Novel Quaternion Kalman Filter

D. CHOUKROUN
UCLA

I. Y. BAR-ITZHACK, Fellow, IEEE

Y. OSHMAN, Senior Member, IEEE
Technion–Israel Institute of Technology
Israel

This paper presents a novel Kalman filter (KF) for estimating the attitude-quaternion as well as gyro random drifts from vector measurements. Employing a special manipulation on the measurement equation results in a linear pseudo-measurement equation whose error is state-dependent. Because the quaternion kinematics equation is linear, the combination of the two yields a linear KF that eliminates the usual linearization procedure and is less sensitive to initial estimation errors. General accurate expressions for the covariance matrices of the system state-dependent noises are developed. In addition, an analysis shows how to compute these covariance matrices efficiently. An adaptive version of the filter is also developed to handle modeling errors of the dynamic system noise statistics. Monte-Carlo simulations are carried out that demonstrate the efficiency of both versions of the filter. In the particular case of high initial estimation errors, a typical extended Kalman filter (EKF) fails to converge whereas the proposed filter succeeds.

Manuscript received May 7, 2004; revised March 29, 2005; released for publication June 10, 2005.

IEEE Log No. T-AES/42/1/870599.

Refereeing of this contribution was handled by M. Ruggieri.

This work was supported by NASA-Goddard Space Flight Center under Grant NAG5-8770.

This work was presented at the 42nd AIAA Guidance, Navigation, and Control Conference, Monterey, CA, Aug. 2002; paper 2002-4460.

Authors' current addresses: D. Choukroun, Dept. of Mechanical and Aerospace Engineering, University of California, Los Angeles, CA, E-mail: (dach@ucla.edu); I. Y. Bar-Itzhack and Y. Oshman, Asher Space Research Institute, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel, E-mail: (ibaritz@technion.ac.il, yaakov.oshman@technion.ac.il).

INTRODUCTION

In the normal course of spacecraft operation, the spacecraft attitude needs to be determined. The task of attitude determination (AD) is that of determining the orientation of the spacecraft relative to some reference frame. This reference frame is usually celestial or Earth-fixed. The mathematical representation of attitude is very diverse [1, 2]. One attitude representation that has proven very useful is the attitude quaternion, which is a 4×1 unit-norm vector in $\mathbb{R}^4$. A well-known attractive feature of this representation is that the formulation of the attitude dynamics in terms of the quaternion is linear and nonsingular. Moreover, with only one redundant parameter, the quaternion is the minimal nonsingular attitude parameterization.

Estimating the attitude quaternion from spacecraft on-board measurements has a long history. Optimal algorithms have been developed over the last four decades following two main approaches; namely, the classical least-squares approach and the Kalman filtering approach. The first approach was originally introduced in 1967, in the so-called Wahba's problem [3], which is a constrained least-squares minimization problem for finding the attitude matrix. A basic assumption was that a batch of at least two simultaneous vector measurements of the attitude was available. Later, in 1971, Davenport formulated and solved Wahba's problem in terms of the attitude quaternion [4]. The algorithm developed is known in the literature as the **q**-method [1, pp. 426–428]. One advantage of the **q**-method is that it yields a closed-form optimal estimate of the quaternion while explicitly preserving the unit-norm property. Numerous algorithms were developed in order to provide the basic **q**-method with other features like the ability of sequentially estimating time-varying attitude [5, 6] and of estimating parameters other than attitude [6, 7]. Furthermore, some works addressed the issue of numerical accuracy [8], while others aimed at providing a probabilistic meaning to the original least-squares problem [6].

The Kalman filtering approach, on the other hand, yields, by design, sequential quaternion estimates that are minimum-variance, and allows the estimation of parameters other than attitude in a straightforward manner. The Kalman filter (KF), however, is not designed to preserve constraints imposed on the estimated state variables. The difficulty of preserving the unit-norm property of the quaternion estimate was overcome in various ways. One way was to develop a measurement update stage in an extended KF (EKF) where the error between the true and estimated quaternion is itself a quaternion and is multiplied (in the sense of quaternion multiplication) with the a priori quaternion estimate to yield the a posteriori estimate. This kind

of EKF is called a multiplicative EKF [9]. Another way was to normalize the a posteriori estimate after a classic (additive) measurement update stage of the EKF [10]. An alternative method to preserve the quaternion unit-norm property was to derive a fictitious quaternion measurement model from the equation expressing the unit-norm property. This model, called pseudo-measurement model was implemented in an additive EKF [11].

In the works mentioned hitherto, the EKFs were operating on a nonlinear quaternion measurement model. In the case of vector observations [10] this model is derived as follows. The vector measurement equation is

$$\mathbf{b} = \mathbf{b}^o + \boldsymbol{\delta}\mathbf{b} \tag{1}$$

where

$$\mathbf{b}^o = A(\mathbf{q})\mathbf{r}. \tag{2}$$

The 3×1 vectors $\mathbf{b}^o$ and $\mathbf{r}$ are the normalized projections of a physical vector along the axes of the body frame $\mathcal{B}$ and the reference frame $\mathcal{R}$, respectively. The vector $\mathbf{b}$ is, usually, the output of a body-fixed sensor while $\mathbf{r}$ is known from an almanac or a model. The vector $\boldsymbol{\delta}\mathbf{b}$ is an additive measurement noise. The 3×3 matrix A is the rotation matrix that brings the axes of $\mathcal{R}$ onto the axes of $\mathcal{B}$; that is, A is the attitude matrix, and $\mathbf{q}$ is the quaternion that corresponds to A. It is well known that the attitude matrix and the attitude quaternion are related by [1, p. 414]

$$A(\mathbf{q}) = (q^2 - \mathbf{e}^T\mathbf{e})I_3 + 2\mathbf{e}\mathbf{e}^T - 2q[\mathbf{e}\times] \tag{3}$$

where $\mathbf{e}$ and q are the vector and scalar part, respectively, of the attitude-quaternion $\mathbf{q}$ and

$$\mathbf{q}^T = [\mathbf{e}^T \;\; q]. \tag{4}$$

The 3×3 matrix $[\mathbf{e}\times]$ denotes a skew-symmetric matrix function of $\mathbf{e}$. This matrix, usually called cross-product matrix, is used to express the cross-product of two vectors $\mathbf{x}$ and $\mathbf{y}$ in a matrix-vector form as follows $\mathbf{x} \times \mathbf{y} = [\mathbf{x}\times]\mathbf{y}$. Using (3) in (1) and (2) leads to a measurement equation that is quadratic with respect to $\mathbf{q}$. This equation is linearized and the linearized model is implemented in an EKF. The linearization procedure induces undesirable effects such as sensitivity to initial conditions, biases in the estimation errors, and, sometimes, an increase in the computation load, in particular when the gradient matrices must be evaluated by numerical methods.

The present work introduces a novel model for quaternion vector measurements. This model avoids the linearization procedure. Using this approach, though, the measurement noise is quaternion dependent. The dependence on the quaternion is linear, and this dependence is typical to a wide class of state-dependent noises. Exact general expressions for the covariance matrices of such state-dependent noises are developed here. An analysis shows how their approximations in the filter design can be efficiently computed. A nice feature of the present quaternion model is that the quaternion-dependent process noise has the same pattern as the measurement noise. It should be mentioned that this pattern has already been emphasized in previous works [9, 10].

An ordinary linear KF is designed here to operate on the developed quaternion state-space model, where the state-dependent coefficients are computed using the best available state estimate. The dynamics model is based on gyro measurements. A realistic model for rate integrating gyros (RIG) is considered, and the state vector is augmented to include random gyro drifts. Nevertheless the state-dependence of the model parameters induces modeling errors. Such errors can be dealt with by means of on-line optimal adaptive filtering. A simple process noise adaptive procedure is proposed here. It is applied to two different cases. The first case is that of a filter that uses a process noise variance which is lower than the actual one. In the second case, the gyro outputs are contaminated by an additional and unmodeled drift.

The performance of this novel quaternion KF is first checked under nominal simulated conditions of noise and initial errors. Then we check, by extensive Monte-Carlo simulations, the conjecture that this filter is less sensitive to initial errors than a baseline EKF. Finally we implement the two adaptive versions of the filter.

In the following section we derive the state-space truth model of the quaternion system. Then we provide a detailed covariance analysis of the state-dependent noises. Next we describe the design model and present the corresponding quaternion KF. An adaptive filter based on the design model is then developed. As an introduction to the simulations that follow, the subsequent section contains a brief review of the ordinary linearized measurement update model in an additive EKF. The simulation results for the nominal nonadaptive filter, for the comparative simulation, and for the adaptive filters are presented in the section that follows. Finally, in the last section, we present the conclusions derived from this work.

MATHEMATICAL MODEL

The Measurement Truth Model

We consider the pair of 3×1 unit column-matrices $\mathbf{b}^o$ and $\mathbf{r}$ obtained when resolving a physical vector, at a given time, in the body frame $\mathcal{B}$ and in the reference frame $\mathcal{R}$, respectively. (The time subscripts are omitted in the following development for clarity.) We define the 4×1 quaternion vectors $\mathbf{b}_q^o$ and $\mathbf{r}_q$ as follows:

$$\mathbf{b}_q^o \triangleq \begin{bmatrix} \mathbf{b}^o \\ 0 \end{bmatrix}, \qquad \mathbf{r}_q \triangleq \begin{bmatrix} \mathbf{r} \\ 0 \end{bmatrix}. \tag{5}$$

It is well known that $\mathbf{b}_q^o$ and $\mathbf{r}_q$ are related by the quaternion of rotation $\mathbf{q}$ as follows [12]:

$$\mathbf{b}_q^o = \mathbf{q}^{-1} \circledast \mathbf{r}_q \circledast \mathbf{q} \tag{6}$$

where $\circledast$ is the quaternion product and $\mathbf{q}^{-1}$ is the quaternion inverse [1, pp. 758, 759]. Let $\Upsilon(\cdot)$ and $\chi(\cdot)$ denote the linear mappings from $\mathbb{R}^3$ to $\mathbb{R}^{4\times 4}$ that are defined as follows:

$$\Upsilon(\mathbf{x}) \triangleq \begin{bmatrix} -[\mathbf{x}\times] & \mathbf{x} \\ -\mathbf{x}^T & 0 \end{bmatrix} \quad \text{and} \quad \chi(\mathbf{x}) \triangleq \begin{bmatrix} [\mathbf{x}\times] & \mathbf{x} \\ -\mathbf{x}^T & 0 \end{bmatrix} \quad \forall \quad \mathbf{x} \in \mathbb{R}^3 \tag{7}$$

where $[\mathbf{x}\times]$ denotes the cross-product matrix that is associated with $\mathbf{x}$. Then the definition of the quaternion product yields the identities

$$\mathbf{q} \circledast \mathbf{b}_q^o = \Upsilon(\mathbf{b}^o)\mathbf{q} \tag{8a}$$

$$\mathbf{r}_q \circledast \mathbf{q} = \chi(\mathbf{r})\mathbf{q}. \tag{8b}$$

Premultiplying (6) by $\mathbf{q}$ and substituting (8) into the resulting equation leads to the following 4×1 vector equation

$$[\Upsilon(\mathbf{b}^o) - \chi(\mathbf{r})]\mathbf{q} = \begin{bmatrix} -[(\mathbf{b}^o + \mathbf{r})\times] & \mathbf{b}^o - \mathbf{r} \\ -(\mathbf{b}^o - \mathbf{r})^T & 0 \end{bmatrix} \mathbf{q} = \mathbf{0}. \tag{9}$$

Defining the 4×1 vectors $\mathbf{s}^o$, $\mathbf{d}^o$, and the 4×4 skew-symmetric matrix H^o as follows

$$\mathbf{s}^o \triangleq \tfrac{1}{2}(\mathbf{b}^o + \mathbf{r}) \tag{10a}$$

$$\mathbf{d}^o \triangleq \tfrac{1}{2}(\mathbf{b}^o - \mathbf{r}) \tag{10b}$$

$$H^o \triangleq \begin{bmatrix} -[\mathbf{s}^o\times] & \mathbf{d}^o \\ -\mathbf{d}^{oT} & 0 \end{bmatrix} \tag{10c}$$

and using (10) in (9) yields

$$H^o\mathbf{q} = \mathbf{0}. \tag{11}$$

This equation, which is linear with respect to the quaternion, is the model equation of an error-free quaternion measurement, where the observation matrix H^o is a function of the data $\mathbf{b}^o, \mathbf{r}$. Equation (11) indicates that $\mathbf{q}$ belongs to the null space of the matrix H^o.

In general the vector $\mathbf{b}_{k+1}$ is the output of a sensor at time t_{k+1}, e.g. magnetometer, Earth sensor, sun sensor, etc. The $\mathbf{r}_{k+1}$ vector is known from tables, almanac, or models and, thus, is relatively accurate. The measurement noise, denoted by $\delta\mathbf{b}_{k+1}$, is defined by

$$\delta\mathbf{b}_{k+1} \triangleq \mathbf{b}_{k+1} - \mathbf{b}_{k+1}^o \tag{12}$$

where $\mathbf{b}_{k+1}^o$ is the true value of the observation. From (12) and (10), we obtain a modified (11) as follows:

$$\begin{aligned} H^o\mathbf{q} &= \begin{bmatrix} -[\frac{1}{2}(\mathbf{b} - \delta\mathbf{b} + \mathbf{r})\times] & \frac{1}{2}(\mathbf{b} - \delta\mathbf{b} - \mathbf{r}) \\ -\frac{1}{2}(\mathbf{b} - \delta\mathbf{b} - \mathbf{r})^T & 0 \end{bmatrix} \mathbf{q} \\ &= \begin{bmatrix} -[\mathbf{s}\times] - [-\frac{1}{2}\delta\mathbf{b}\times] & \mathbf{d} - \frac{1}{2}\delta\mathbf{b} \\ -\mathbf{d}^T - (-\frac{1}{2}\delta\mathbf{b})^T & 0 \end{bmatrix} \mathbf{q} \\ &= H\mathbf{q} - \tfrac{1}{2}\Upsilon(\delta\mathbf{b})\mathbf{q} \end{aligned} \tag{13}$$

where the vectors $\mathbf{s}$ and $\mathbf{d}$, and the matrix $\Upsilon(\delta\mathbf{b})$ are implicitly defined by the second and third equalities, respectively. Using (11), we rewrite (13) as

$$\mathbf{0} = H\mathbf{q} - \tfrac{1}{2}\Upsilon(\delta\mathbf{b})\mathbf{q}. \tag{14}$$

Using the identity (see [13, Appendix A])

$$\Upsilon(\mathbf{x})\mathbf{y} = \Xi(\mathbf{y})\mathbf{x} \qquad \forall \quad (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^3 \times \mathbb{R}^4 \tag{15}$$

where $\mathbf{x} = \delta\mathbf{b}$, $\mathbf{y} = \mathbf{q}$, and

$$\Xi(\mathbf{q}) = \begin{bmatrix} [\mathbf{e}\times] + qI_3 \\ -\mathbf{e}^T \end{bmatrix}. \tag{16}$$

Equation (14) can be written at time t_{k+1} as

$$\mathbf{0} = H_{k+1}\mathbf{q}_{k+1} - \tfrac{1}{2}\Xi_{k+1}\delta\mathbf{b}_{k+1}. \tag{17}$$

where Ξ_{k+1} denotes $\Xi(\mathbf{q}_{k+1})$. Equation (17) describes the quaternion pseudo-measurement model at time t_{k+1}. The signal term $H_{k+1}\mathbf{q}_{k+1}$ is a linear function of the quaternion. The noise term, $-\frac{1}{2}\Xi_{k+1}\delta\mathbf{b}_{k+1}$, is an additive 4×1 quaternion-dependent vector. It is assumed that the 3×1 measurement noise vector $\delta\mathbf{b}_{k+1}$ is a zero-mean white-noise process with a covariance matrix R_{k+1}; that is

$$E\{\delta\mathbf{b}_{k+1}\} = \mathbf{0}, \qquad E\{\delta\mathbf{b}_{k+1}\delta\mathbf{b}_{k+1}^T\} = R_{k+1}. \tag{18}$$

The Process Truth Model

The quaternion continuous-time process truth model has been developed in previous works [9, 10]. Here, we develop the discrete-time process equation. We consider the following discrete-time process of the attitude kinematics [1, pp. 511, 512]

$$\mathbf{q}_{k+1} = \Phi_k^o\mathbf{q}_k, \qquad k = 0, 1, \ldots \tag{19}$$

with the initial conditions $\mathbf{q}_0$. The vector $\mathbf{q}_k$ is the quaternion of the rotation from a given reference frame $\mathcal{R}$ onto the body frame $\mathcal{B}^o$ at time t_k. Using $\boldsymbol{\omega}_k^o$, the angular velocity vector of $\mathcal{B}^o$ with respect to $\mathcal{R}$ resolved in $\mathcal{B}^o$, define

$$\Omega_k^o \triangleq \Upsilon(\boldsymbol{\omega}_k^o). \tag{20}$$

Then, for relatively short time intervals, $\Delta t = t_{k+1} - t_k$, the 4×4 orthogonal transition matrix, Φ_k^o, can be expressed by [1, pp. 511, 512]

$$\Phi_k^o = \exp(\tfrac{1}{2}\Omega_k^o\Delta t). \tag{21}$$

Defining a 3×1 vector of integrated rates $\boldsymbol{\theta}_k^o$ and the corresponding matrix Θ_k^o as follows,

$$\boldsymbol{\theta}_k^o \triangleq \boldsymbol{\omega}_k^o \Delta t \tag{22a}$$

$$\Theta_k^o \triangleq \Upsilon(\boldsymbol{\theta}_k^o). \tag{22b}$$

Equation (19) is rewritten as

$$\mathbf{q}_{k+1} = \exp(\tfrac{1}{2}\Theta_k^o)\mathbf{q}_k. \tag{23}$$

In practice the ideal angular increment, $\boldsymbol{\theta}_k^o$, is not known. In this work, it is assumed that a triad of RIG measures the incremental rotation over small time intervals Δt. The RIG output $\boldsymbol{\theta}_k$ is modeled with an additive error $\boldsymbol{\delta\theta}_k$

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^o + \boldsymbol{\delta\theta}_k. \tag{24}$$

A "measured" transition matrix can be computed using $\boldsymbol{\theta}_k$ instead of $\boldsymbol{\theta}_k^o$ in (22) and (23). This matrix, denoted by Φ_k, differs from Φ_k^o by $\Delta\Phi_k$; that is,

$$\Phi_k = \Phi_k^o + \Delta\Phi_k. \tag{25}$$

The error matrix $\Delta\Phi_k$ can be expressed as a matrix power series as follows:

$$\begin{aligned}
\Delta\Phi_k &= \Phi_k - \Phi_k^o \\
&= \exp(\tfrac{1}{2}\Theta_k) - \exp(\tfrac{1}{2}\Theta_k^o) \\
&= [I_4 + \tfrac{1}{2}\Theta_k + \tfrac{1}{2}(\tfrac{1}{2}\Theta_k)^2 + \text{HOT}] \\
&\quad - [I_4 + \tfrac{1}{2}\Theta_k^o + \tfrac{1}{2}(\tfrac{1}{2}\Theta_k^o)^2 + \text{HOT}] \\
&= \tfrac{1}{2}(\Theta_k - \Theta_k^o) + \tfrac{1}{8}(\Theta_k^2 - \Theta_k^{o2}) + \text{HOT} \\
&= \tfrac{1}{2}\delta\Theta_k + \tfrac{1}{8}[\Theta_k^2 - (\Theta_k - \delta\Theta_k)^2] + \text{HOT} \\
&= \tfrac{1}{2}\delta\Theta_k + \tfrac{1}{8}(\Theta_k\delta\Theta_k + \delta\Theta_k\Theta_k - \delta\Theta_k^2) + \text{HOT}.
\end{aligned} \tag{26}$$

We recall here that the matrix Θ_k is an incremental matrix based on the incremental time Δt; that is, Θ_k is of order Δt. Therefore, the term HOT (which stands for higher order terms) in the last line of (26) refers to terms of order $\delta\Theta_k^3$, or $\delta\Theta_k^2\Delta t$, or $\delta\Theta_k\Delta t^2$, and higher. Assuming that the norm of the noise matrix $\delta\Theta_k$ is small and that the time interval Δt between two gyro readouts is also small, the matrix $\Delta\Phi_k$ is approximated by the first term on the right-hand side (RHS) of (26); that is

$$\Delta\Phi_k \simeq \tfrac{1}{2}\delta\Theta_k. \tag{27}$$

The expression for the 4×4 matrix $\delta\Theta_k$ is obtained, of course, by substituting $\boldsymbol{\delta\theta}_k$ into $\boldsymbol{\theta}_k^o$ of (22b). Inserting (27) in (25) and using the resulting equation in the nominal discrete dynamics equation (23) yields

$$\mathbf{q}_{k+1} \simeq \Phi_k\mathbf{q}_k - \tfrac{1}{2}\delta\Theta_k\mathbf{q}_k. \tag{28}$$

Then, using (15) where $\mathbf{x} = \boldsymbol{\theta}_k^o$ and $\mathbf{y} = \mathbf{q}_k$ we obtain

$$\mathbf{q}_{k+1} \simeq \Phi_k\mathbf{q}_k - \tfrac{1}{2}\Xi_k\boldsymbol{\delta\theta}_k \tag{29}$$

where Ξ_k denotes $\Xi(\mathbf{q}_k)$. Equation (29) is a first-order approximation in $\boldsymbol{\delta\theta}_k$ of the exact process equation (23). One can notice that (29) does not preserve the unit-norm property of the quaternion. However, simulations show that for sufficiently small values of $\boldsymbol{\delta\theta}_k$, e.g., $\|\boldsymbol{\delta\theta}_k\| = 10^{-4}$ rad, the difference between (23) and (29) is negligible for all practical purposes. Note that the value of 10^{-4} rad is conservative, and corresponds to low-grade gyros. For missions with high attitude accuracy requirements, using, e.g., star-trackers as attitude sensing devices, such an error will probably not be reached, which makes the proposed model even less sensitive to modeling errors. The mathematical modeling of the RIG noises chosen here follows the one presented in [1, pp. 268–270]. The RIG noise $\boldsymbol{\delta\theta}_k$ is modeled as being composed of electronic noise, float torque noise, and float torque derivative noise. Electronic noise $\mathbf{n}_{1,k}$ and float torque noise $\mathbf{n}_{2,k}$ are modeled here as zero-mean white Gaussian sequences of standard deviations σ_1 and $\sigma_2\sqrt{\Delta t}$, respectively, where Δt denotes the time interval between two consecutive RIG readouts. The principal source of error is drift rate instability. The gyro output drift rate $\boldsymbol{\mu}_k$ is modeled as a random walk driven by a zero-mean white Gaussian sequence $\mathbf{n}_{3,k}$ of standard deviation $\sigma_3\sqrt{\Delta t}$. The three sequences $\mathbf{n}_{1,k}$, $\mathbf{n}_{2,k}$, and $\mathbf{n}_{3,k}$, are assumed to be uncorrelated with one another and with the initial values of $\boldsymbol{\mu}_k$ and $\mathbf{q}_k$. Consequently, the resulting equations describing the RIG output model are

$$\boldsymbol{\theta}_k = \boldsymbol{\theta}_k^o + \boldsymbol{\delta\theta}_k \tag{30a}$$

$$\boldsymbol{\delta\theta}_k = \mathbf{n}_{1,k} + \mathbf{n}_{2,k} + \boldsymbol{\mu}_k\Delta t \tag{30b}$$

$$\boldsymbol{\mu}_{k+1} = \boldsymbol{\mu}_k + \mathbf{n}_{3,k}. \tag{30c}$$

Using (30b) in (29) yields

$$\mathbf{q}_{k+1} = \Phi_k\mathbf{q}_k - \frac{\Delta t}{2}\Xi_k\boldsymbol{\mu}_k - \frac{1}{2}\Xi_k\mathbf{n}_{1,k} - \frac{1}{2}\Xi_k\mathbf{n}_{2,k}. \tag{31}$$

For this model we use the customary technique of state vector augmentation (see e.g. [14, p. 52]) to incorporate the drift variable $\boldsymbol{\mu}_k$ in the state vector. Define an augmented seven-dimensional state vector $\mathbf{x}_k$ as

$$\mathbf{x}_k^T \triangleq [\mathbf{q}_k^T \;\; \boldsymbol{\mu}_k^T]. \tag{32}$$

Using (30c) and (31), the derivation of the process equation of the augmented state-space model is straightforward and is summarized in the following.

Summary of the Truth Model

The process and measurement equations of the system are as follows

$$\mathbf{x}_{k+1} = \Psi_k\mathbf{x}_k + \Gamma_k\mathbf{n}_k \tag{33}$$

$$\mathbf{0} = \bar{H}_{k+1}\mathbf{x}_{k+1} - \tfrac{1}{2}\Xi_{k+1}\boldsymbol{\delta}\mathbf{b}_{k+1} \tag{34}$$

where the 7×7 transition matrix Ψ_k, the 7×9 process noise input matrix Γ_k, and the 9×1 augmented gyro noise vector $\mathbf{n}_k$ are

$$\Psi_k = \begin{bmatrix} \Phi_k & -\frac{\Delta t}{2}\Xi_k \\ O & I \end{bmatrix}$$

$$\Gamma_k = \begin{bmatrix} -\frac{1}{2}\Xi_k & -\frac{1}{2}\Xi_k & O \\ O & O & I \end{bmatrix} \tag{35}$$

$$\mathbf{n}_k = \begin{bmatrix} \mathbf{n}_{1,k} \\ \mathbf{n}_{2,k} \\ \mathbf{n}_{3,k} \end{bmatrix}$$

where Ξ_k is defined from (16), and O and I are the null matrix and the identity matrix, respectively, of appropriate dimensions. The matrix Φ_k is computed from the measured angular displacement $\boldsymbol{\theta}_k$,

$$\Theta_k = \Upsilon(\boldsymbol{\theta}_k) \tag{36}$$

$$\Phi_k = \exp(\tfrac{1}{2}\Theta_k). \tag{37}$$

The augmented observation matrix $\bar{H}_{k+1}$ is computed from the vector observations at t_{k+1}, $(\mathbf{b}_{k+1}, \mathbf{r}_{k+1})$,

$$\mathbf{s}_{k+1} = \tfrac{1}{2}(\mathbf{b}_{k+1} + \mathbf{r}_{k+1}) \tag{38a}$$

$$\mathbf{d}_{k+1} = \tfrac{1}{2}(\mathbf{b}_{k+1} - \mathbf{r}_{k+1}) \tag{38b}$$

$$H_{k+1} = \begin{bmatrix} -[\mathbf{s}_{k+1}\times] & \mathbf{d}_{k+1} \\ -\mathbf{d}_{k+1}^T & 0 \end{bmatrix} \tag{38c}$$

$$\bar{H}_{k+1} = [H_{k+1} \;\; O]. \tag{38d}$$

The 4×3 matrices Ξ_k and Ξ_{k+1} denote $\Xi(\mathbf{q}_k)$ and $\Xi(\mathbf{q}_{k+1})$, respectively, and the mapping $\Xi(\cdot)$ is defined in (7). The augmented gyro noise $\mathbf{n}_k$ is a 9×1 zero-mean white Gaussian sequence with covariance matrix $Q_k = \mathrm{diag}\{\sigma_1^2 I, \sigma_2^2 \Delta t I, \sigma_3^2 \Delta t I\}$. The measurement noise $\boldsymbol{\delta}\mathbf{b}_{k+1}$ is a 3×1 zero-mean white Gaussian noise sequence with covariance matrix R_{k+1}. All noises are statistically independent from one another and from the initial state.

NOISES STOCHASTIC MODELS

In this section, exact expressions for the covariance matrices of the state-dependent system noises are provided. The derivation of these expressions takes advantage of the fact that the input matrices of the process noises $\mathbf{n}_{1,k}$ and $\mathbf{n}_{2,k}$ (35), and of the measurement noise $\boldsymbol{\delta}\mathbf{b}_k$ (34) are linear matrix functions of the state. The similarity between the expressions for the state-dependent noises lends itself to a general treatment, which is provided in Appendix A. We apply these general results to our particular model here. First, we treat the measurement equation in detail, then, due to the similarity, the process equation is handled in a similar manner but more concisely. As shown in the following, an extra term in the expressions for the covariance matrices of the state-dependent noises yields an improvement over the usual first-order approximation.

Covariance Matrix of the Novel Measurement

Let $\hat{\mathbf{q}}_{k+1}$ denote the expected value of $\mathbf{q}_{k+1}$, let $\mathbf{y}_{k+1}$ denote the measurement vector at time t_{k+1} (which is $\mathbf{0}$ in the proposed model), and let $\mathbf{v}_{k+1}$ denote the quaternion measurement error. The measurement equation is rewritten here for convenience

$$\mathbf{y}_{k+1} = H_{k+1}\mathbf{q}_{k+1} + \mathbf{v}_{k+1} \tag{39}$$

where

$$\mathbf{v}_{k+1} = -\tfrac{1}{2}\Xi(\mathbf{q}_{k+1})\boldsymbol{\delta}\mathbf{b}_{k+1} \tag{40}$$

and, for the sake of clarity, the explicit dependence on $\mathbf{q}_{k+1}$ is expressed. Let P_{k+1}^q and P_{k+1}^v denote the covariance matrices of $\mathbf{q}_{k+1}$ and $\mathbf{v}_{k+1}$, respectively. Assuming that $\boldsymbol{\delta}\mathbf{b}_{k+1}$ and $\mathbf{q}_{k+1}$ are independent, and that $\boldsymbol{\delta}\mathbf{b}_{k+1}$ is a zero-mean white sequence with covariance matrix R_{k+1}, then the exact expression for P_{k+1}^v is

$$P_{k+1}^v = \tfrac{1}{4}\Xi(\hat{\mathbf{q}}_{k+1})R_{k+1}\Xi^T(\hat{\mathbf{q}}_{k+1}) + \tfrac{1}{4}\bar{\mathcal{E}}(R_{k+1} \otimes P_{k+1}^q)\bar{\mathcal{E}}^T \tag{41}$$

where $\hat{\mathbf{q}}_{k+1}$ denotes the expected value of $\mathbf{q}_{k+1}$ and $\otimes$ denotes the Kronecker product. The matrix $\bar{\mathcal{E}} \in \mathbb{R}^{4\times 12}$ in (41) is defined as follows:

$$\bar{\mathcal{E}} = [E_1 \;\; E_2 \;\; E_3] \tag{42}$$

where

$$E_i = \Upsilon(\mathbf{e}_i), \qquad i = 1,2,3 \tag{43}$$

The mapping $\Upsilon(\cdot)$ is defined in (7), and the vectors $\mathbf{e}_i$, $i = 1,2,3$, are the standard basis vectors in $\mathbb{R}^3$. The derivation of (41) is a direct application of the proposition presented in Appendix A.

REMARK 1 One notices that the first term on the RHS of (41) constitutes the usual approximation to the covariance matrix of a state-dependent noise; that is, the true state is simply replaced by its expected value. The second term on the RHS of (41) is due to the incorporation of the error $\mathbf{q}_{k+1} - \hat{\mathbf{q}}_{k+1}$ into the developments. That term involves expectations of products of the components of $\boldsymbol{\delta}\mathbf{b}_{k+1}$ with the components of the error $\mathbf{q}_{k+1} - \hat{\mathbf{q}}_{k+1}$. As shown in Appendix A, the development of (41) exploits the linearity of the state-dependent noise with respect to $\mathbf{q}_{k+1}$. Note that the second term on the RHS of (41) is beneficial in that it renders P_{k+1}^v nonsingular (notice that the rank of the first term on the RHS of (41) is at most three. In practical implementation of the filter, the vector $\hat{\mathbf{q}}_{k+1}$ and the matrix P_{k+1}^q are replaced by their best available estimates; that is, $\hat{\mathbf{q}}_{k+1/k}$ and $P_{k+1/k}$, respectively.

REMARK 2 No particular assumptions concerning the structure of the matrix R_{k+1} have been made up to now. Thus (41) allows to use any type of covariance matrix for $\delta\mathbf{b}_k$. On the other hand, the high dimensions of the matrix $\bar{\mathcal{E}}$ and the use of the Kronecker product for computing the extra term increase the computational burden of the model. Moreover, since this correction is only of second order, the real value in its computation might be questionable. Fortunately, however, the type of the matrix R_{k+1} that is usually considered in attitude determination problems from vector measurements will allow us to obtain fairly simple expressions for P_k^v.

Special Cases for the Covariance Matrix of the State-Dependent Measurement Noise: We consider the following three cases for R_{k+1} and show how to simplify the expression for the covariance matrix of the state-dependent noise.

Case 1 $R_{k+1} = \rho_{k+1} I_3$. We first use a known property of the matrix $\Xi(\cdot)$; namely $\Xi(\hat{\mathbf{q}}_k)\Xi^T(\hat{\mathbf{q}}_k) = I_4 - \hat{\mathbf{q}}_k\hat{\mathbf{q}}_k^T$, where it is assumed that $\hat{\mathbf{q}}_k$ is of unit-norm, to obtain

$$\Xi(\hat{\mathbf{q}}_{k+1})R_{k+1}\Xi^T(\hat{\mathbf{q}}_{k+1}) = \rho_{k+1}(I_4 - \hat{\mathbf{q}}_{k+1}\hat{\mathbf{q}}_{k+1}^T). \quad (44)$$

Then, substituting $R_{k+1} = \rho_{k+1} I_3$ in the correction term of (41) yields (see [13, Appendix C])

$$\bar{\mathcal{E}}(R_{k+1} \otimes P_{k+1}^q)\bar{\mathcal{E}}^T = \rho_{k+1}[\mathrm{tr}(P_{k+1}^q) I_4 - P_{k+1}^q]. \quad (45)$$

Using (44) and (45) in (41) yields

$$\begin{aligned} P_{k+1}^v &= \tfrac{1}{4}\rho_{k+1}\{[1 + \mathrm{tr}(P_{k+1}^q)] I_4 - (\hat{\mathbf{q}}_{k+1}\hat{\mathbf{q}}_{k+1}^T + P_{k+1}^q)\} \\ &= \tfrac{1}{4}\rho_{k+1}[\mathrm{tr}(M_{k+1}) I_4 - M_{k+1}] \end{aligned} \quad (46)$$

where

$$M_{k+1} = \hat{\mathbf{q}}_{k+1}\hat{\mathbf{q}}_{k+1}^T + P_{k+1}^q. \quad (47)$$

Let $\Lambda = [\mathrm{tr}(M_{k+1}) I_4 - M_{k+1}]$. If P_{k+1}^q is positive definite, then Λ is also positive definite as shown next. Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4 > 0$ denote the four positive eigenvalues of P_{k+1}^q, then the spectrum of Λ can be easily shown to be $\{1 + \lambda_1 + \lambda_2 + \lambda_3, 1 + \lambda_1 + \lambda_2 + \lambda_4, 1 + \lambda_1 + \lambda_3 + \lambda_4, \lambda_2 + \lambda_3 + \lambda_4\}$ where all the eigenvalues are positive. Therefore, the measurement noise covariance matrix is positive definite, as it should be.

Case 2 $R_{k+1} = \rho_{k+1}(I_3 - \mathbf{b}_{k+1}\mathbf{b}_{k+1}^T)$. This is a fairly accurate approximation for the covariance matrix of a unit-vector measurement [15]. Consider the expression

$$\Xi(\hat{\mathbf{q}})\mathbf{b}\mathbf{b}^T\Xi(\hat{\mathbf{q}}) + \bar{\mathcal{E}}(\mathbf{b}\mathbf{b}^T \otimes P^q)\bar{\mathcal{E}}^T \quad (48)$$

which is the part of the covariance that corresponds to $\mathbf{b}_{k+1}\mathbf{b}_{k+1}^T$, and where the time indices were omitted for the sake of notational simplicity. Let $B \in \mathbb{R}^{4\times 4}$ denote the skew-symmetric matrix $\Upsilon(\mathbf{b})$, then, according to (15)

$$\Xi(\hat{\mathbf{q}})\mathbf{b}\mathbf{b}^T\Xi^T(\hat{\mathbf{q}}) = B\hat{\mathbf{q}}\hat{\mathbf{q}}^T B^T. \quad (49)$$

It can be shown (see [13, Appendix D]) that

$$\bar{\mathcal{E}}(\mathbf{b}\mathbf{b}^T \otimes P^q)\bar{\mathcal{E}}^T = BP^qB^T. \quad (50)$$

Subtracting (49) and (50) from (46) yields

$$P_{k+1}^v = \tfrac{1}{4}\rho_{k+1}[\mathrm{tr}(M_{k+1}) I_4 - M_{k+1} - B_{k+1}M_{k+1}B_{k+1}^T] \quad (51)$$

where M_{k+1} is defined in (47).

Case 3 $R_{k+1} = \sum_{i=1}^3 \nu_i \mathbf{u}_i\mathbf{u}_i^T$. Let $U_i \in \mathbb{R}^{3\times 3}$ denote $\Upsilon(\mathbf{u}_i)$, $i = 1,2,3$, then P_{k+1}^v is expressed as (see [13, Appendix E])

$$P_{k+1}^v = \frac{1}{4}\sum_{i=1}^3 \nu_i U_i M_{k+1} U_i^T \quad (52)$$

where M_{k+1} is defined in (47).

Covariance Matrix of the State-Dependent Process Noise

Similar developments to those of the previous section lead us to an exact and a simplified expression for the covariance matrix of the state-dependent process noise in (33). Since we assume that both gyro noises $\mathbf{n}_{1,k}$ and $\mathbf{n}_{2,k}$ have covariance matrices of the form $\sigma^2 I_3$, we can directly implement the result of Case 1 of the previous subsection. We obtain the following expression for the covariance matrix of $\mathbf{w}_k = \Gamma_k\mathbf{n}_k$ in (33) as

$$\begin{aligned} P_k^w &\triangleq \mathrm{cov}\{\Gamma_k\mathbf{n}_k\} \\ &= \begin{bmatrix} (\sigma_1^2 + \sigma_2^2\Delta t)\tfrac{1}{4}[\mathrm{tr}(M_k) I_4 - M_k] & O \\ O & \sigma_3^2\Delta t I_3 \end{bmatrix} \end{aligned} \quad (53)$$

with

$$M_k = \hat{\mathbf{q}}_k\hat{\mathbf{q}}_k^T + P_k^q \quad (54)$$

where $\hat{\mathbf{q}}_k$ and P_k^q denote, respectively, the expectation and the covariance matrix of the state $\mathbf{q}_k$. The expression for P_k^w in (53) will serve as a basis for designing the covariance in the prediction stage of the KF.

QUATERNION KALMAN FILTER

The issue of state-dependence of the model parameters must be addressed in order to implement a KF. We adopt here the common approach of replacing every unknown variable in the model parameters by its best available estimate. First, the true quaternion $\mathbf{q}_k$ is replaced by its a posteriori estimate $\hat{\mathbf{q}}_{k/k}$ when computing the transition matrix Ψ_k (35). Next, the state-dependent process noise covariance matrix is

computed as follows (see (53) and (54))

$$P_k^w = \begin{bmatrix} (\sigma_1^2 + \sigma_2^2 \Delta t)\frac{1}{4}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}] & O \\ O & \sigma_3^2 \Delta t I_3 \end{bmatrix} \tag{55}$$

where

$$\hat{M}_{k/k} = \hat{\mathbf{q}}_{k/k}\hat{\mathbf{q}}_{k/k}^T + P_{k/k}^q \tag{56}$$

and $P_{k/k}^q$ is the computed a posteriori quaternion estimation error covariance matrix. It is assumed that the measurement vectors $\mathbf{b}_{k+1}$ are unitized, as is commonly done when using line-of-sight measurements. In this case, the suitable model for the matrix R_{k+1} is [15]: $R_{k+1} = \rho_{k+1}(I_3 - \mathbf{b}_{k+1}\mathbf{b}_{k+1}^T)$, which corresponds to Case 2 of the previous section, and where ρ_{k+1} is a given scalar variance parameter. The adequate expression for the state-dependent noise covariance matrix is thus provided in (51) and (47). Accordingly, the measurement noise covariance matrix used by the filter is

$$P_{k+1}^v = \tfrac{1}{4}\rho_{k+1}[\mathrm{tr}(\hat{M}_{k+1/k})I_4 - \hat{M}_{k+1/k} - B_{k+1}\hat{M}_{k+1/k}B_{k+1}^T] \tag{57}$$

where

$$\hat{M}_{k+1/k} = \hat{\mathbf{q}}_{k+1/k}\hat{\mathbf{q}}_{k+1/k}^T + P_{k+1/k}^q \tag{58}$$

$\hat{\mathbf{q}}_{k+1/k}$ is the a priori quaternion estimate, $P_{k+1/k}^q$ is the a priori quaternion estimation error covariance matrix, and $B_{k+1} = \Upsilon(\mathbf{b}_{k+1})$.

The KF is not designed to preserve any particular relationship among the components of the estimate vector. It is known however that, in order for a quaternion to represent a rotation, its length has to be equal to one. Moreover, some of the previous developments on the state-dependent noise covariance matrices relied on the unity-assumption of $\hat{\mathbf{q}}_{k/k}$. There are several ways of introducing constraints in a KF. Adding a pseudo-measurement[11] would increase the filter computational burden. Truncating or reducing the state-space model[9] would prevent us from using the proposed linear measurement model equation (see (39)). Therefore we opt for the simple quaternion normalization; namely,

$$\mathbf{q}_{k+1/k+1}^* = \frac{\hat{\mathbf{q}}_{k+1/k+1}}{\|\hat{\mathbf{q}}_{k+1/k+1}\|} \tag{59}$$

which was efficiently incorporated in a previous quaternion KF [10], and is also used one way or another in the multiplicative EKF [9].

Summary of the Algorithm

Using the previous developments, the novel quaternion KF is summarized as follows.

Filter Initialization: Choose the appropriate values for σ_1, σ_2, σ_3 and ρ_{k+1}, and choose an approximate value for the initial estimate of the state vector and for the initial estimation error covariance matrix $P_{0/0}$. In the absence of such initial estimate, choose the zero attitude quaternion; that is, $\hat{\mathbf{q}}_{0/0}^T = [0,0,0,1]$, and zero initial drift estimates, $\hat{\boldsymbol{\mu}}_{0/0}^T = [0,0,0]$, to form

$$\hat{\mathbf{x}}_{0/0}^T = [\hat{\mathbf{q}}_{0/0}^T, \hat{\boldsymbol{\mu}}_{0/0}^T]. \tag{60}$$

Time Propagation: Given $\hat{\mathbf{x}}_{k/k}^T = [{\mathbf{q}_{k/k}^*}^T, \hat{\boldsymbol{\mu}}_{k/k}^T]$ and a RIG readout $\boldsymbol{\theta}_k$,

$$\hat{\boldsymbol{\theta}}_{k/k} = \boldsymbol{\theta}_k - \hat{\boldsymbol{\mu}}_{k/k}\Delta t \tag{61a}$$

$$\hat{\Theta}_{k/k} = \Upsilon(\hat{\boldsymbol{\theta}}_{k/k}) \tag{61b}$$

$$\hat{\Phi}_{k/k} = \exp(\tfrac{1}{2}\hat{\Theta}_{k/k}) \tag{61c}$$

$$\hat{\mathbf{x}}_{k+1/k} = \begin{bmatrix} \hat{\Phi}_{k/k} & O \\ O & I \end{bmatrix} \hat{\mathbf{x}}_{k/k} \tag{61d}$$

$$\Theta_k = \Upsilon(\boldsymbol{\theta}_k) \tag{61e}$$

$$\Phi_k = \exp(\tfrac{1}{2}\Theta_k) \tag{61f}$$

$$\hat{\Xi}_{k/k} = \Xi(\hat{\mathbf{q}}_{k/k}) \tag{61g}$$

$$\Psi_k = \begin{bmatrix} \Phi_k & -\dfrac{\Delta t}{2}\hat{\Xi}_{k/k} \\ O & I \end{bmatrix} \tag{61h}$$

Extract $P_{k/k}^q$ from $P_{k/k}$, then

$$\hat{M}_{k/k} = \hat{\mathbf{q}}_{k/k}\hat{\mathbf{q}}_{k/k}^T + P_{k/k}^q \tag{61i}$$

$$P_k^w = \begin{bmatrix} (\sigma_1^2 + \sigma_2^2 \Delta t)\frac{1}{4}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}] & O \\ O & \sigma_3^2 \Delta t I_3 \end{bmatrix} \tag{61j}$$

$$P_{k+1/k} = \Psi_k P_{k/k} \Psi_k^T + P_k^w. \tag{61k}$$

Notice that the estimate transition matrix in (61d) does not contain off-diagonal terms, whereas the transition matrix in (61h) does. The former transition matrix is due to the nonlinear propagation model equations, as given in (19) and (30c). The correction in the quaternion estimate due to the drift estimate is made via the correction of the incremental rotation, $\hat{\boldsymbol{\theta}}_{k/k}$, in (61a). Thus, the propagation stage preserves the unit-norm property of the quaternion estimate. The transition matrix Ψ_k, on the other hand, stems from the process model equation as given in (33) and (35), and is used for the covariance propagation. The off-diagonal term is necessary for the convergence of the filter. In the absence of that term, and because of the structure of the augmented measurement matrix [see (62d)], the system would not be completely observable, and the drifts would not be estimated.

Measurement Update: Given $\hat{\mathbf{x}}_{k+1/k} = [\hat{\mathbf{q}}^T_{k+1/k}, \hat{\boldsymbol{\mu}}^T_{k+1/k}]$ and the new pair $(\mathbf{b}_{k+1}, \mathbf{r}_{k+1})$,

$$\mathbf{s}_{k+1} = \tfrac{1}{2}(\mathbf{b}_{k+1} + \mathbf{r}_{k+1}) \tag{62a}$$

$$\mathbf{d}_{k+1} = \tfrac{1}{2}(\mathbf{b}_{k+1} - \mathbf{r}_{k+1}) \tag{62b}$$

$$H_{k+1} = \begin{bmatrix} -[\mathbf{s}_{k+1}\times] & \mathbf{d}_{k+1} \\ -\mathbf{d}^T_{k+1} & 0 \end{bmatrix} \tag{62c}$$

$$\bar{H}_{k+1} = [H_{k+1} \;\; O] \tag{62d}$$

Extract $P^q_{k+1/k}$ from $P_{k+1/k}$, then

$$\hat{M}_{k+1/k} = \hat{\mathbf{q}}_{k+1/k}\hat{\mathbf{q}}^T_{k+1/k} + P^q_{k+1/k} \tag{62e}$$

$$B_{k+1} = \Upsilon(\mathbf{b}_{k+1}) \tag{62f}$$

$$P^v_{k+1} = \tfrac{1}{4}\rho_{k+1}[\mathrm{tr}(\hat{M}_{k+1/k})I_4 - \hat{M}_{k+1/k} - B_{k+1}\hat{M}_{k+1/k}B^T_{k+1}] \tag{62g}$$

$$S_{k+1/k} = H_{k+1}P^q_{k+1/k}H^T_{k+1} + P^v_{k+1} \tag{62h}$$

$$K_{k+1} = P_{k+1/k}\bar{H}^T_{k+1}S^{-1}_{k+1/k} \tag{62i}$$

$$\hat{\mathbf{q}}_{k+1/k+1} = (I_4 - K_{k+1}\bar{H}_{k+1})\hat{\mathbf{q}}_{k+1/k} \tag{62j}$$

$$P_{k+1/k+1} = (I_4 - K_{k+1}\bar{H}_{k+1})P_{k+1/k}(I_4 - K_{k+1}\bar{H}_{k+1})^T + K_{k+1}P^v_{k+1}K^T_{k+1} \tag{62k}$$

$$\mathbf{q}^*_{k+1/k+1} = \frac{\hat{\mathbf{q}}_{k+1/k+1}}{\|\hat{\mathbf{q}}_{k+1/k+1}\|}. \tag{62l}$$

ADAPTIVE FILTER

Motivation

It is well known that the KF is an unbiased minimum-variance state estimator under the assumptions that the system truth model is linear and the design model is identical to the truth model. However, the KF becomes suboptimal and sensitive to initial conditions if there are errors in the design model. This is the case of the current model because, due to the difference between the physical and the mathematical models, the system equations, (33) and (34), contain modeling errors, and because the state-dependent noise covariance matrices, P^w_k and P^v_{k+1} in (55) and (57), are approximated in the filter by estimate-dependent expressions as given in (61j) and (62g), respectively. An adequate approach to on line enhancement of the filter performance is adaptive filtering [16, pp. 311–318]. In this section we present an adaptive filter that processes the measurement residuals in order to optimally compensate for modeling errors according to some performance index. The technique of process noise adaptive estimation is employed; that is, the filter process noise level is changed adaptively in order to compensate for modeling errors. The adaptive filter is applied to two types of modeling errors: 1) errors in the values of the gyro noise levels, σ_1, σ_2, and σ_3, and 2) unmodeled biases in the random walk model of the gyro drifts (30c).

Approach

The approach presented herein is inspired by Jazwinski [16, p. 311], who considered the case of a scalar Gaussian measurement noise with a single predicted residual processed by an adaptive filter. In this work, however, we consider vector measurements and propose an ad-hoc extension of Jazwinski's technique. Assume that a measurement was acquired at time t_{k+1}. In our case of zero measurement (see (17)), $\boldsymbol{\nu}_{k+1/k}$, the measurement residuals process at t_{k+1}, is given by

$$\boldsymbol{\nu}_{k+1/k} = -H_{k+1}\hat{\mathbf{q}}_{k+1/k}. \tag{63}$$

Furthermore, the covariance matrix Q_k of the augmented gyro noise vector is modeled by

$$Q_k = \mathrm{diag}\{\eta_1 I_3, \eta_2 \Delta t I_3, \eta_3 \Delta t I_3\} \tag{64}$$

where η_1, η_2, and η_3 are positive parameters to be estimated. Defining $\boldsymbol{\eta} \triangleq [\eta_1, \eta_2, \eta_3]^T \in \mathbb{R}^3$, and recalling that $S_{k+1/k}$, the covariance matrix of $\boldsymbol{\nu}_{k+1/k}$, is computed in the filter and is a function of the value of $\boldsymbol{\eta}$ assumed in the filter, we propose to solve the following minimization problem:

$$\min_{\boldsymbol{\eta}\geq 0}\{J(\boldsymbol{\eta}) = \|\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}^T_{k+1/k} - S_{k+1/k}(\boldsymbol{\eta})\|^2\} \tag{65}$$

where the norm $\|A\|$ is the Frobenius norm; that is $\|A\|^2 \triangleq \mathrm{tr}(AA^T)$, and the inequality $\boldsymbol{\eta} \geq 0$ is component-wise. The value of $\boldsymbol{\eta}$ that is computed from (65) represents the optimal estimate of the process noise levels.

The rationale for the cost function (65) is explained as follows. The matrix $\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}^T_{k+1/k}$ is the residual sample covariance matrix, whereas $S_{k+1/k}$ is the covariance matrix of $\boldsymbol{\nu}_{k+1/k}$, computed using the filter design model parameters. In particular, $S_{k+1/k}$ is a function of the a priori value of $\boldsymbol{\eta}$. Hence, a good value of $\boldsymbol{\eta}$ is one that brings consistency between the sample covariance matrix $\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}^T_{k+1/k}$ and its predicted value $S_{k+1/k}$. This consistency condition is satisfied through the minimization described in (65).

Solution

Recall that $S_{k+1/k}$ is computed according to (62h), rewritten here using the augmented matrices $\bar{H}_{k+1}$ and $P_{k+1/k}$,

$$S_{k+1/k} = \bar{H}_{k+1}P_{k+1/k}\bar{H}^T_{k+1} + P^v_{k+1}. \tag{66}$$

Using (61k) in (66), yields

$$S_{k+1/k} = \bar{H}_{k+1}(\Psi_k P_{k/k} \Psi_k^T + P_k^w)\bar{H}_{k+1}^T + P_{k+1}^v$$
$$= \bar{H}_{k+1} P_k^w \bar{H}_{k+1}^T + \bar{H}_{k+1}\Psi_k P_{k/k} \Psi_k^T \bar{H}_{k+1}^T + P_{k+1}^v. \quad (67)$$

Next we evaluate the first term on the RHS of (67) using (62d) and (61j),

$$\bar{H}_{k+1} P_k^w \bar{H}_{k+1}^T$$
$$= [H_{k+1} \;\; O] \begin{bmatrix} (\eta_1 + \eta_2 \Delta t)\frac{1}{4}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}) & O \\ O & \eta_3 \Delta t I_3 \end{bmatrix}$$
$$\times [H_{k+1} \;\; O]^T$$
$$= (\eta_1 + \eta_2 \Delta t)\tfrac{1}{4} H_{k+1}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}]H_{k+1}^T$$
$$= \eta L_1 \quad (68)$$

where $\eta \triangleq \eta_1 + \eta_2 \Delta t$, and L_1 is the matrix part in the second equality of (68). Defining N_1 as

$$N_1 \triangleq \bar{H}_{k+1}\Psi_k P_{k/k} \Psi_k^T \bar{H}_{k+1}^T + P_{k+1}^v \quad (69)$$

and using (68) and (69) in (67) yields

$$S_{k+1/k}(\eta) = \eta L_1 + N_1. \quad (70)$$

It appears that η_1 and η_2 cannot be evaluated separately; only the combination η can be computed for the adaptation purposes. Moreover, the parameter η_3 does not play any role since it does not participate in the computation of the covariance matrix $S_{k+1/k}$. This stems directly from the fact that one measures only the quaternion and not the drift. The zero components in the augmented matrix $\bar{H}_{k+1}$ cancel the drift-related gyro noise variance η_3. Then, performing the minimization

$$\min_{\eta \geq 0}\{J(\eta) = \|\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T - S_{k+1/k}(\eta)\|^2\} \quad (71)$$

yields the following minimizing η

$$\eta^* = \frac{\mathrm{tr}[(\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T - N_1)L_1^T]}{\mathrm{tr}(L_1 L_1^T)}. \quad (72)$$

The derivation of (72) is provided in Appendix B, in which $A = \boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T$, $B = S_{k+1/k}$, $C = N_1$, and $D = L_1$. If η^* is positive, the previous value of η is replaced by η^* and the latter value is used to update the covariance matrix P_k^w in (61j). If η^* is negative, the value of η in P_k^w is set to 0. The reason for which η^* can be negative is as follows. We can interpret the matrix N_1 in (72) as the residual covariance matrix that is predicted by the filter as if there were no process noise in the system. On the other hand, the matrix $\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T$ is an approximate value for the residual covariance as computed from the actual measurements. If N_1 happens to be greater than $\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T$, this means that the filter is too conservative and should lower its level of process noise in order to fit the actual residual level. Since it can do this only by regulating the value of η, it would choose a negative value. However, we rule out such a value in order to ensure that the time propagation stage only increases the estimation error covariance, as should be the case in a well tuned KF.

Algorithm

Assume that the a posteriori estimation error $P_{k/k}$ and the a priori estimate $\hat{\mathbf{q}}_{k+1/k}$ were computed. Then the adaptive process noise algorithm is as follows

$$\boldsymbol{\nu}_{k+1/k} = -H_{k+1}\hat{\mathbf{q}}_{k+1/k} \quad (73a)$$
$$L_1 = \tfrac{1}{4} H_{k+1}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}]H_{k+1}^T \quad (73b)$$
$$N_1 = \bar{H}_{k+1}\Psi_k P_{k/k}\Psi_k^T \bar{H}_{k+1}^T + P_{k+1}^v \quad (73c)$$
$$\eta^* = \frac{\mathrm{tr}[(\boldsymbol{\nu}_{k+1/k}\boldsymbol{\nu}_{k+1/k}^T - N_1)L_1^T]}{\mathrm{tr}(L_1 L_1^T)} \quad (73d)$$
$$P_k^w = \begin{bmatrix} \eta^* \frac{1}{4}[\mathrm{tr}(\hat{M}_{k/k})I_4 - \hat{M}_{k/k}] & O \\ O & \eta_3 \Delta t I_3 \end{bmatrix}. \quad (73e)$$

Equations (73) describe the adaptive part of the quaternion KF. The result of this algorithm, that is the value of P_k^w, is substituted into (61j) in the general algorithm.

Simplified Computation of η^*: The denominator in (73d), $\mathrm{tr}(L_1 L_1^T)$, can be approximated to second-order in the measurement and estimation errors by the value of $\frac{1}{8}$. The above approximation is justified as follows. First, in the evaluation of $\hat{M}_{k/k}$ in (61i), we neglect the covariance matrix $P_{k/k}^q$ with respect to the matrix $\hat{\mathbf{q}}_{k/k}\hat{\mathbf{q}}_{k/k}^T$, so that $\hat{M}_{k/k} \simeq \hat{\mathbf{q}}_{k/k}\hat{\mathbf{q}}_{k/k}^T$, and therefore, $\mathrm{tr}(\hat{M}_{k/k}) \simeq 1$. This yields

$$L_1 \simeq \tfrac{1}{4}[H_{k+1}H_{k+1}^T - (H_{k+1}\hat{\mathbf{q}}_{k/k})(H_{k+1}\hat{\mathbf{q}}_{k/k})^T]. \quad (74)$$

Then, in (74), we neglect the second matrix in the brackets with respect to $H_{k+1}H_{k+1}^T$. This is justified by the fact that $\mathbf{q}_{k/k}$ is close to the null-space of H_{k+1} provided that the filter converged and that Δt is small. The previous argument yields the following approximation for $L_1 L_1^T$ to second order in the measurement and the estimation errors,

$$L_1 \simeq \tfrac{1}{4} H_{k+1}H_{k+1}^T. \quad (75)$$

Finally we use the property that the matrix H_{k+1} is skew-symmetric with eigenvalues $\{0, 0, j, -j\}$ (provided that $\mathbf{b}_{k+1}$ and $\mathbf{r}_{k+1}$ are unitized), which yields

$$\mathrm{tr}(L_1 L_1^T) \simeq \frac{\mathrm{tr}(H_{k+1}^4)}{16} = \frac{1}{8}. \quad (76)$$

TABLE I
Noise Standard Deviations

		σ_1 arcs	σ_2 arcs/s$^{1/2}$	σ_3 arcs/s$^{3/2}$	σ_b
Case A	system	0.5	6	7.10^{-3}	1 deg
	QKF	0.5	6	10^{-2}	0.2 deg
Case B	system	0.5	6	7.10^{-3}	100 arcs
	QKF	0.5	6	10^{-2}	100 arcs
	AEKF	0.5	6	7.10^{-3}	100 arcs
Case C	system	0.5	60	7.10^{-3}	100 arcs
	QKF	0.05	6	7.10^{-3}	100 arcs
Case D	system	0.5	6	7.10^{-3}	100 arcs
	QKF	0.5	6	7.10^{-3}	100 arcs

This analysis was supported by simulations where the error in the approximation remained at the order of 1% for reasonable levels of noises (see Table I).

BRIEF REVIEW OF THE QUATERNION ADDITIVE EXTENDED KALMAN FILTER

One purpose of the simulation is to show the advantage of the new algorithm over the quaternion additive extended Kalman filter (AEKF) [10]. To investigate the effect of the novel measurement model, we note that the AEKF and the novel quaternion KF differ only in their measurement update stage, because that is where linearization is being used in the AEKF. The goal of this section is to briefly review the typical measurement update stage of the AEKF. For convenience we rewrite the nonlinear vector measurement equation (2)

$$\mathbf{b}_{k+1} = A(\mathbf{q}_{k+1})\mathbf{r}_{k+1} + \boldsymbol{\delta}\mathbf{b}_{k+1} \tag{77}$$

where A is a known function of $\mathbf{q}_{k+1}$ (see (3)). Denoting by $\mathbf{v}(\mathbf{q}_{k+1})$ the nonlinear measurement function $A(\mathbf{q}_{k+1})\mathbf{r}_{k+1}$, the 3×4 linearized measurement matrix, when evaluated at $\hat{\mathbf{q}}_{k+1/k}$, is expressed as follows

$$\begin{aligned}\hat{H}_{k+1} &= \nabla_{\mathbf{q}}\mathbf{v}(\hat{\mathbf{q}}_{k+1/k}) \\ &= 2[\mathbf{e}^T\mathbf{r}I_3 + \mathbf{e}\mathbf{r}^T - \mathbf{r}\mathbf{e}^T + 2q[\mathbf{r}\times], q\mathbf{r} + [\mathbf{r}\times]\mathbf{e}]|_{\mathbf{q}=\hat{\mathbf{q}}_{k+1/k}}\end{aligned} \tag{78}$$

where $\mathbf{e}$ and q are the vector and scalar part of $\mathbf{q}$, respectively, and where the time indices are omitted for the sake of notational simplicity. The measurement update stage in the AEKF is then formulated as follows

$$S_{k+1/k} = \hat{H}_{k+1}P^q_{k+1/k}\hat{H}^T_{k+1} + R_{k+1} \tag{79a}$$

$$\bar{\hat{H}}_{k+1} = [\hat{H}_{k+1} \;\; O] \tag{79b}$$

$$K_{k+1} = P_{k+1/k}\bar{\hat{H}}^T_{k+1}S^{-1}_{k+1/k} \tag{79c}$$

$$\hat{\mathbf{q}}_{k+1/k+1} = \hat{\mathbf{q}}_{k+1/k} + K_{k+1}[\mathbf{b}_{k+1} - A(\hat{\mathbf{q}}_{k+1/k})\mathbf{r}_{k+1}] \tag{79d}$$

$$\begin{aligned}P_{k+1/k+1} &= (I_4 - K_{k+1}\bar{\hat{H}}_{k+1})P_{k+1/k}(I_4 - K_{k+1}\bar{\hat{H}}_{k+1})^T \\ &\quad + K_{k+1}R_{k+1}K^T_{k+1}\end{aligned} \tag{79e}$$

where R_{k+1}, the measurement error covariance matrix, is equal to $\rho_{k+1}(I_3 - \mathbf{b}_{k+1}\mathbf{b}^T_{k+1})$. In the implementation of the AEKF, (79) replace (62a) to (62k) of the novel quaternion KF. Notice that the estimation errors, which are present in the matrix $\hat{H}_{k+1}$, contaminate more terms of the covariance computation than in the case of the novel quaternion KF. Moreover, the effect of these errors is weakened in the novel quaternion KF due to the multiplication by the small noise variance parameters [$(\sigma_1^2 + \sigma_2^2\Delta t)$ in (61j), and ρ_{k+1} in (62g)]. It is anticipated, thus, that the novel quaternion KF will be less sensitive to large initial estimation errors than the AEKF. These heuristic statements will be verified through simulations.

SIMULATION STUDY

Extensive Monte-Carlo simulations were performed in order to test the nonadaptive quaternion KF, the adaptive quaternion KF, and to compare the performance of the nonadaptive quaternion KF to that of the AEKF.

Simulations Outline

In all simulations the reference coordinate system $\mathcal{R}$ was assumed to be inertial. The body coordinate system $\mathcal{B}$ was rotating with an angular velocity vector given by

$$\boldsymbol{\omega}(t) = \begin{bmatrix}1\\1\\1\end{bmatrix}\sin(2\pi t/150)\frac{\text{deg}}{\text{s}} \tag{80}$$

in $\mathcal{B}$. The initial state of the system was systematically taken as

$$\begin{aligned}\mathbf{q}_0 &= [0.3780 \;\; -0.3780 \;\; 0.7560 \;\; 0.3780]^T \\ \boldsymbol{\mu}_0 &= [1 \;\; -1 \;\; 0.5]^T\frac{\text{deg}}{\text{hr}}\end{aligned} \tag{81}$$

The time interval between two consecutive RIG readouts was $\Delta t = 250$ ms. This conservative value (gyros may have much higher sampling rates) was chosen in order to test the estimators under unfavorable conditions. The time interval between two consecutive vector measurements was 5 s. In the sequel, the standard deviation of the vector measurement error will be denoted by σ_b; that is, $\sigma_b = \sqrt{\rho_{k+1}}$. The sequence of unitized vector measurements $\mathbf{b}_{k+1}$ was generated as follows. First we generated a random sequence of unitized $\mathbf{r}_{k+1}$. Next we solved

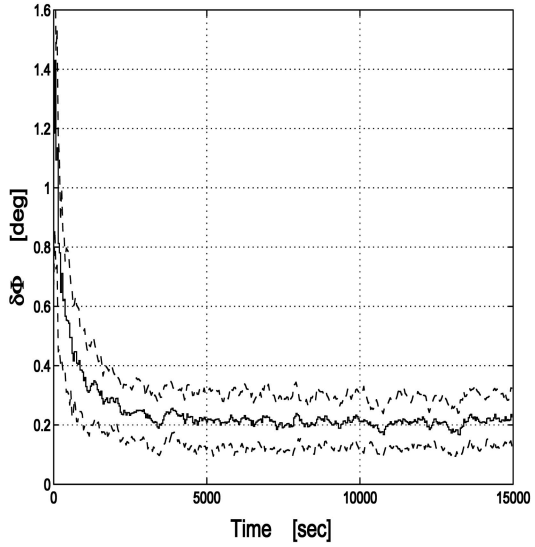


Fig. 1. Case A. Monte-Carlo mean (solid line) and $\pm 1\ \sigma$ envelope (dashed lines) of angular estimation error $\delta\phi$ of nonadaptive filter over 100 runs.

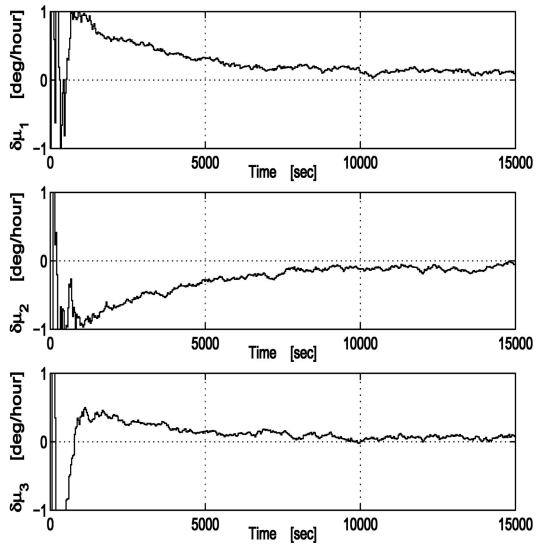


Fig. 2. Case A. Monte-Carlo means of RIG drift rates estimation errors of nonadaptive filter over 100 runs.

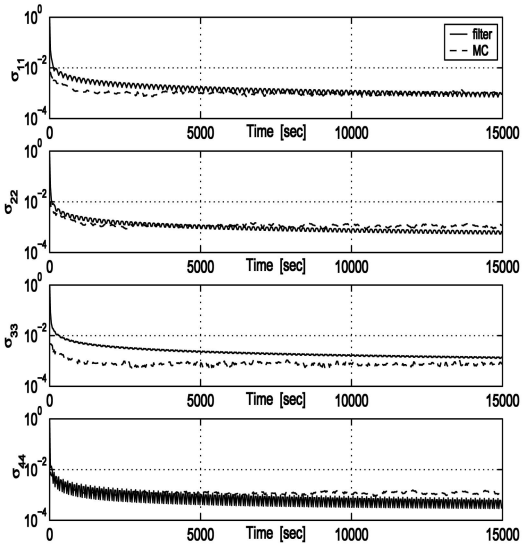


Fig. 3. Case A. Monte-Carlo standard deviations and single-run filter standard deviations in quaternion estimation errors of nonadaptive filter.

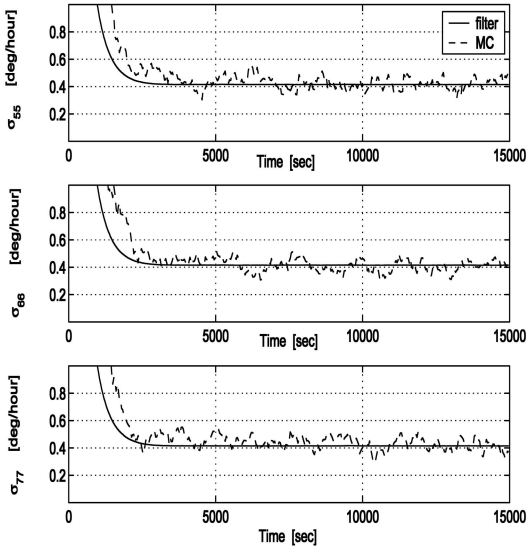


Fig. 4. Case A. Monte-Carlo standard deviations and single-run filter standard deviations in RIG drift rates estimation errors of nonadaptive filter.

(23) using the given $\omega(t)$ vector. This produced the true quaternion. Then $\mathbf{r}_{k+1}$ was rotated using the known direction cosine matrix expression which was based on the quaternion (see (3)). A low-intensity white Gaussian noise was added to the true value of $\mathbf{b}_{k+1}$, and the resulting vector was normalized. The above-mentioned quantities were common to all subsequent simulations. In each simulation the Monte-Carlo average was computed over an ensemble of 100 runs.

Case A. Nonadaptive Quaternion Kalman Filter

This section presents the performance of the nonadaptive quaternion KF, which was presented previously. The levels of the process and measurement noises in the system are summarized in Table I. The angular standard deviation $\sigma_b = 1$ deg is typical to magnetometer measurements. The filter needed some tuning for σ_3 and σ_b (See Table I for the actual values.) The filter initialization was done using the following values for the initial estimate and the initial estimation error covariance matrix

$$\hat{\mathbf{q}}_{0/0} = [0\ \ 0\ \ 0\ \ 1]^T, \qquad \hat{\boldsymbol{\mu}}_{0/0} = [0\ \ 0\ \ 0]^T, \qquad P_{0/0} = 5I_7. \tag{82}$$

Each Monte-Carlo run lasted 15000 s, which is typically the duration of three revolutions of a low-Earth-orbit satellite.

The results for the nonadaptive quaternion KF are summarized in Figs. 1–4. Fig. 1 shows the

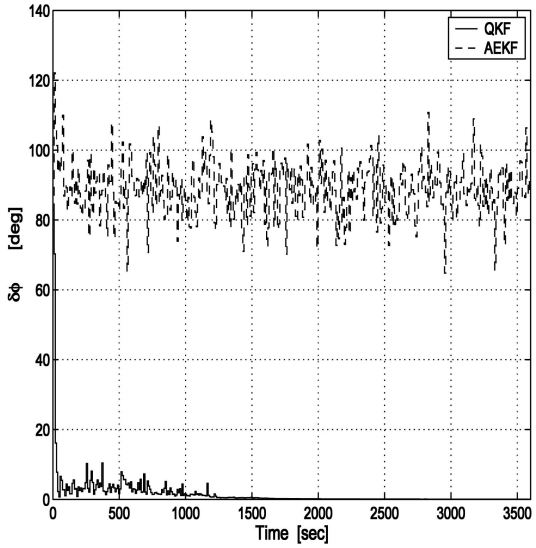


Fig. 5. Case B. Monte-Carlo means of angular estimation error $\delta\phi$ of QKF and AEKF over 100 runs.

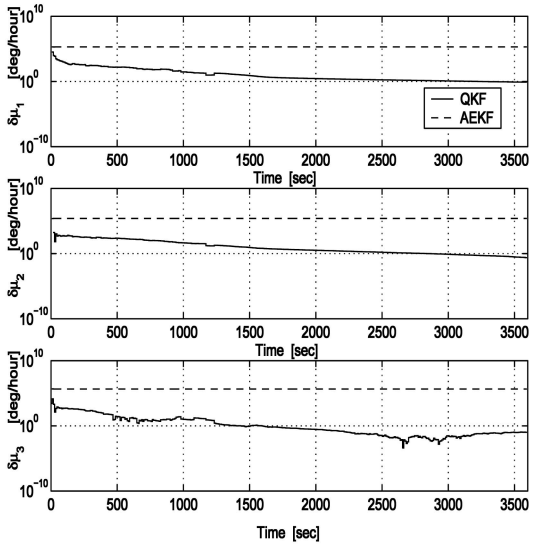


Fig. 6. Case B. Monte-Carlo means of RIG drift rates estimation errors of QKF and AEKF over 100 runs.

Monte-Carlo mean and the Monte-Carlo $\pm 1\sigma$ envelope of the angular estimation error $\delta\phi$. The angle $\delta\phi$ is defined as the angle of the small rotation that brings the estimated body frame $\hat{\mathcal{B}}$ onto the true body frame $\mathcal{B}$. This angle is obtained as follows. First, the quaternion of the rotation from $\hat{\mathcal{B}}$ to $\mathcal{B}$, denoted by $\boldsymbol{\delta q}$, is evaluated, then the rotation angle $\delta\phi$ is computed from δq, the scalar component of $\boldsymbol{\delta q}$, using the known relation [1, p. 414] $\delta\phi = 2\arccos(\delta q)$. After a transient of approximately 2500 s, the Monte-Carlo mean and the Monte-Carlo standard deviation of $\delta\phi$ settle on steady-state values of 0.2 deg and 80 mdeg, respectively. Fig. 2 depicts the Monte-Carlo means of the gyro drift estimation error $\boldsymbol{\delta}\hat{\boldsymbol{\mu}}_{k/k}$. The three components show a fast transient phase of about 1000 s followed by a slow convergence phase; their final values are equal to or below 10 mdeg/hr. Two types of standard deviations are plotted in Fig. 3 for each of the four components of the quaternion estimation error. The solid lines represent the standard deviation computed in the filter; that is, the square roots of the diagonal elements of $P^q_{k/k}$. The dashed lines represent the Monte-Carlo standard deviations of the estimation errors computed over 100 runs. The two different standard deviations are plotted in order to check the consistency of the quaternion KF under normal operational conditions. It is concluded that the filter predicts correctly the values of the estimation error variances when the level of the noises is well known. The filter variables, however, appear to be much more oscillatory than the Monte-Carlo variables. This was expected since the filter covariance computation is estimate dependent, therefore the dynamics become more visible in the estimation error variances. Fig. 4 shows the Monte-Carlo standard deviations and the filter standard deviations corresponding to the three components of the gyro drift estimation error. After a transient of 2500 s all standard deviations settle on a steady-state value of approximately 0.4 deg/hr. As seen in Fig. 4 there is a good agreement between the Monte-Carlo results and the filter computation.

Case B. Quaternion KF versus the Additive Extended Kalman Filter

In this case, the system has the same characteristics as in the previous case except for σ_b (see Table I) and for the initial estimation errors. We choose here $\sigma_b = 100$ arcs, which is a reasonable value considering the use of a star-tracker. Compared with Case A, the level of vector measurement noise was lowered in order for the effect of the initial estimation errors on the estimator performance to surface. High initial estimation errors were assumed in the filters both in attitude and in the gyro drift rates. Of course, for the sake of comparison between the quaternion KF and the AEKF, the characteristics of both filters are identical except for the measurement update formulation, as was pointed out previously. Each one of the 100 Monte-Carlo runs lasted 3600 s. The estimation error and its covariance matrix were chosen as follows

$$\begin{aligned} \boldsymbol{\delta}\hat{q}_{0/0} &= [0.0985 \;\; 0.9853 \;\; -0.0985 \;\; 0.0985]^T \\ \boldsymbol{\delta}\hat{\boldsymbol{\mu}}_{0/0} &= 200[1 \;\; 1 \;\; 1]^T \frac{\text{deg}}{\text{hr}}, \qquad P_{0/0} = 5I_7. \end{aligned} \tag{83}$$

The comparison results between the quaternion KF and the AEKF are summarized in Figs. 5–6. Essentially, these figures emphasize the fact that the AEKF fails to converge due to high initial estimation errors, whereas the quaternion KF converges. Fig. 5

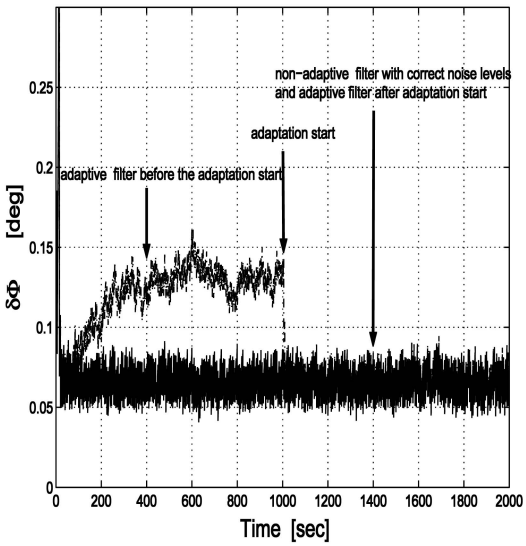


Fig. 7. Case C. Monte-Carlo mean of angular estimation error $\delta\phi$ of adaptive filter over 100 runs.

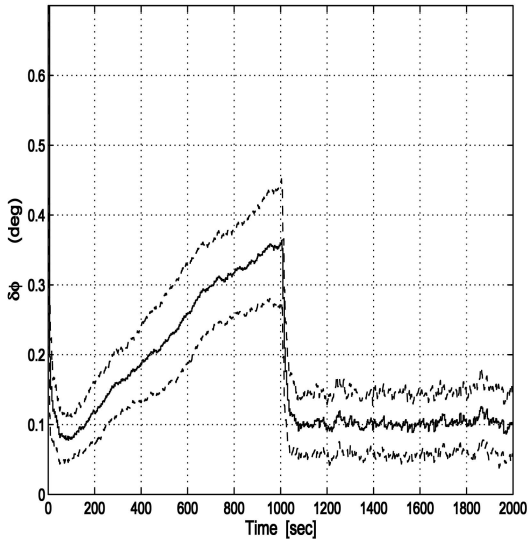


Fig. 8. Case D. Monte-Carlo mean (solid line) and $\pm 1\ \sigma$ envelope (dashed lines) of angular estimation error $\delta\phi$ of adaptive filter over 100 runs.

depicts the Monte-Carlo means of the angular estimation errors, $\delta\phi_{\text{Quaternion KF}}$ and $\delta\phi_{\text{AEKF}}$, of the quaternion KF and AEKF, respectively. It clearly shows that the Monte-Carlo mean of $\delta\phi_{\text{Quaternion KF}}$ converges, whereas that of $\delta\phi_{\text{AEKF}}$ oscillates around the value of 90 deg. The latter stems from the fact that almost all the realizations (97 out of 100) of $\delta\phi_{\text{AEKF}}$ diverge and oscillate between 10 deg and 180 deg. The divergence of the Monte-Carlo means of the gyro drifts estimation errors in the AEKF is presented in Fig. 6. The means reach steady-state values of the order of 10^5 deg/hr. Divergence of all variables in the AEKF occurs just after the first measurement update.

Case C. Adaptive Quaternion KF for Mismodeled Gyro Noise Levels

The noise levels of the system and of the filter are summarized in Table I. Notice that the filter assumes for σ_1 and σ_2 values which are 10 times lower than in the system. The filter was initialized with small estimation errors in order for the effect of the adaptation procedure to surface

$$\boldsymbol{\delta}\hat{\boldsymbol{\mu}}_{0/0} = [5\ \ 5\ \ 5]^T \frac{\text{deg}}{\text{hr}}, \qquad P_{0/0} = 5I_7. \tag{84}$$

Each Monte-Carlo run lasted 2000 s. In each run the adaptive algorithm was activated at time $t = 1000$ s. The results of the adaptive estimation are summarized in Fig. 7. Fig. 7 shows the Monte-Carlo mean of $\delta\phi$ for the adaptive filter $\delta\phi_a$ (dashed line), and the Monte-Carlo mean of $\delta\phi$ for the nonadaptive filter with the correct noise levels $\delta\phi_{na}$ (solid line). Before the adaptation starts, $\delta\phi_a$ settles on approximately 0.13 deg, which is higher than $\delta\phi_{na}$ (approximately 0.07 deg). As soon as the adaptation procedure is activated, $\delta\phi_a$ drops to the level of $\delta\phi_{na}$. This result proves the efficiency of the adaptive filter; it recovers the performance of the ideal case that was lost because of modeling errors.

Case D. Adaptive Quaternion KF for Unmodeled Gyro Drift Rates

The noise standard deviations in the system and in the filter are summarized in Table I. In addition, the system process noise $\mathbf{n}_{3,k}$ had a constant bias $\bar{\mathbf{n}}_3$; that is,

$$\mathbf{n}_{3,k} \sim \mathcal{N}(\bar{\mathbf{n}}_3, \sigma_3^2 \Delta t I_3), \qquad \bar{\mathbf{n}}_3 = [2\ \ 2\ \ 2]^T \frac{\text{deg}}{\text{hr}}. \tag{85}$$

The filter was designed without taking the constant bias $\bar{\mathbf{n}}_3$ into consideration. As is shown next, the adaptation technique copes with that unmodeled bias, too. Each one of the Monte-Carlo simulation runs lasted 2000 s. The adaptive algorithm was activated at $t = 1000$ s in each run. The results are summarized in Fig. 8, which depicts the Monte-Carlo mean and the $\pm 1\sigma$ envelope of the angular estimation error $\delta\phi$. At first, the Monte-Carlo mean of $\delta\phi$ increases almost linearly with time up to a value of approximately 0.35 deg. Then, at 1000 s, when the adaptive filter is activated, it drops to a steady-state value around 0.1 deg. This value is higher than that obtained by a nonadaptive filter without modeling errors (see the solid line in Fig. 7). Nevertheless, the divergence is avoided, and the Monte-Carlo mean and standard deviation of the estimation error remain on an acceptable level. We thus see that in spite of its simple formulation, the proposed adaptive filter can compensate for severe modeling errors such as unknown constant drifts in the RIG outputs.

SUMMARY AND CONCLUSION

A novel KF for estimating the quaternion of rotation using vector measurements was presented in this work. The special feature of the algorithm was its measurement model. This model was obtained by a certain manipulation of the 3×1 vector measurement equation which resulted in an augmented pseudo-measurement equation with a signal term that is linear in the quaternion. Thanks to this favorable property, the usual linearization procedure and, thus, its associated approximation errors, was avoided. The inherent nonlinearity of the quaternion vector measurement was, however, present in the quaternion-dependent noise but with no effect on the filter performance. The process equation (quaternion propagation) was based on gyro measurements. The case of gyro outputs contaminated by random walk drifts and white noises was considered. The gyro drifts were estimated as part of the state vector. The process noise and the measurement noises exhibited the same kind of linear state-dependence. That property was used to develop exact expressions for the respective noise covariance matrices, thus leading to meaningful approximations in the filter implementation. A main benefit of these new expressions was to address in a non ad-hoc way the issue of singularity in the measurement noise covariance matrix. This was done, however, at the expense of the filter computational simplicity. Exploiting the covariance matrix structures, computationally efficient expressions were developed. It was shown, by means of extensive Monte-Carlo simulations, that the filter behaved well under nominal conditions. The proposed filter was tested against the AEKF in the case of high initial estimation errors. The new filter behaved satisfactorily while the AEKF failed to converge. It was suggested to apply adaptive filtering to compensate for modeling errors, which were partly induced by the state-dependence of the noises. A simple procedure for process noise adaptive estimation was developed. As demonstrated through the results of Monte-Carlo simulations, it was successfully applied to the case of a "fine tuning" of the process noise level, and to the harder case of unmodeled constant biases in the gyro outputs. The comparison of the proposed filter with the multiplicative EKF of [9] merits investigation and will be the subject of future work.

APPENDIX A. COVARIANCE MATRIX OF $\mathbf{v}_k = g(\mathbf{x}_k)\mathbf{u}_k$

The proposition formulated in the following is inspired by an approximate result presented in [16, pp. 90–91], where the case of a scalar stochastic sequence with a nonlinear function $G(x)$ was considered. We consider here the case of a vector sequence and provide exact results thanks to the linear-in-x property of $G(\mathbf{x})$.

PROPOSITION *Let $\{\mathbf{x}_k\} \in \mathbb{R}^n$ denote a random sequence with mean $\hat{\mathbf{x}}_k = E\{\mathbf{x}_k\}$ and covariance matrix P_k^x, and let $\{\mathbf{u}_k\} \in \mathbb{R}^m$ denote a zero-mean white random sequence with covariance matrix P_k^u. Assume that $\{\mathbf{x}_k\}$ and $\{\mathbf{u}_k\}$ are statistically independent. Let $\{\mathbf{v}_k\} \in \mathbb{R}^p$ be defined as*

$$\mathbf{v}_k = G(\mathbf{x}_k)\mathbf{u}_k \tag{86}$$

where $G(\cdot) : \mathbb{R}^n \to \mathbb{R}^{p\times m}$ is a linear matrix function of its argument. Then the exact expression for the covariance matrix of $\mathbf{v}_k$, denoted by P_k^v, is given as follows

$$P_k^v = G(\hat{\mathbf{x}}_k)P_k^u G^T(\hat{\mathbf{x}}_k) + \Gamma\,(P_k^u \otimes P_k^x)\Gamma^T \tag{87}$$

where $\otimes$ denotes the Kronecker product, $\Gamma \in \mathbb{R}^{p\times mn}$ is defined as

$$\Gamma \triangleq [G_{c1}\ \ G_{c2}\cdots G_{cm}] \tag{88}$$

and each matrix $G_{ci} \in \mathbb{R}^{p\times n}$, $i = 1,2,\ldots,m$ is defined according to the identity

$$G_{ci}\mathbf{x} = G(\mathbf{x})\mathbf{e}_i. \tag{89}$$

The column-vectors $\mathbf{e}_i$ in (89) are the standard unit vectors in $\mathbb{R}^m$ with 1 at position i and 0 elsewhere; that is, G_{ci} is mapping $\mathbf{x} \in \mathbb{R}^n$ to the ith column of the matrix $G(\mathbf{x}) \in \mathbb{R}^{p\times m}$. Note that G_{ci} can be expressed as a linear mapping of the components of $\mathbf{e}_i$, for $i = 1,2,\ldots,m$.

PROOF See [13, Appendix B].

REMARK 1 The result presented above applies to a large class of sequences. This is the case, in particular, for sequences described by the following equation $y_k = (A + \Delta A)\mathbf{x}_k$, where the matrix ΔA denotes an additive zero-mean white-noise error related to the matrix A. In this case, the error $\Delta A\mathbf{x}_k$ has the form assumed in (86), so that under the assumption of independence between ΔA and $\mathbf{x}_k$, the result of the proposition holds.

REMARK 2 The first term in (87) corresponds to the usual first-order approximation which is made in the framework of extended Kalman filtering. The second term in (87) thus yields an exact expression for the covariance matrix of the modeled sequence.

APPENDIX B. DERIVATION OF (72)

Let A, C, D, with $D \neq 0$ (not all the elements of D are zero) denote three square matrices of the same dimension, and let $B(\eta)$ denote the matrix function of the real scalar η defined by $B(\eta) = C + \eta D$. Consider the following unconstrained minimization problem:

$$\min_{\eta}\{J(\eta) = \|A - B(\eta)\|^2\} \tag{90}$$

where $\|M\|$ denotes the Frobenius norm of the matrix M; that is $\|M\|^2 = \text{tr}(MM^T)$. We show that the solution of this problem, described in (90), is

$$\hat{\eta} = \frac{\text{tr}[(A-C)D^T]}{\text{tr}(DD^T)}. \tag{91}$$

We start by writing

$$\begin{aligned} J(\eta) &= \|A - C - \eta D\|^2 \\ &= \text{tr}[(A-C-\eta D)(A-C-\eta D)^T] \\ &= \eta^2 \text{tr}(DD^T) - \eta \text{tr}[(A-C)D^T + D(A-C)^T] \\ &\quad + \text{tr}[(A-C)(A-C)^T] \\ &= \eta^2 \text{tr}(DD^T) - 2\eta \text{tr}[(A-C)D^T] \\ &\quad + \text{tr}[(A-C)(A-C)^T]. \end{aligned}$$

The first derivative of the cost function $J(\eta)$ is then

$$\frac{dJ}{d\eta} = 2\,\text{tr}(DD^T)\eta - 2\,\text{tr}[(A-C)D^T]. \tag{92}$$

The necessary condition for $\hat{\eta}$ to yield a minimum of J is $dJ/d\eta = 0$. This gives

$$\hat{\eta} = \frac{\text{tr}[(A-C)D^T]}{\text{tr}(DD^T)}. \tag{93}$$

Note that $D \neq 0$ implies that $\text{tr}(DD^T) \neq 0$, and, thus, (93) always has a solution. The second derivative of $J(\eta)$ is

$$\frac{d^2J}{d\eta^2} = 2\,\text{tr}(DD^T). \tag{94}$$

It is well known that for any square matrix D the matrix DD^T is positive semi-definite, therefore its trace is nonnegative. Since the matrix $D \neq 0$ then $DD^T \neq 0$, so that its trace is positive. Therefore, the sufficient condition for $\hat{\eta}$ to identify a minimum of the function $J(\hat{\eta})$ is satisfied.

REFERENCES

[1] Wertz, J. R. (Ed.)
Spacecraft Attitude Determination and Control.
Dordrecht, The Netherlands: D. Reidel, 1984.

[2] Markley, F. L.
Spacecraft attitude determination methods.
Presented at the 40th Israel Annual Conference on Aerospace Sciences, Tel Aviv, Israel, Feb. 23–24, 2000.

[3] Wahba, G.
A least squares estimate of spacecraft attitude.
SIAM Review, **7**, 3 (1965), 409.

[4] Keat, J.
Analysis of least squares attitude determination routine DOAOP.
CSC/TM-77/6034, Computer Sciences Corp., Silver Spring, MD, Feb. 1977.

[5] Bar-Itzhack, I. Y.
REQUEST: A recursive QUEST algorithm for sequential attitude determination.
Journal of Guidance, Control, and Dynamics, **19**, 5 (Sept.–Oct. 1996), 1034–1038.

[6] Psiaki, M. L.
Attitude determination filtering via extended quaternion estimation.
Journal of Guidance, Control, and Dynamics, **23**, 2 (2000), 206–214.

[7] Markley, F. L.
Attitude determination and parameter estimation using vector observations: Theory.
Journal of the Astronautical Sciences, **37**, 1 (Jan.–Mar. 1989), 41–58.

[8] Mortari, D.
A closed-form solution to the Wahba problem.
Journal of the Astronautical Sciences, **45**, 2 (Apr.–June 1997), 195–204.

[9] Lefferts, E. J., Markley, F. L., and Shuster, M. D.
Kalman filtering for spacecraft attitude estimation.
Journal of Guidance, Control, and Dynamics, **5** (Sept.–Oct. 1982), 417–429.

[10] Bar-Itzhack, I. Y., and Oshman, Y.
Attitude determination from vector observations: Quaternion estimation.
IEEE Transactions on Aerospace and Electronic Systems, **AES-21** (Jan. 1985), 128–136.

[11] Bar-Itzhack, I. Y., Deutschmann, J., and Markley, F. L.
Quaternion normalization in additive EKF for spacecraft attitude determination.
Presented at the Guidance, Navigation, and Control Conference, New Orleans, LA, Aug. 1991, AIAA Paper 91-2706.

[12] Wilcox, J. C.
A new algorithm for strapped-down inertial navigation.
IEEE Transactions on Aerospace and Electronic Systems, **AES-3**, 5 (Sept. 1967), 796–802.

[13] Choukroun, D., Bar-Itzhack, I. Y., and Oshman, Y.
A novel quaternion Kalman filter.
TAE Report 930, Faculty of Aerospace Engineering, Technion–Israel Institute of Technology, Haifa, Israel, Jan. 2004,
http://www.aero.technion.ac.il/library/tae_view.php?id=5.

[14] Gelb, A. (Ed.)
Applied Optimal Estimation.
Cambridge, MA: MIT Press, 1979.

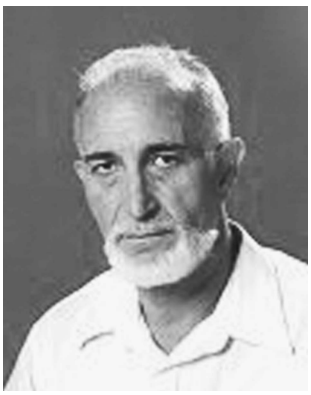
[15] Markley, F. L.
Attitude determination using vector observations and the singular value decomposition.
Journal of the Astronautical Sciences, **36**, 3 (July–Sept. 1988), 245–258.

[16] Jazwinski, A. H.
Stochastic Processes and Filtering Theory.
New York: Academic, 1970.

Daniel Choukroun received the B.Sc. (summa cum laude), M.Sc., and Ph.D. degrees in aerospace engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1997, 2000, and 2003, respectively. He also received the title Engineer under Instruction from the Ecole Nationale de l'Aviation Ciule, Toulouse, France, in 1994.

From 1998 to 2003, he was a teaching and research assistant in the field of automatic control at the Technion–Israel Institute of Technology. Since 2003, he has been a postdoctoral fellow and a lecturer at the University of California, Los Angeles, in the department of mechanical and aerospace engineering. His research interests are in optimal estimation and control theory with applications to aerospace systems.

Dr. Choukroun received the Miriam and Aaron Gutwirth Special Excellency Award for achievement in research from the Technion–Israel Institute of Technology.



Itzhack Y. Bar-Itzhack (M'73—SM'84—F'95) received his B.Sc. and M.Sc. degrees in electrical engineering, in 1961 and 1964, respectively, both from Technion–Israel Institute of Technology. In 1968 he received the Ph.D. degree from the University of Pennsylvania, also in electrical engineering.

From 1968 to 1971 he served as a member of the technical staff at Bellcomm, Inc., in Washington, D.C., where he worked on the Apollo project. In 1971 he joined the Faculty of Aerospace Engineering of The Technion where he is now a Sophie and William Shambam Professor of Aeronautical Engineering, and a member of the Technion Asher Space Research Institute. During the 1977–78 academic year he spent his sabbatical with The Analytic Sciences Corporation (TASC), in Reading, MA, working in research and development of multisensor integrated navigation for the TRIDENT Fleet Ballistic Missile Submarines. During the 1987–89 academic year he was a national research council research associate at NASA–Goddard Space Flight Center, working on attitude determination problems. In 1993 and 1994 he served as Dean of the Faculty of Aerospace Engineering of the Technion. Dr. Bar-Itzhack was back at NASA–Goddard in the 1995–96, the 2000–01 and the 2005–2006 academic years, again as a national research council research associate working on attitude, angular velocity, and orbit determination problems. He has published over 75 journal papers, over 130 conference papers, numerous technical reports, and contributed to two books. He was a member of the AIAA Technical Committee on Guidance Navigation and Control, and was an international advisor of the *Journal of Guidance, Control, and Dynamics*. He is an IEEE Aerospace and Electronic Systems Society Distinguished Lecturer in the distinguished lectures series. He is the recipient of the NASA–Goddard Space Flight Center Group Achievement Award for "Exceptional Achievement in Advanced Attitude Determination and Sensor Calibration Technology" and of the IEEE Aerospace and Electronic Systems Society 2004 Kershner Award in recognition of his "outstanding achievements" in the technology of navigation. (http://www.ewh.ieee.org/soc/aess/plns/content_pages/kershner.html). In 2004 he was also awarded by NASA Administrator the NASA Exceptional Technology Achievement Medal.

His research interests include inertial navigation, spacecraft attitude and orbit determination, applied estimation, and guidance. He is a member of Sigma Xi, and AIAA Fellow.

Yaakov Oshman (AM'96—SM'97) received his B.Sc. (summa cum laude) and D.Sc. degrees, both in aeronautical engineering, from the Technion–Israel Institute of Technology, Haifa, Israel, in 1975 and 1986, respectively.

From 1975 to 1981 he was with the Israeli Air Force, where he worked in the areas of structural dynamics and flutter analysis and flight testing. In 1987 he was a research associate at the Department of Mechanical and Aerospace Engineering of the State University of New York at Buffalo, where he was, in 1988, a visiting professor. Since 1989 he has been with the Department of Aerospace Engineering at the Technion–Israel Institute of Technology, where he is currently a professor. During the 1996/1997 and 1997/1998 academic years he spent a sabbatical with the Guidance, Navigation and Control Center of NASA's Goddard Space Flight Center, where he worked in research and development of spacecraft attitude estimation algorithms. He has consulted to RADA Electronic Industries, Ltd., RAFAEL Ltd., and the Israeli Ministry of Defense. His research interests are in advanced optimal estimation, information fusion, and control methods and their application in aerospace guidance, navigation, and control systems. Of particular interest are interdisciplinary aerospace systems, including structural estimation and control, and health monitoring/fault detection and isolation systems.

Dr. Oshman is the President of the Israeli Association for Automatic Control, a national member organization of the International Federation of Automatic Control (IFAC). He headed the Philadelphia Flight Control Laboratory at the Technion (1993–1996). He published over 120 journal and conference papers and book chapters. He is a coauthor of the paper that was awarded the Best Paper Award of the 2002 AIAA Astrodynamics Specialist Conference, and a coauthor and advisor of the paper that has been awarded the Best Paper Award of the 2004 AIAA Guidance, Navigation and Control Conference. He has received the Technion's Raymond and Miriam Klein Research Prize for his research on enhanced air-to-air missile tracking using target orientation observations (2002), and the Technion's Meir Hanin Research Prize for his work on spacecraft angular velocity estimation (2004). The latter work has been put to use in the Israeli AMOS-2 communication satellite. He has been on the program committees of over a dozen international conferences. He is an editor for Guidance and Control Systems for the *IEEE Transactions on Aerospace and Electronic Systems*, and a member of the editorial board of the AIAA *Journal of Guidance, Control and Dynamics*. He is an Associate Fellow of the AIAA.