3} \\ [\hat{\mathbf{b}}_k^-\times] & I_{3\times3} \end{bmatrix}$
Update	$\begin{bmatrix} \hat{\mathbf{q}}_k^+ \\ \hat{\mathbf{b}}_k^+ \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{q}}_k^- \\ \hat{\mathbf{b}}_k^- \end{bmatrix} + C_k^-\bar{K}_k\left[\tilde{\mathbf{y}}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k^-)\right]$ $\hat{\mathbf{q}}_k^+ \leftarrow \hat{\mathbf{q}}_k^+ / \|\hat{\mathbf{q}}_k^+\|$ $\mathbf{h}_k(\hat{\mathbf{x}}_k^-) = \begin{bmatrix} A(\hat{\mathbf{q}}_k^-)\mathbf{r}_{k1} \\ \vdots \\ A(\hat{\mathbf{q}}_k^-)\mathbf{r}_{kn} \end{bmatrix}$ $\mathcal{P}_k^+ = \bar{M}_k\left\{[I_{6\times6} - \bar{K}_k\bar{H}_k]\mathcal{P}_k^-[I_{6\times6} - \bar{K}_k\bar{H}_k]^T + \bar{K}_k R_k \bar{K}_k^T\right\}\bar{M}_k^T$ $\bar{M}_k = \begin{bmatrix} \Xi^T(\hat{\mathbf{q}}_k^+)\Xi(\hat{\mathbf{q}}_k^-) & 0_{3\times3} \\ [\hat{\mathbf{b}}_k^-\times] - [\hat{\mathbf{b}}_k^+\times]\Xi^T(\hat{\mathbf{q}}_k^+)\Xi(\hat{\mathbf{q}}_k^-) & I_{3\times3} \end{bmatrix}$
Propagation	$\dot{\hat{\mathbf{q}}} = \frac{1}{2}\Xi(\hat{\mathbf{q}})\left(\tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}}\right)$ $\dot{\hat{\mathbf{b}}} = \mathbf{0}$ $\dot{\mathcal{P}} = F_g\mathcal{P} + \mathcal{P}F_g^T + G_g Q G_g^T$ $F_g(t) = \begin{bmatrix} -[\tilde{\boldsymbol{\omega}}\times] & -I_{3\times3} \\ [\hat{\mathbf{b}}\times][\tilde{\boldsymbol{\omega}}\times] & [\hat{\mathbf{b}}\times] \end{bmatrix}, \quad G_g(t) = \begin{bmatrix} -I_{3\times3} & 0_{3\times3} \\ [\hat{\mathbf{b}}\times] & I_{3\times3} \end{bmatrix}$

error to additive error depends on the pole $\hat{\mathbf{q}}_k^-$ suggests that a given error at one attitude is not numerically equivalent to that attitude error for another absolute attitude. As a result, the *a priori* error covariance $\mathcal{P}_k^-$ represented in $\mathbf{d}\boldsymbol{\alpha}_k^-$ space must be updated according to the standard EKF update with gain $\bar{K}_k$ and then transformed to $\mathbf{d}\boldsymbol{\alpha}_k^+$ space through the transformation $\bar{M}_k$. Other recent work has also acknowledged this fact [18]. It is argued in the literature that this transformation can be neglected since when $\hat{\mathbf{q}}_k^+ \approx \hat{\mathbf{q}}_k^-$, the upper-left element in $\bar{M}_k$ approaches identity [19]. Caution should be taken however because a large initial attitude error covariance implies poor attitude knowledge. An accurate attitude observation would produce a high update gain. A significant update would follow which is incompatible with the assumption that $\hat{\mathbf{q}}_k^+ \approx \hat{\mathbf{q}}_k^-$.

The GEKF is a rigorous adaptation preserving the foundation of the EKF. The system equations are linearized in order to predict an approximate conditional mean and error covariance as well as to perform a linear, additive update. Constraints and common frame vector errors are maintained to within first-order. For quaternion error which is quantified by the components of its projection onto a tangent hyperplane, the dynamics and non-uniformity of this projection are respected to within first-order. Truncating the dynamics in this case are found to not be equivalent to projected and then linearized dynamics. This fact is acknowledged by the covariance transformation $\bar{M}_k$.

D. Discrete-Time Attitude Estimation

In application, it is advantageous to operate in discrete time. A discrete time formulation of the MEKF is well-established [17], and it is desire that one be available for the GEKF. The state and error covariance can be propagated using numerical integration techniques. This can however be computationally taxing. Fortunately, an approximate analytical discrete-time solution can be derived. The discrete error state transition matrix can be derived using a power series approach [15]. Rather than follow this method directly, note that

$$F_g = T^{-1} F_{MEKF} T \tag{62}$$

where F_{MEKF} is the matrix

$$F_{MEKF} = \begin{bmatrix} -[(\tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}})\times] & -I_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix} \tag{63}$$

and the time-dependent transformation T is defined according to

$$T \triangleq \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \\ [\hat{\mathbf{b}}\times] & I_{3\times 3} \end{bmatrix} \tag{64}$$

The matrix F_{MEKF} arises in the MEKF linearized error state dynamics whose discrete-time equations are already developed [17]. Given Eq. (62) and the properties of square matrices [20], the GEKF error state transition matrix Φ_k can be written as

$$\Phi_k = T^{-1} \Phi_{MEKF} T \tag{65}$$

where Φ_{MEKF} is the MEKF state transition matrix and the state estimate $\hat{\mathbf{x}}$ is assumed constant over the sampling interval. The inverse of T is given by

$$T^{-1} = \begin{bmatrix} I_{3\times 3} & 0_{3\times 3} \\ [\hat{\mathbf{b}}\times]^T & I_{3\times 3} \end{bmatrix} \tag{66}$$

The MEKF error state transition matrix is given by [17]

$$\Phi_{MEKF} = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ 0_{3\times 3} & I_{3\times 3} \end{bmatrix} \tag{67}$$

with elements defined as

$$\Phi_{11} = I_{3\times 3} - [\hat{\boldsymbol{\omega}}\times]\frac{\sin(\|\hat{\boldsymbol{\omega}}\|\Delta t)}{\|\hat{\boldsymbol{\omega}}\|} + [\hat{\boldsymbol{\omega}}\times]^2 \frac{\{1-\cos(\|\hat{\boldsymbol{\omega}}\|\Delta t)\}}{\|\hat{\boldsymbol{\omega}}\|^2} \tag{68a}$$

$$\Phi_{12} = [\hat{\boldsymbol{\omega}}\times]\frac{\{1-\cos(\|\hat{\boldsymbol{\omega}}\|\Delta t)\}}{\|\hat{\boldsymbol{\omega}}\|^2} - I_{3\times 3}\Delta t - [\hat{\boldsymbol{\omega}}\times]^2 \frac{\{\|\hat{\boldsymbol{\omega}}\|\Delta t - \sin(\|\hat{\boldsymbol{\omega}}\|\Delta t)\}}{\|\hat{\boldsymbol{\omega}}\|^3} \tag{68b}$$

and where Δt is the propagation time interval.

The matrix G_g is also related to its MEKF counterpart G_{MEKF} according to

$$G_g = T^{-1} G_{MEKF} \tag{69}$$

This relationship can be leveraged since the spectral density matrix conversion to a discrete-time covariance for the MEKF is also available. It is approximately given by

$$\mathcal{Q}_{MEKF} = \begin{bmatrix} \left(\sigma_v^2 \Delta t + \frac{1}{3}\sigma_u^2 \Delta t^3\right) I_{3\times 3} & \left(\frac{1}{2}\sigma_u^2 \Delta t^2\right) I_{3\times 3} \\ \left(\frac{1}{2}\sigma_u^2 \Delta t^2\right) I_{3\times 3} & \left(\sigma_u^2 \Delta t\right) I_{3\times 3} \end{bmatrix} \tag{70}$$

under the assumption that the sampling rate is below Nyquist's limit [15]. A complete closed-form solution can be used if this condition is not true [15]. The GEKF discrete error is propagated according to

$$\mathcal{P}_{k+1}^{-} = \Phi_k \mathcal{P}_k^{+} \Phi_k^T + \mathcal{Q}_k \tag{71}$$

where

$$\mathcal{Q}_k \triangleq T_k \mathcal{Q}_{MEKF} T_k^T \tag{72}$$

An alternative approach given by Van Loan provides a numerical solution for the discrete-time model [17,21]. However, this requires a numerical matrix exponential which can be quite costly. The discrete-time system using Eqs. (71) and (72) has an advantage in this respect.

The quaternion estimates can also be propagated using a discrete-time model, given by

$$\hat{\mathbf{q}}_{k+1}^{-} = \left[\cos\left(\frac{1}{2}\|\hat{\boldsymbol{\omega}}_k^{+}\|\Delta t\right) I_{4\times 4} + \Omega(\hat{\boldsymbol{\psi}}_k^{+})\right] \hat{\mathbf{q}}_k^{+} \tag{73}$$

where

$$\hat{\boldsymbol{\psi}}_k^{+} \equiv \frac{\sin\left(\frac{1}{2}\|\hat{\boldsymbol{\omega}}_k^{+}\|\Delta t\right)}{\|\hat{\boldsymbol{\omega}}_k^{+}\|} \hat{\boldsymbol{\omega}}_k^{+} \tag{74}$$

and $\hat{\boldsymbol{\omega}}_k^{+} = \tilde{\boldsymbol{\omega}}_k - \hat{\mathbf{b}}_k^{+}$. The bias estimates are propagated simply through $\hat{\mathbf{b}}_{k+1}^{-} = \hat{\mathbf{b}}_k^{+}$.

V. Numerical Results

This section presents results obtained by implementing the MEKF and GEKF for spacecraft attitude estimation. Multiple experiments are considered in order to emphasize any differing performance. Subsection A examines extreme test cases. Subsection B considers a normalized error square (NES) test is to evaluate filter consistency [10].

A. Simulations

These simulations are performed using an Earth-pointing spacecraft with Keplerian orbital parameters given by

$$\begin{aligned} e &= 0.0001353 & i &= 0.6102090\,(\text{rad}) \\ m_0 &= 6.0868\,(\text{rad}) & \omega &= 4.6551753\,(\text{rad}) \\ a &= 6777.2090\,(\text{km}) & \Omega &= 4.5264800\,(\text{rad}) \end{aligned} \tag{75}$$

where e is the orbit eccentricity, m_0 the mean anomaly at epoch, a the semi-major axis, i the orbital plane inclination, ω the argument of perigee and Ω the right ascension of the ascending node. The mean anomaly m_0 is given with epoch Oct. 21, 2015 at 16:29:00. The ephemeris in Eq. (75) corresponds to a near-circular, 90-minute orbit at an altitude of 350 km. This geometry is based on the orbit for the Tropical Rainfall Measuring Mission [22]. At nominal operation, in order to maintain Earth-pointing, the spacecraft rotates clockwise once per orbit about its $y-$axis. This yields a rotation rate of 236 deg/h or about 0.00113 rad/s.

To determine the initial vehicle attitude, the body $z-$axis is chosen as the Earth-pointing axis which initially points along $\boldsymbol{\ell}_z = -\mathbf{r}_0/\|\mathbf{r}_0\|$ where $\mathbf{r}_0$ is the initial vehicle position. The body $y-$axis is oriented along the negative orbital momentum vector $\boldsymbol{\ell}_y = -(\mathbf{r}_0 \times \mathbf{v}_0)/\|\mathbf{r}_0 \times \mathbf{v}_0\|$ where $\mathbf{v}_0$ is the vehicle initial velocity vector. As a result, the initial body attitude matrix is given by

$$A(\mathbf{q}_0) = \begin{bmatrix} (\boldsymbol{\ell}_y \times \boldsymbol{\ell}_z) & \boldsymbol{\ell}_y & \boldsymbol{\ell}_z \end{bmatrix}^T \tag{76}$$

The attitude matrix in Eq. (76) corresponds to the initial quaternion

$$\mathbf{q}_0 = \begin{bmatrix} 0.2063 & -0.4244 & 0.7144 & -0.5167 \end{bmatrix}^T \tag{77}$$

which is used to initialize the simulation.

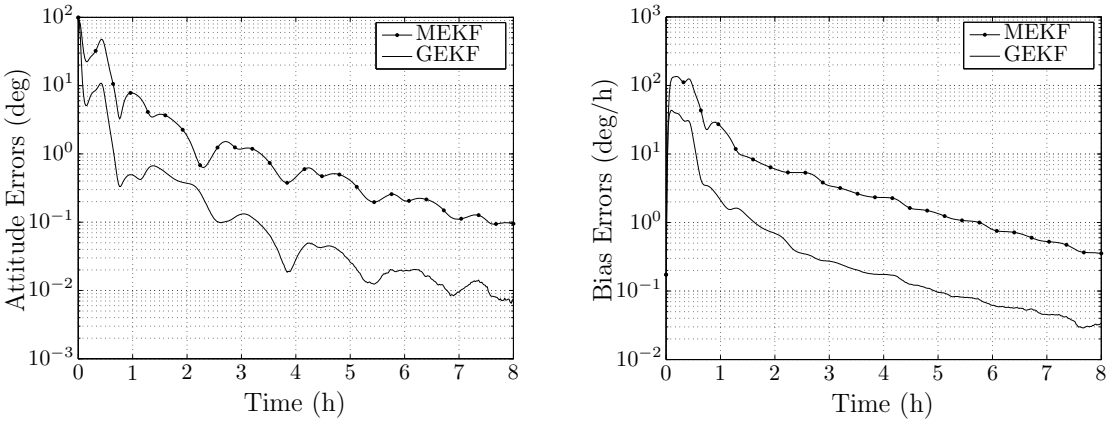
The following examples assume that the attitude sensor hardware consists of a three-axis magnetometer (TAM) and gyroscopic rate measurements. The magnetic field reference is modeled using a 10^{th}-order International Geomagnetic Reference Field model [23]. Magnetometer measurement noise is characterized by a standard deviation of $\sigma_m = 50\,\text{nT}$. The gyro parameters are set according to $\sigma_u = \sqrt{10}\times 10^{-10}\,\text{rad/s}^{3/2}$ and $\sigma_v = \sqrt{10}\times 10^{-7}\,\text{rad/s}^{1/2}$ with an initial bias of 0.1 deg/h on each axis. The gyro and TAM measurements are both sampled at 1 Hz. Gyroscope output is simulated using Eq. (7), and TAM measurements are generated according to Eqs. (44) and (46). The discrete-time MEKF and GEKF are implemented, and results compared for each example.

1. *Large Initial Attitude Error*

These simulation cases consider a large initial attitude error characterized by a 90° yaw and a 90° roll for a 3-2-1 Euler rotation sequence (or a 120° error about the principal rotation axis). This initial quaternion estimate is given by

$$\hat{\mathbf{q}}_0 = \begin{bmatrix} -0.7246 & -0.2164 & 0.4142 & -0.5065 \end{bmatrix}^T \tag{78}$$

The initial bias estimates are set to zero. The initial state error covariance $\mathcal{P}_k^-$ is diagonal with attitude error variances set to $(30\,\text{deg})^2$ and bias variances set to $(0.2\,\text{deg/h})^2$. The filter results are shown in Figure 2 over a time interval of 8 hours. In Figure 2(a) the GEKF norm attitude error converges below 1 deg in under an hour while the MEKF has a 1 deg settling time of over 2 hours. Norm bias estimate errors are shown in Figure 2(b). The GEKF norm bias error converges below 0.1 deg/h at about the 5 hour mark whereas the MEKF bias does not reach this point in the 8 hours simulated. GEKF errors are lower than those of the MEKF for the entire simulation.



(a) **Norm Attitude Error** (b) **Norm Bias Error**

Figure 2. Case 1 Results

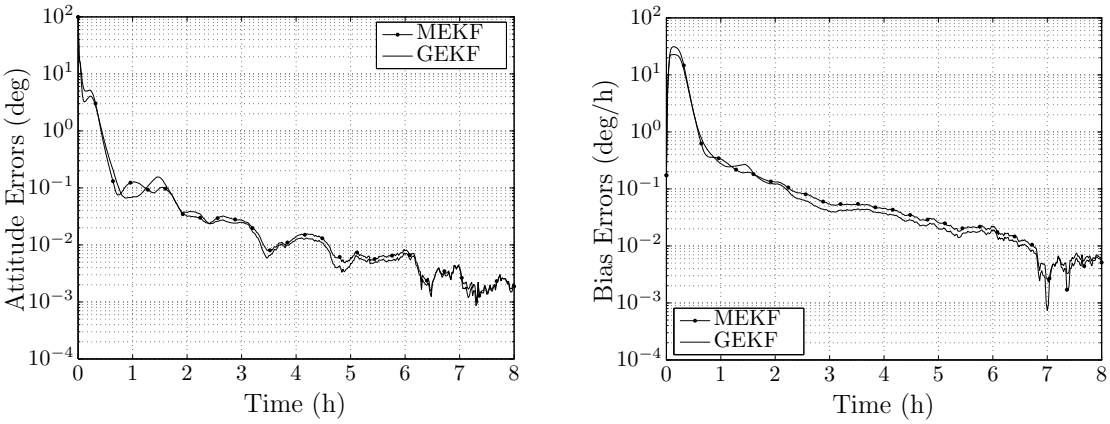
A second case is considered in which the initial mean anomaly m_0 of the orbit is changed. This corresponds to an initial true quaternion given by

$$\mathbf{q}_0 = \begin{bmatrix} 0.7435 & 0.3713 & 0.0010 & -0.5560 \end{bmatrix}^T \tag{79}$$

and with the same initial attitude error as in the previous case yields the initial estimate

$$\hat{\mathbf{q}}_0 = \begin{bmatrix} 0.2745 & -0.4592 & -0.0869 & -0.8404 \end{bmatrix}^T \tag{80}$$

All other parameters are identical to those of Case 1. The primary difference then is that the quaternion begins at a different point along the trajectory it sweeps out on the unit hypersphere. Figure 3 shows the results of this second test case. In this simulation, the MEKF and GEKF perform similarly.

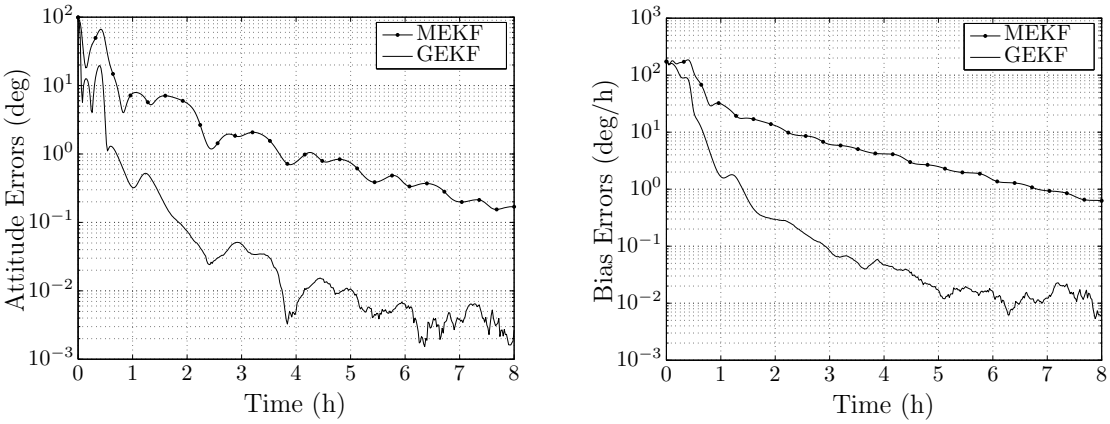


(a) **Norm Attitude Error** (b) **Norm Bias Error**

Figure 3. Case 2 Results

2. Large Initial Attitude and Gyro Failure

A simulation is now conducted to consider the case of gyro failure. The conditions and states for Case 1 are employed with the modification that a gyro failure is simulated by a true initial bias given by 100 deg/h on each axis. Results for this case are shown in Figure 4.



(a) **Norm Attitude Error** (b) **Norm Bias Error**

Figure 4. Case 3 Results

In this situation, both the norm attitude error and the norm bias error converge much more rapidly for the

GEKF. During the final simulated hours, the GEKF errors are more than an order of magnitude smaller than those of the MEKF.

B. Consistency Experiment

As might be expected, the simulations considered in this study do in fact yield comparable steady-state results. This is not illustrated here but it is easy to deduce that once converged, nonlinearities neglected by linearization are less important and further, the error definition employed by the GEKF approaches that of the MEKF. Examining Eq. (22b), as $\mathbf{dq}$ decreases, $\mathbf{db}$ approaches $\mathbf{\Delta b}$. Consider then the filter transients where these terms are generally more important.

Assume now that the spacecraft is undergoing motion characterized by

$$\mathbf{q}(t_0) = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T \tag{81a}$$

$$\boldsymbol{\omega}(t) = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}^T \text{ deg/s} \tag{81b}$$

Suppose also increased initial attitude and bias knowledge. The MEKF and GEKF algorithms are initialized accordingly using an initial geometric error covariance

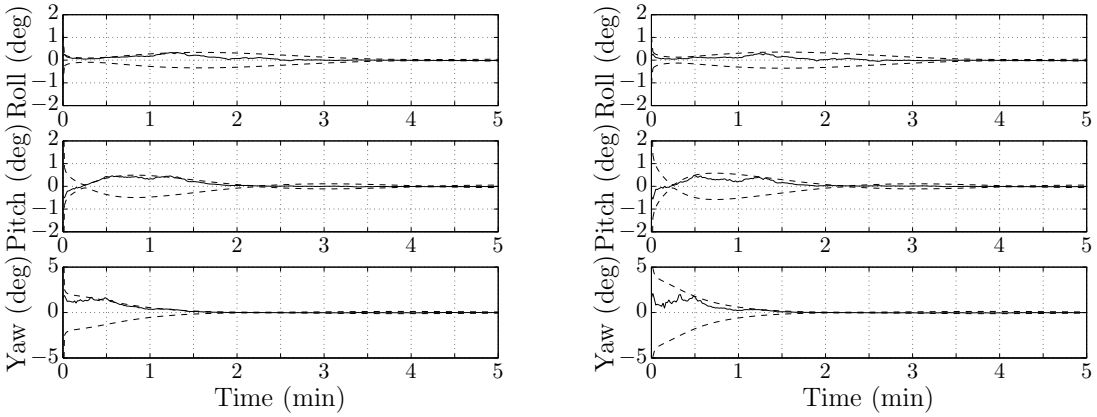
$$\mathcal{P}_0 = \begin{bmatrix} (5I_{3\times 3}\,\text{deg})^2 & 0_{3\times 3} \\ 0_{3\times 3} & (0.2I_{3\times 3}\,\text{deg/h})^2 \end{bmatrix} \tag{82}$$

as well as the initial bias estimate

$$\hat{\mathbf{b}}_0^- = \begin{bmatrix} -0.02 & 0.20 & 0.42 \end{bmatrix}^T \text{ deg/h} \tag{83}$$

and an initial attitude error corresponding to a 5° roll, −5° pitch and −15° yaw for a 3-2-1 Euler sequence.

Figure 5 details the attitude error transient response for the simulation characterized by Eqs. (81)−(83). The roll, pitch and yaw errors and 3-sigma confidence bounds are shown for the initial five minute interval. This test case suggests that the GEKF predicts attitude error 3-sigma bounds that are more consistent with the true attitude error (most apparent in the yaw component). A larger sample set is analyzed to reinforce this finding.



(a) **Multiplicative Extended Kalman Filter** (b) **Geometric Extended Kalman Filter**

Figure 5. Transient Attitude Errors and 3-Sigma Bounds for the MEKF and GEKF

A Monte Carlo experiment is conducted to support the assertion that the GEKF more accurately predicts the attitude error 3-sigma bounds during the transient period. The normalized error square (NES) test is

employed to accomplish this. The NES is defined by

$$\epsilon_k = \mathbf{e}_k^T \mathcal{E}_k^{-1} \mathbf{e}_k \tag{84}$$

where $\mathbf{e}_k$ is a discrete error which is known to be a zero-mean, Gaussian white-noise process with covariance $\mathcal{E}_k$. An appropriate verification of the NES is to numerically demonstrate that

$$E\{\epsilon_k\} = n \tag{85}$$

is satisfied where n is the dimension of the error state vector [17]. With M Monte Carlo sample runs, the average NES is computed according to

$$\bar{\epsilon}_k = \frac{1}{M} \sum_{i=1}^{M} \mathbf{e}_{k,i}^T \mathcal{E}_{k,i}^{-1} \mathbf{e}_{k,i} \tag{86}$$

where $\mathbf{e}_{k,i}$ denotes the i^{th} run at time t_k. To what degree Eq. (85) is satisfied for the average NES $\bar{\epsilon}_k$ provides a quantitative measure of filter consistency.

Filter consistency during the transient period is studied by determining the $\bar{\epsilon}_k$ sequence over the initial 5 minute interval. For the MEKF, $\mathbf{e}_{k,i} = \left(\mathbf{d}\boldsymbol{\alpha}_{k,i}, \Delta\mathbf{b}_{k,i}\right)$ and $\mathcal{E}_{k,i} = P_{k,i}$ into Eq. (86). Likewise, the test is applied to the GEKF with $\mathbf{e}_{k,i} = \mathbf{dx}_{k,i}$ and $\mathcal{E}_{k,i} = \mathcal{P}_{k,i}$.

The true state trajectory is fixed and $M = 500$ sample filter estimate trajectories are simulated using initial errors sampled from the initial error covariance. The resulting discrete, average NES sequences are displayed in Figure 6. It is important to note that both filters are initialized with the same state and covariance. An update occurs at the first measurement point, which explains the difference in the starting values shown in Figure 6. The theoretical value is $\bar{\epsilon}_k = 6$, the dimension of the state error vector. Figure 6 illustrates the improved predicted distribution accuracy obtained by the GEKF. While the average NES associated with the MEKF converges to within 6.0 ± 0.5 after 5 minutes, the GEKF attains this at 1 m 45 s. At the end of 5 minutes, the GEKF average NES converges to within 6.0 ± 0.05.

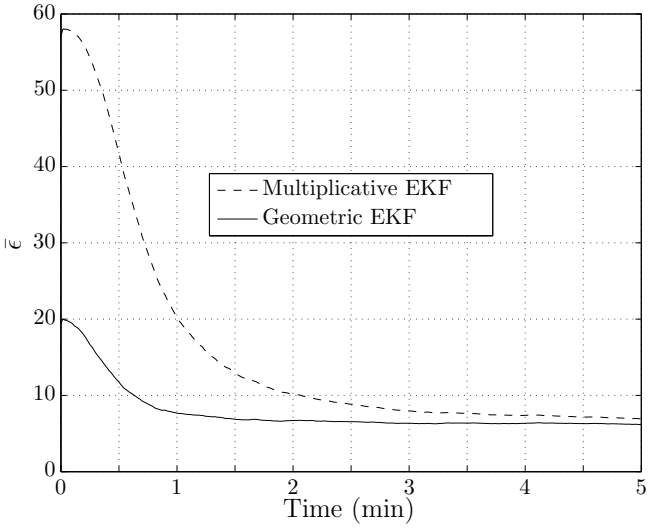


Figure 6. Mean Normalized Error Square

It is arguably more important to have an accurate confidence bound than to have a more accurate estimate. In practice, the accuracy of an estimate is not exactly known while confidence intervals provide the primary accuracy metric. Under the Gaussian error assumption, the GEKF is shown to provide a more consistent assessment of state knowledge during the transient period. It is believed that the GEKF outperforms the MEKF for two reasons: 1.) comparing the linearized state matrix of the GEKF in Eq. (40) to the linearized state matrix of the MEKF in Eq. (63) shows that there are more coupling effects in the bias propagation for

the GEKF than the MEKF, and 2.) the update in the GEKF provides an extra transport term through the $[\hat{\mathbf{b}}_k^- \times]$ matrix in the C_k^- matrix, which is not present in the MEKF.

VI. Conclusions

It is clear that just as classical filtering algorithms are unconstrained, they also do not explicitly acknowledge with respect to which coordinate frame a state vector quantity is expressed. This is readily demonstrated by implementing the extended Kalman Filter for attitude estimation which yields error realizations that in general are expressed with respect to different coordinate frames. A mean and covariance are not sensible in this situation since these quantities require summation over the sample space. A new frame-consistent filter based on the extended Kalman Filter was developed which is based on a new state vector error definition. The resulting algorithm, the Geometric Extended Kalman Filter, acknowledges that the state vector is expressed with respect to a specific reference frame. Results obtained by numerical simulation demonstrate similar or better state convergence for the cases considered. Furthermore, the Geometric Extended Kalman Filter outperformed the Multiplicative Extended Kalman Filter in a Gaussian error consistency test. This fact is arguably more important since in application, the actual state error is not known and an error bound that is truly compatible is more indicative of the state errors in practice. Finally, a discrete time Geometric Extended Kalman Filter is presented for practical application.

Acknowledgements

Special thanks to F. Landis Markley of NASA Goddard Space Flight Center, Edward Corbett of MIT Lincoln Laboratory, Yang Cheng of Mississippi State University and Puneet Singla from the University at Buffalo for their contributions, discussions and comments which played a critical role in the development and presentation of this work.

Appendix: Alternate Derivation

The covariance propagation described by Eqs. (40)−(42) can also be derived directly in the same way that the MEKF is derived in the reduced covariance approach [1]. This eliminates the pseudoinverse used in determining Eq. (32).

Taking the time-derivative of Eq. (22b) and using Eq. (11b) as well as the fact that $\dot{\hat{\mathbf{b}}} = \mathbf{0}$ leads to

$$\dot{\mathbf{db}} = A^T(\mathbf{dq})\boldsymbol{\eta}_u + \dot{A}^T(\mathbf{dq})\mathbf{b} \tag{A.1}$$

Using $A(\mathbf{dq}) = A(\mathbf{q})A^T(\hat{\mathbf{q}})$ allows one to deduce that

$$\dot{A}(\mathbf{dq}) = -A(\mathbf{dq})[\mathbf{d}\boldsymbol{\omega}\times] \tag{A.2}$$

where

$$\mathbf{d}\boldsymbol{\omega} \triangleq A^T(\mathbf{dq})\boldsymbol{\omega} - \hat{\boldsymbol{\omega}} \tag{A.3}$$

Using Eq. (A.2) in Eq. (A.1) leads to

$$\dot{\mathbf{db}} = A^T(\mathbf{dq})\boldsymbol{\eta}_u + [\mathbf{d}\boldsymbol{\omega}\times]A^T(\mathbf{dq})\mathbf{b} \tag{A.4}$$

Substituting Eq. (7a) into Eq. (A.3) gives

$$\mathbf{d}\boldsymbol{\omega} = A^T(\mathbf{dq})(\tilde{\boldsymbol{\omega}} - \mathbf{b} - \boldsymbol{\eta}_v) - (\tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}}) \tag{A.5}$$

Using the definition of $\mathbf{db}$ in Eq. (22b) gives

$$\mathbf{d}\boldsymbol{\omega} = A^T(\mathbf{dq})\tilde{\boldsymbol{\omega}} - \mathbf{db} - A^T(\mathbf{dq})\boldsymbol{\eta}_v - \tilde{\boldsymbol{\omega}} \tag{A.6}$$

Assuming small angle errors, the attitude matrix in Eq. (A.6) can be replaced using Eq. (23) leading to

$$\mathbf{d}\boldsymbol{\omega} = [\mathbf{d}\boldsymbol{\alpha}\times]\,\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - [\mathbf{d}\boldsymbol{\alpha}\times]\,\boldsymbol{\eta}_v \tag{A.7}$$

Also using Eq. (23) in Eq. (A.4) yields

$$\dot{\mathbf{db}} = \boldsymbol{\eta}_u + [\mathbf{d}\boldsymbol{\alpha}\times]\,\boldsymbol{\eta}_u + [\mathbf{d}\boldsymbol{\omega}\times]\,\mathbf{b} + [\mathbf{d}\boldsymbol{\omega}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{b} \tag{A.8}$$

Substituting Eq. (A.7) into Eq. (A.8) gives

$$\begin{aligned}\dot{\mathbf{db}} = \boldsymbol{\eta}_u + [\mathbf{d}\boldsymbol{\alpha}\times]\,\boldsymbol{\eta}_u + [(\mathbf{d}\boldsymbol{\alpha}\times\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - \mathbf{d}\boldsymbol{\alpha}\times\boldsymbol{\eta}_v)\times]\,\mathbf{b} \\ + [(\mathbf{d}\boldsymbol{\alpha}\times\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - \mathbf{d}\boldsymbol{\alpha}\times\boldsymbol{\eta}_v)\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{b}\end{aligned} \tag{A.9}$$

Solving Eq. (22b) for $\mathbf{b}$ and using Eq. (23) gives

$$\mathbf{b} = \mathbf{db} - [\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{db} + \hat{\mathbf{b}} - [\mathbf{d}\boldsymbol{\alpha}\times]\,\hat{\mathbf{b}} \tag{A.10}$$

Substituting Eq. (A.10) into Eq. (A.9) gives

$$\begin{aligned}\dot{\mathbf{db}} = \boldsymbol{\eta}_u + [\boldsymbol{\gamma}\times]\,\mathbf{db} - [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{db} + [\boldsymbol{\gamma}\times]\,\hat{\mathbf{b}} - [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\hat{\mathbf{b}} + [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{db} \\ - [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\mathbf{db} + [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\hat{\mathbf{b}} - [\boldsymbol{\gamma}\times][\mathbf{d}\boldsymbol{\alpha}\times][\mathbf{d}\boldsymbol{\alpha}\times]\,\hat{\mathbf{b}}\end{aligned} \tag{A.11}$$

where the vector $\boldsymbol{\gamma}$ is defined by

$$\boldsymbol{\gamma} \triangleq \mathbf{d}\boldsymbol{\alpha}\times\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - \mathbf{d}\boldsymbol{\alpha}\times\boldsymbol{\eta}_v \tag{A.12}$$

Retaining only first-order terms gives

$$\dot{\mathbf{db}} = [\hat{\mathbf{b}}\times][\tilde{\boldsymbol{\omega}}\times]\,\mathbf{d}\boldsymbol{\alpha} + [\hat{\mathbf{b}}\times]\,\mathbf{db} + [\hat{\mathbf{b}}\times]\,\boldsymbol{\eta}_v + \boldsymbol{\eta}_u \tag{A.13}$$

From [1] the quaternion attitude error is given as

$$\dot{\mathbf{dq}} = \frac{1}{2}\left\{\begin{bmatrix}\boldsymbol{\omega}\\0\end{bmatrix}\otimes\mathbf{dq} - \mathbf{dq}\otimes\begin{bmatrix}\hat{\boldsymbol{\omega}}\\0\end{bmatrix}\right\} \tag{A.14}$$

without error. To derive the error-attitude kinematics, solve Eq. (A.3) for $\boldsymbol{\omega}$ and substitute the resultant into Eq. (A.14) to yield

$$\dot{\mathbf{dq}} = \frac{1}{2}\left\{\begin{bmatrix}A(\mathbf{dq})\mathbf{d}\boldsymbol{\omega}\\0\end{bmatrix}\otimes\mathbf{dq} + \begin{bmatrix}A(\mathbf{dq})\hat{\boldsymbol{\omega}}\\0\end{bmatrix}\otimes\mathbf{dq} - \mathbf{dq}\otimes\begin{bmatrix}\hat{\boldsymbol{\omega}}\\0\end{bmatrix}\right\} \tag{A.15}$$

Substituting Eq. (23) into Eq. (A.15) and noting that to within first-order, $\mathrm{d}q_4 \approx 1$, leads to

$$\dot{\mathbf{dq}} = \frac{1}{2}\left\{\begin{bmatrix}\mathbf{d}\boldsymbol{\omega}\\0\end{bmatrix} + 2\begin{bmatrix}[\mathbf{d}\boldsymbol{\omega}\times]\,\mathbf{d}\boldsymbol{\varrho}\\0\end{bmatrix}\right\} \tag{A.16}$$

Using the first-order approximation $\mathbf{d}\boldsymbol{\alpha} \approx 2\,\mathbf{d}\boldsymbol{\varrho}$ leads to

$$\dot{\mathbf{d}\boldsymbol{\alpha}} = [\mathbf{d}\boldsymbol{\omega}\times]\,\mathbf{d}\boldsymbol{\alpha} + \mathbf{d}\boldsymbol{\omega} \tag{A.17}$$

Substituting Eq. (A.7) into Eq. (A.17) gives

$$\begin{aligned}\dot{\mathbf{d}\boldsymbol{\alpha}} = [(\mathbf{d}\boldsymbol{\alpha}\times\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - \mathbf{d}\boldsymbol{\alpha}\times\boldsymbol{\eta}_v)\times]\,\mathbf{d}\boldsymbol{\alpha} \\ + [\mathbf{d}\boldsymbol{\alpha}\times]\,\tilde{\boldsymbol{\omega}} - \mathbf{db} - \boldsymbol{\eta}_v - [\mathbf{d}\boldsymbol{\alpha}\times]\,\boldsymbol{\eta}_v\end{aligned} \tag{A.18}$$

Retaining only first-order terms gives

$$\dot{\mathbf{d}\boldsymbol{\alpha}} = -[\tilde{\boldsymbol{\omega}}\times]\,\mathbf{d}\boldsymbol{\alpha} - \mathbf{db} - \boldsymbol{\eta}_v \tag{A.19}$$

Therefore, the linearized error-dynamics in Eqs. (A.19) and (A.13) can be assembled into

$$\begin{bmatrix}\dot{\mathbf{d}\boldsymbol{\alpha}}\\ \dot{\mathbf{db}}\end{bmatrix} = \begin{bmatrix}-[\tilde{\boldsymbol{\omega}}\times] & -I_{3\times3}\\ [\hat{\mathbf{b}}\times][\tilde{\boldsymbol{\omega}}\times] & [\hat{\mathbf{b}}\times]\end{bmatrix}\begin{bmatrix}\mathbf{d}\boldsymbol{\alpha}\\ \mathbf{db}\end{bmatrix} + \begin{bmatrix}-I_{3\times3} & 0_{3\times3}\\ [\hat{\mathbf{b}}\times] & I_{3\times3}\end{bmatrix}\begin{bmatrix}\boldsymbol{\eta}_v\\ \boldsymbol{\eta}_u\end{bmatrix} \tag{A.20}$$

Equation (A.20) agrees with the result obtained in Eq. (39).

As is the case with the approach taken in Section IV, reversing the definition of $\mathbf{db}$ to that of Eq. (43) produces the result obtained in Eq. (A.13). However, this is only true to within first-order. More specifically, the expression counterpart to Eq. (A.11) would differ in its higher order terms.

References

[1]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, 1982, pp. 417–429, doi:10.2514/3.56190.

[2]Savage, P.G., "Strapdown Inertial Navigation Integration Algorithm Design Part 1: Attitude Algorithms," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 1, Jan.-Feb. 1998, pp. 19–28, doi:10.2514/2.4228.

[3]Miller, P.A., Farrell, J.A., Zhao, Y., and Djapic, V., "Autonomous Underwater Vehicle Navigation," *IEEE Journal of Oceanic Engineering*, Vol. 35, No. 3, July 2010, pp. 663–678, doi:10.1109/JOE.2010.2052691.

[4]Barshan, B. and Durrant-Whyte, H.F., "Inertial Navigation Systems for Mobile Robots," *IEEE Transactions on Robotics and Automation*, Vol. 11, No. 3, June 1995, pp. 328–342, doi:10.1109/70.388775.

[5]Farrell, J.L., "Attitude Determination by Kalman Filtering," *Automatica*, Vol. 6, No. 5, 1970, pp. 419–430, doi:10.1016/0005-1098(70)90057-9.

[6]Karlgaard, C.D. and Schaub, H., "Nonsingular Attitude Filtering Using Modified Rodrigues Parameters," *The Journal of the Astronautical Sciences*, Vol. 57, No. 4, 2009, pp. 777–791, doi:10.1007/BF03321529.

[7]Crassidis, J.L., Markley, F.L., and Cheng, Y., "Survey of Nonlinear Attitude Estimation Methods," *Journal of Guidance, Control, and Dynamics*, Vol. 30, No. 1, Jan.-Feb. 2007, pp. 12–28, doi:10.2514/1.22452.

[8]Markley, F.L., "Attitude Error Representations for Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, 2003, pp. 311–317, doi:10.2514/2.5048.

[9]Julier, S.J., Uhlmann, J.K., and Durrant-Whyte, H.F., "A New Approach for Filtering Nonlinear Systems," *American Control Conference*, June 1995, pp. 1628–1632, doi:10.1109/ACC.1995.529783.

[10]Crassidis, J.L. and Markley, F.L., "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, July-Aug. 2003, pp. 536–542, doi:10.2514/2.5102.

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

[12]Farrenkopf, R.L., "Analytic Steady-State Accuracy Solutions for Two Common Spacecraft Attitude Estimators," *Journal of Guidance and Control*, Vol. 1, No. 4, 1978, pp. 282–284, doi:10.2514/3.55779.

[13]Hughes, P.C., *Spacecraft Attitude Dynamics*, Dover Publications, New York, NY, 1st ed., 2004, pp. 6–30.

[14]Gelb, A., *Applied Optimal Estimation*, The M.I.T. Press, Cambridge, MA, 1st ed., 1974, pp. 180–192.

[15]Markley, F.L. and Crassidis, J.L., *Fundamentals of Spacecraft Attitude Determination and Control*, Springer, New York, NY, 2014, pp. 246–249, 258–259, 345–359, doi:10.1007/978-1-4939-0802-8.

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

[17]Crassidis, J.L. and Junkins, J.L., *Optimal Estimation of Dynamic Systems*, CRC Press, Boca Raton, FL, 2nd ed., 2012, pp. 451–460, 611–614.

[18]Reynolds, R.G., "Asymptotically Optimal Attitude Filtering with Guaranteed Convergence," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 1, 2015/02/09 2008, pp. 114–122.

[19]Markley, F.L., "Lessons Learned," *The Journal of the Astronautical Sciences*, Vol. 57, No. 1-2, 2009, pp. 3–29.

[20]Chen, C.T., *Linear System Theory and Design*, Holt, Rinehart, and Winston, New York, NY, 1st ed., 1984, pp. 134–141.

[21]Golub, G. and Van Loan, C., *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, 1996, pp. 572–578.

[22]Crassidis, J.L. and Markley, F.L., "Attitude Estimation Using Modified Rodrigues Parameters," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA-Goddard Space Flight Center, Greenbelt, MD, May 1996, pp. 71–83.

[23]Langel, R.A., "International Geomagnetic Reference Field: The Sixth Generation," *Journal of Geomagnetism and Geoelectricity*, Vol. 44, No. 9, 1992, pp. 679–707.