Quaternion Estimation from Vector Observations using a Matrix Kalman Filter

D. CHOUKROUN, Member, IEEE
Delft University of Technology

H. WEISS, Member, IEEE
RAFAEL

I. Y. BAR-ITZHACK, Fellow, IEEE
Y. OSHMAN, Fellow, IEEE
Technion–Israel Institute of Technology

A novel two-stage quaternion estimator from vector observations that is a synthesis between Wahba's approach and the Kalman filtering approach is presented. The first stage features an optimal denoising procedure of the elements of a time-varying noisy K-matrix. The second stage produces a quaternion estimate from the filtered K-matrix via any eigenvalue-eigenvector solver. This work's contribution consists in performing the denoising via Kalman filtering. For that purpose, a matrix Kalman filter (MKF) is developed, which has the advantage of preserving the natural formulation of the matrix plant equations. As a result, two aspects of a previous algorithm, called Optimal-REQUEST (OPREQ), are improved: the K-matrix update estimation stage uses a matrix gain rather than a scalar gain, and that gain is optimized with respect to the classical minimum-variance cost. This work assumes that the sensed lines of sight (LOS) are time invariant as seen in the chosen reference frame. This assumption fits in various operational mission architectures. An exact Kalman filter is developed that accounts for the state-multiplicative noise in the process equation. A reduced estimator is also developed assuming simple expressions for the filter covariance matrices. A constrained estimator, which enforces the symmetry and null-trace of the estimated matrix, is designed using the pseudomeasurement (PM) technique. Extensive Monte-Carlo simulations illustrate the performance of the novel filters with a spinning and nutating spacecraft (SC) as a case study. Extensive Monte-Carlo simulations show that the proposed estimator outperforms OPREQ. As illustrated by additional Monte-Carlo simulations, the constrained MKF exhibits a better transient and a better steady-state accuracy than the unconstrained filter for large initial disturbances in the symmetry and null-trace properties.

Manuscript received November 18, 2008; revised October 1, 2009; released for publication November 6, 2011.

IEEE Log No. T-AES/48/4/944196.

Refereeing of this contribution was handled by P. Willett.

This work was presented as Paper 2005-6397 at the AIAA Guidance, Navigation and Control Conference, San Francisco, CA, Aug. 15–18, 2005.

A part of this work was performed while I. Y. Bar-Itzhack held a National Research Council Research Associateship Award at NASA–Goddard Space Flight Center.

Authors' addresses: D. Choukroun, Faculty of Aerospace Engineering, TUD–Delft University of Technology, Delft, 2629 HS, Netherlands, E-mail: (d.choukroun@tud.nl); H. Weiss, RAFAEL, Ministry of Defense, PO Box 2250, Department 35, Haifa 31021, Israel; I. Y. Bar-Itzhack (deceased, 9 may 2007), Asher Space Research Institute, Technion–Israel Institute of Technology, Haifa, 32000, Israel; Y. Oshman, Faculty of Aerospace Engineering, Technion–Israel Institute of Technology, Haifa, 32000, Israel.

INTRODUCTION

The problem of spacecraft (SC) attitude determination (AD) from vector observations has been investigated for the last 40 years, and has given rise to numerous algorithms. A widely used class of these algorithms is concerned with the estimation of the four Euler symmetric parameters [1, pp. 414–416], which form the 4×1 quaternion-of-rotation [1, p. 758–759]. Although a three-axis attitude representation requires a minimum of three parameters, the quaternion $\mathbf{q}$ has become very popular because it is the minimal nonsingular set for global attitude description [1, p. 415], and because the rigid-body kinematics are described by means of a linear differential equation in $\mathbf{q}$.

An optimal estimator of the quaternion typically falls into two categories. The first category has its origin in a constrained least-squares problem introduced by Wahba in 1965 [2]. Wahba's problem was formulated and solved in terms of the quaternion of rotation by Davenport who introduced the celebrated q-method [1, pp. 426–428]. In that method, the optimal quaternion is computed as the eigenvector of a special matrix, the so-called K-matrix, that is associated with the maximal positive eigenvalue. The highlights of the q-method are that it is a closed-form nonlinear optimal estimator of the quaternion, where no a priori estimate is needed, the whole quaternion is estimated rather than some correction, and the unit-norm constraint on the quaternion estimate is explicitly and optimally preserved. Besides these features, other attributes were added to the original q-method along a rich list of AD algorithms: numerical simplicity [3, 4], approximate covariance analysis of the quaternion estimation error [3], estimates of parameters other than the quaternion [5, 6], ability of processing the data recursively [7, 8], and stochastic optimal filtering of time-propagation noises [9]. It is, however, difficult to combine all these enhancements in a single algorithm.

On the other hand, the second category of optimal quaternion estimators, which belongs to the general class of extended Kalman filters, benefits by design from desired properties such as approximate minimum-variance estimation errors, and the straightforward ability of estimating additional states, other than the quaternion, by means of the state augmentation technique [10, p. 350]. The drawbacks of that approach, however, are the well-known linearization effects and the suboptimal procedures that are applied in order to preserve the quaternion's unit-norm constraint [11, 12].

In the present work a novel quaternion optimal estimator is proposed as a synthesis of the approaches mentioned above. Similarly to any $\mathbf{q}$-method-based quaternion estimator, the proposed estimator consists of two stages. The first stage features an optimal denoising procedure of the elements of a time-varying noisy K-matrix. Then, the second stage produces a quaternion estimate from the filtered K-matrix via any eigenvalue-eigenvector solver. This work's contribution resides in a novel and enhanced design of the first stage, i.e., in performing the denoising via Kalman filtering. Therefore, two aspects of the approach introduced in [9] are improved: the K-matrix update estimation stage uses a matrix gain rather than a scalar gain, and that gain is optimized with respect to the classical minimum-variance cost (in [9] the scalar gain is selected nonoptimally). As a result, the first stage computes a more accurate K-matrix. The rational for focusing on enhancing the K-matrix estimation stage is manyfold. First, it is clear that a better knowledge on the K-matrix yields a more accurate extracted quaternion estimate. Second, the K-matrix system equations are linear, which allows a straightforward development of a Kalman filter. Third, the proposed approach circumvents one serious drawback of previous **q**-method-based estimators, which is the difficulty of easily estimating, in a probabilistic framework, other parameters than the quaternion, such as gyro biases. Indeed, using state augmentation techniques [10, p. 350] renders this task straightforward in a Kalman filtering framework. In this work, however, we restrict the scope to quaternion-only estimation. Casting the K-matrix in a state-space framework requires a specific but important type of information pattern; we consider the case where the sensed lines of sight (LOS) remain constant in time along the axes of the chosen reference frame. This assumption is clarified in the next section and several examples of the mission's architectures are provided, showing that the proposed approach can be successfully implemented in practice.

Another original contribution of this work consists of estimating the K-matrix using a matrix Kalman filter (MKF) [13], where the matrix structure of the original K-matrix plant is preserved. Analytical expressions for the covariance matrices of that matrix system's noises are developed. Due to the additive gyro noises, the K-matrix process noise is bilinear with respect to the state and the gyro white noise. Hinging on previous results about Kalman filtering with state-multiplicative noises (e.g. [14–16]), an exact Kalman filter is developed, along with approximate, computationally simpler, versions. Furthermore, the linear constraints of symmetry and zero-trace of the K-matrix are easily incorporated in the MKF paradigm using the pseudomeasurement (PM) approach. That approach, which essentially accounts for substituting soft constraints to hard constraints in the underlying optimization problem, presents conceptual as well as practical advantages. It was successfully applied to quaternion normalization [17] and to direction cosine matrix orthogonalization [18]. Reference [19] presents a comprehensive survey on constrained Kalman filtering via PM (a.k.a pseudo-observations) and projections, it elaborates a successful combination of these approaches and illustrates it in a constrained quaternion estimation problem. Extensive Monte-Carlo simulations illustrate the improvement in the attitude estimation performances of the proposed novel filter when compared with the estimator of [9]. A simple analysis is also provided that validates the quantitative improvement.

The remainder of this paper is organized as follows. The next section is a preliminary section presenting the general linear matrix dynamical model, for which the general MKF is developed. Then the mathematical model for the K-matrix system is developed. The following section contains the development of the MKF of the K-matrix. The issue of constrained estimation is treated afterwards. The comparative numerical study is then presented. Finally, conclusions are drawn in the last section.

THE GENERAL LINEAR MATRIX PLANT

The state MKF [13] can handle linear discrete-time plants that are described by the following matrix equations

$$X_{k+1} = \sum_{r=1}^{\mu} \Theta_k^r X_k \Psi_k^r + W_k \tag{1}$$

$$Y_{k+1} = \sum_{s=1}^{\nu} H_{k+1}^s X_{k+1} G_{k+1}^s + V_{k+1} \tag{2}$$

where $X_k \in \mathbb{R}^{m\times n}$ is the state matrix, $\Theta_k^r \in \mathbb{R}^{m\times m}$, and $\Psi_k^r \in \mathbb{R}^{n\times n}$, $r = 1,2,\ldots,\mu$, are transition matrices, $W_k \in \mathbb{R}^{m\times n}$ is the process noise; the matrix $Y_{k+1} \in \mathbb{R}^{p\times q}$ is the measurement, $H_{k+1}^s \in \mathbb{R}^{p\times m}$, and $G_{k+1}^s \in \mathbb{R}^{n\times q}$, $s = 1,2,\ldots,\nu$, are measurement matrices, and $V_{k+1} \in \mathbb{R}^{p\times q}$ is the measurement noise. The scalars μ and ν are problem dependent. The usual assumptions concerning the noise stochastic models are adopted. That is, the system noises, W_k and V_k, are zero-mean white Gaussian sequences; they are uncorrelated with one another, and uncorrelated with the initial state X_0. Also, the covariances of the noises are known. The covariance of a matrix sequence, say U_k, is defined here as the covariance of its vec-transform, denoted by $\mathrm{vec}(U_k)$, where vec is the vec-operator [23]. The vec-operator operates on an arbitrary matrix, $M \in \mathbb{R}^{m\times n}$, by stacking the columns of M one over the other, and returning the mn-dimensional column-vector, $\mathrm{vec}(M)$. The MKF combines the statistical properties of an ordinary Kalman filter with the advantage of a compact notation. It produces a Kalman filter matrix estimate in terms of the original plant matrices. The algorithm, its proof, and examples of its use are presented in [13]. It is summarized next for convenience.

State Matrix Kalman Filter

The symbol $\otimes$ used thereafter denotes the Kronecker product [24, p. 243].

1) Initialization:

$$\hat{X}_{0/0} = \bar{X}_0, \qquad P_{0/0} = \Pi_0. \tag{3}$$

2) Time Update equations:

$$\hat{X}_{k+1/k} = \sum_{r=1}^{\mu} \Theta_k^r \hat{X}_{k/k} \Psi_k^r \tag{4}$$

$$\mathcal{F}_k = \sum_{r=1}^{\mu} [(\Psi_k^r)^T \otimes \Theta_k^r] \tag{5}$$

$$P_{k+1/k} = \mathcal{F}_k P_{k/k} \mathcal{F}_k^T + Q_k. \tag{6}$$

3) Measurement Update equations:

$$\tilde{Y}_{k+1} = Y_{k+1} - \sum_{s=1}^{\nu} H_{k+1}^s \hat{X}_{k+1/k} G_{k+1}^s \tag{7}$$

$$\mathcal{H}_{k+1} = \sum_{s=1}^{\nu} [(G_{k+1}^s)^T \otimes H_{k+1}^s] \tag{8}$$

$$S_{k+1} = \mathcal{H}_{k+1} P_{k+1/k} \mathcal{H}_{k+1}^T + R_{k+1} \tag{9}$$

$$K_{k+1} = P_{k+1/k} \mathcal{H}_{k+1}^T S_{k+1}^{-1} \tag{10}$$

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{n} \sum_{l=1}^{q} K_{k+1}^{jl} \tilde{Y}_{k+1} E^{lj} \tag{11}$$

where K_{k+1}^{jl} is a $m \times p$ submatrix of the $mn \times pq$ matrix K_{k+1} defined by

$$K_{k+1} = \underbrace{\begin{bmatrix} K_{k+1}^{11} & \cdots & K_{k+1}^{1l} & \cdots \\ \vdots & \ddots & \vdots & \ddots \\ K_{k+1}^{j1} & \cdots & K_{k+1}^{jl} & \cdots \\ \vdots & \ddots & \vdots & \ddots \end{bmatrix}}_{q \text{ matrices}} \left.\vphantom{\begin{bmatrix} K \\ \vdots \\ K \\ \vdots \end{bmatrix}}\right\} n \text{ matrices} \tag{12}$$

and E^{lj} is a $q \times n$ matrix with 1 at position (lj) and 0 elsewhere.

$$\begin{aligned} P_{k+1/k+1} &= (I_{mn} - K_{k+1}\mathcal{H}_{k+1}) P_{k+1/k} (I_{mn} - K_{k+1}\mathcal{H}_{k+1})^T \\ &\quad + K_{k+1} R_{k+1} K_{k+1}^T \end{aligned} \tag{13}$$

where I_{mn} is the $mn \times mn$ identity matrix.

The variance and the covariance associated with $\Delta X[i,j]$ (the element (ij) in the matrix ΔX) are

$$\mathrm{var}\{\Delta X[i,j]\} = P[(j-1)m + i, (j-1)m + i] \tag{14a}$$

$$\mathrm{cov}\{\Delta X[i,j], \Delta X[k,l]\} = P[(j-1)m + i, (l-1)m + k] \tag{14b}$$

where $i,k = 1\ldots m$, and $j,l = 1\ldots n$. The variable ΔX denotes either the a posteriori or the a priori estimation error as applicable, and P is the associated covariance matrix.

THE MATHEMATICAL MODEL

In this section the state-space model equations of the K-matrix system are formulated, and explicit expressions for the system noise covariance matrices are provided.

Preliminaries

The State K-Matrix: We consider: 1) two Cartesian coordinate frames, $\mathcal{R}$ and $\mathcal{B}$, which are the reference frame and the SC body frame, respectively, 2) batches of $N(k)$; $k = 1,2\ldots$ physical vectors $\{\bar{\mathbf{v}}_i(k)\}_{i=1}^{N(k)}$, e.g. LOS vectors, which are observed at each sampling time t_k, 3) two sets of projections for each LOS vector $\bar{\mathbf{v}}_i$, onto the frames $\mathcal{R}$ and $\mathcal{B}$, denoted as $\{\mathbf{r}_i(k)\}_{i=1}^{N(k)}$ and $\{\mathbf{b}_i(k)\}_{i=1}^{N(k)}$, respectively, 4) the matrices $K(k)$, which are known functions of the sets $\{\mathbf{r}_i(k), \mathbf{b}_i(k)\}_{i=1}^{N(k)}$, and are defined as follows:

$$K(k) = \begin{bmatrix} S - \sigma I_3 & \mathbf{z} \\ \mathbf{z}^T & \sigma \end{bmatrix} \tag{15}$$

where

$$S = B + B^T \tag{16}$$

$$B = \sum_{i=1}^{N(k)} a_i \mathbf{b}_i \mathbf{r}_i^T \tag{17}$$

$$\mathbf{z} = \sum_{i=1}^{N(k)} a_i \mathbf{b}_i \times \mathbf{r}_i^T \tag{18}$$

$$\sigma = \mathrm{tr}(B) \tag{19}$$

and a_i are positive scalar weights. The cornerstone of the proposed approach is that the elements of the matrix $K(k)$ define a set of state variables, which we aim at estimating using a sequence of noisy measurements of the $\mathbf{b}_i$s. This idea of working with an (albeit redundant) representation of the attitude implies that it should not take two different numerical values when describing the same attitude. From this premise stems the need to use the same batch of LOS vectors in order to define the (ideal) K-matrix. In other words, every sampled LOS vector should have its projection on the reference frame remain identical, i.e., the vectors $\mathbf{r}_i(k)$ should remain identical at each t_k. We now clarify and expand on this assumption.

1) Assume, for simplicity (the assumption will be relaxed later on), that $\mathcal{B}$ and $\mathcal{R}$ coincide at each t_k and notice that both frames can themselves be rotating with respect to an inertial frame. The relative attitude of $\mathcal{B}$ with respect to $\mathcal{R}$ is thus constantly zero. Assume that, at t_0, two LOS vectors $\bar{\mathbf{v}}_1, \bar{\mathbf{v}}_2$ are available and that they coincide with the axes $\mathcal{R}_x$ and

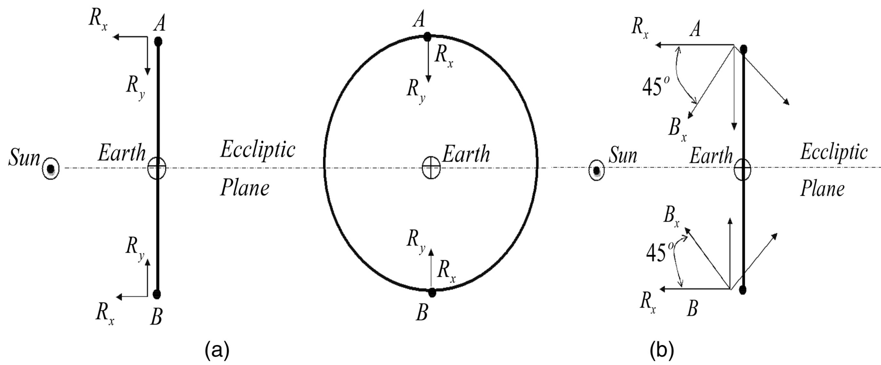


Fig. 1. Sun-pointing/nadir-pointing LEO satellite. Attitude with respect to trajectory frame. (a) Zero attitude. (b) 45° attitude.

$\mathcal{R}_y$. Clearly, we have

$$\mathbf{b}_1 = \mathbf{r}_1 = [1,0,0]^T, \qquad \mathbf{b}_2 = \mathbf{r}_2 = [0,1,0]^T.$$

From (15), the K-matrix that is computed using these LOS vectors is then (with unit weighting coefficients):

$$K(t_0) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -2 & \cdot \\ \cdot & \cdot & \cdot & 2 \end{bmatrix} \tag{20}$$

where the dots in (20) represent zeros. Equation (20) shows the true values of the 16 state variables at t_0. Next, assume that at t_1, the second LOS vector is not available and that a third LOS vector $\bar{\mathbf{v}}_3$ can be acquired, which lies along the axis $\mathcal{R}_z$, such that

$$\mathbf{b}_3 = \mathbf{r}_3 = [0,0,1]^T.$$

Then, using $(\mathbf{r}_1, \mathbf{r}_3)$ in (15) yields the following K-matrix:

$$K(t_1) = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & -2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 2 \end{bmatrix}. \tag{21}$$

Equation (21) shows the values of the state variables at t_1. Clearly, (20) and (21) are inconsistent: although the attitude has not changed from t_0 to t_1, the state variables assume different numerical values, which depend on the values of the $\mathbf{r}$s used. Thus, choosing the $\mathbf{r}$s to be identical at t_0 and t_1 allows for the K-matrix to consistently define a representation of the attitude of $\mathcal{B}$ with respect to $\mathcal{R}$.

2) Although the $\mathbf{r}$s are required to be constant, the observed LOS vectors are not constrained to be the same physical directions at all times. This is illustrated by the following example. Consider the case of a Sun-pointing and Earth-pointing (nadir) low Earth orbit (LEO) satellite. The reference frame is defined as the trajectory frame, i.e., with $\mathcal{R}_x$ pointing normal to the orbit plane, $\mathcal{R}_y$ coinciding with the local nadir, and $\mathcal{R}_z$ pointing forward along the in-track orbit tangent. These LEO satellites orbits are usually Sun-synchronous, with inclinations close or equal to 90°, which avoids Sun eclipses. See the illustration in Fig. 1 for a 90° inclination. Due to the very high ratio between the Sun–SC and the Earth–SC distances (around 10^4), the axis $\mathcal{R}_x$ can be identified with the Sun–SC LOS vector, for all practical purposes. This first LOS vector can be observed via Sun sensors. The nadir (the $\mathcal{R}_y$ axis) provides a second LOS vector and can be observed by Earth sensors. Thus, thanks to this choice of the reference frame, we are in the case discussed in 2 above, where:

$$\mathbf{r}_1 = [1,0,0]^T, \qquad \mathbf{r}_2 = [0,1,0]^T.$$

In the case of a zero attitude, the body frame coincides with the reference frame (see Fig. 1(a)), and the time-invariant state matrix is thus obtained from (20), i.e.,

$$X = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -2 & \cdot \\ \cdot & \cdot & \cdot & 2 \end{bmatrix}. \tag{22}$$

The above example illustrates an operational configuration where the two observed LOS vectors keep the same projections in the reference frame at all sampling times. Nonetheless, the nadir-LOS vector is continuously changing with respect to the inertial frame.

3) We now show how to relax the requirement for the body frame to be fixed with respect to the reference frame. Assume that $\mathcal{B}(k)$ rotates with respect to the reference frame $\mathcal{R}$ with an angular velocity $\boldsymbol{\omega}_k$ and that $\boldsymbol{\omega}_k$ is known along the axes of $\mathcal{B}(k)$. For the sake of illustration, consider a roll-only motion, i.e., $\mathcal{B}$ rotates around its axis $\mathcal{B}_z$, the in-track orbit tangential axis, with a magnitude of $\pi/4$ rad/s, such that:

$$\boldsymbol{\omega}_k = [0,0,\pi/4]^T. \tag{23}$$

This may arise from a requirement of scanning or tracking of a body-mounted camera. Assume that at t_0, $\mathcal{B}$ and $\mathcal{R}$ coincide, such that, according to the previous computations, $X(t_0)$ is obtained from (22). Using $\boldsymbol{\omega}_k$, the matrix X can be propagated from t_0 to

$t_1 = t_0 + 1$ [8]. Considering the present example, the computations are

$$X(t_1) = \Phi X(t_o)\Phi^T$$
$$= \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -\sqrt{2} & \sqrt{2} \\ \cdot & \cdot & \sqrt{2} & \sqrt{2} \end{bmatrix} \tag{24}$$

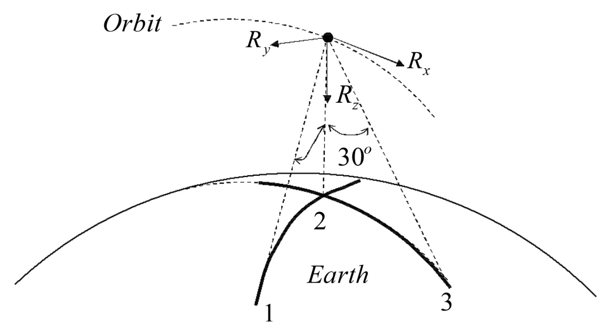
where

$$\Phi = \exp\left\{ \frac{1}{2} \begin{bmatrix} \cdot & \frac{\pi}{4} & \cdot & \cdot \\ -\frac{\pi}{4} & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \frac{\pi}{4} \\ \cdot & \cdot & -\frac{\pi}{4} & \cdot \end{bmatrix} \right\}. \tag{25}$$

Equation (24) provides the true value of the state matrix that represents the attitude of $\mathcal{B}(t_1)$ with respect to $\mathcal{R}$. This can be visualized in Fig. 1(b). The dynamical model where $\boldsymbol{\omega}_k$ is measured with noises is described in the next subsection. To conclude, the proposed approach can handle time-varying attitudes, provided that a measurement of the angular velocity of $\mathcal{B}$ with respect to $\mathcal{R}$ is available.[1]

Mission Architectures: Following are various and important examples of real missions architectures in which the proposed approach can be successfully implemented.

1) Consider a LEO SC whose attitude is simultaneously stabilized with respect to the Sun–SC LOS vector (for power production via the solar arrays) and to the SC–nadir LOS vector (for Earth observation). We have already used this architecture for previous illustration (see Fig. 1). The knowledge of the relative position of the SC with respect to the Sun/Earth in some reference frame (from navigation computations) and in the body frame (from on-board measurements) allows for AD and this fits the framework proposed in this work. In the NASA LANDSAT [20] mission, the Landsat 7 satellite is designed for a Sun-synchronous, Earth mapping orbit. Its payload is a single nadir-pointing instrument and power is provided by a single Sun-tracking solar array. The satellite attitude control must maintain the SC platform within 0.015 deg of Earth pointing. In the NASA SAMPEX [21] mission, SAMPEX is a momentum-biased, Sun-pointing SC that maintains the experiment-view axis in a zenith direction as much as possible. It points its solar array at the Sun by aiming the momentum vector toward the Sun and rotating the SC one revolution per orbit about the Sun–SC axis. A two-axis digital Sun sensor and a set of five



$$\mathbf{r}_1 = \begin{bmatrix} 0 \\ 1/2 \\ \sqrt{3}/2 \end{bmatrix} \quad \mathbf{r}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{r}_3 = \begin{bmatrix} 1/2 \\ 0 \\ \sqrt{3}/2 \end{bmatrix}$$

Fig. 2. LEO satellite. Attitude with respect to trajectory frame.

coarse Sun sensors are used for AD, which provides the quasi antinadir pointed attitude required by the science experience.

Consider also another architecture of LEO SC mission, e.g., a SC monitoring the Earth surface via some imaging sensor and equipped with an Earth sensor. Figure 2 illustrates a configuration of three observed LOS vectors, which are the nadir (aligned with the $\mathcal{R}_z$-axis), and two additional LOS vectors inclined by 30° with respect to the nadir in the in-track plane and in the off-track plane, respectively. The corresponding values of the **r**s that appear in Fig. 2 are given in the trajectory frame $(\mathcal{R}_x, \mathcal{R}_y, \mathcal{R}_z)$. Therefore, the proposed approach can be successfully applied provided that the Earth imaging sensor is continuously observing these three LOS vectors.

2) Consider the case of a geostationary SC. Once stabilized along the nadir, the SC is able to observe the same locations on the Earth surface all the time. Not only two of them but as many as the field of view allows for. On a geosynchronous orbit (approximately 36000 km), the SC sees the Earth with an angle of about 22 deg (assuming an Earth radius of 6400 km). Within that field of view, any digital camera can track many locations on Earth. The NASA GOES mission features a range of geostationary satellites. Figure 3 visualizes the configuration of three sensed LOS vectors, namely, the nadir, and two additional LOS vectors inclined by 10° with respect to the nadir in the in-track plane. The values of the LOS vectors' projections in the reference frame are given in Fig. 3 for this specific configuration.

3) Consider an SC whose attitude is inertially stabilized with respect to some celestial coordinate frame. Although rotating about some equilibrium position, the SC is assumed to be stabilized enough such that the same physical directions can be observed at any time. These can be directions to stars and in that case their inertial projections are close to be time-invariant. Notice that 3-D stabilization is not needed but that even spin stabilized SC are adequate, provided that their sensors are oriented along the

[1]Notice that body-mounted gyros provide the inertial angular velocity of $\mathcal{B}$. However, knowing the inertial angular velocity of $\mathcal{R}$ allows us to compute the sought angular velocity $\boldsymbol{\omega}_k$.

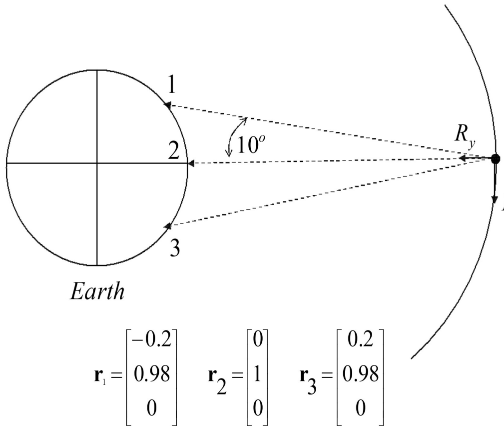


Fig. 3. Geostationary satellite. Attitude with respect to trajectory frame.

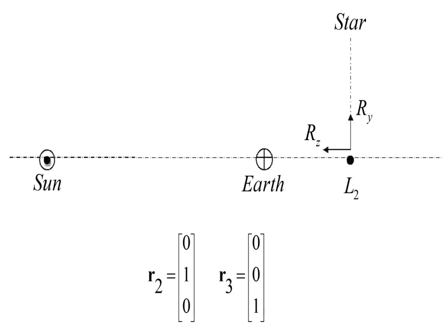


Fig. 4. WMAP-Wilkison maximum anisotropy probe at L_2 Lagrangian point. Attitude with respect to Sun rotating frame.

directions of the inertially stabilized axes. In the NASA MAP [22] mission, the MAP SC is stabilized about the $L2$ Lagrange point, scanning the celestial sphere around the anti Sun–nadir direction. MAP spins every 2 min and its spin axis maintains a fixed angle of 22.5 deg to the Sun–Earth line. The spin axis moves around the Sun–Earth line. The SC uses Sun sensors, star trackers, and gyroscopes for AD. Figure 4 illustrates the case for two sensed LOS vectors, e.g., the Sun–SC LOS vector and the direction to a star perpendicular to it, and provides the values of the $\mathbf{r}$s.

Process Equation

As shown in [9] the dynamics of the true K-matrix can be modeled by the following first-order stochastic linear matrix equation

$$X_{k+1} = \Phi_k X_k \Phi_k^T + W_k \tag{26}$$

where X_k denotes the ideal noise-free K-matrix at time t_k and Φ_k is computed using the angular velocity vector $\boldsymbol{\omega}_k$ (as measured by a triad of body-mounted gyroscopes during a small time increment Δt), i.e.,

$$\Omega_k = \frac{1}{2}\begin{bmatrix} -[\boldsymbol{\omega}_k\times] & \boldsymbol{\omega}_k \\ -\boldsymbol{\omega}_k^T & 0 \end{bmatrix} \tag{27}$$

$$\Phi_k = \exp\{\Omega_k \Delta t\} \tag{28}$$

where, in (27), $[\boldsymbol{\omega}_k\times]$ denotes the cross-product matrix which is defined by the identity $\boldsymbol{\omega}_k \times \mathbf{u} = [\boldsymbol{\omega}_k\times]\mathbf{u}$ for any vector $\mathbf{u} \in \mathbb{R}^3$. One can see that (26) is a special case of the general process equation described in (1), and is, therefore, implementable in an MKF. As shown in Appendix I, the matrix W_k can be expressed as follows

$$W_k = (X_k\mathcal{E}_k - \mathcal{E}_k X_k)\Delta t \tag{29}$$

where X_k is the state matrix at t_k, the time increment is denoted by Δt, and $\mathcal{E}_k$ is the following 4×4 skew-symmetric matrix

$$\mathcal{E}_k = \frac{1}{2}\begin{bmatrix} -[\boldsymbol{\epsilon}_k\times] & \boldsymbol{\epsilon}_k \\ -\boldsymbol{\epsilon}_k^T & 0 \end{bmatrix} \tag{30}$$

The 3×1 vector $\boldsymbol{\epsilon}_k$ denotes the additive error in the measured value of the body angular velocity. In this work we consider the special case where $\boldsymbol{\epsilon}_k$ is a zero-mean white noise process with covariance matrix $Q_k^{\epsilon}/\Delta t$. Moreover, it is assumed that $\boldsymbol{\epsilon}_k$ is uncorrelated with the initial state, X_0. Using these assumptions it can easily be shown that the matrix noise W_k is a zero-mean process.

State-Dependent Process Noise: Taking advantage of the bilinear structure of W_k with respect to $\boldsymbol{\epsilon}_k$ and to X_k, we seek an analytic expression for its covariance matrix Q_k. Previous works on Kalman filtering with this type of noise state-dependence can be found in [14]–[16].

PROPOSITION 1 *Let* $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *denote the three columns of the identity matrix* I_3 *in* $\mathbb{R}^3$. *Let* M *and* Γ_k *denote the following* 16×3 *and* 16×16 *matrices:*

$$\begin{aligned} M^T = [-[\mathbf{e}_1\times] \quad -\mathbf{e}_1 \quad -[\mathbf{e}_2\times] \quad -\mathbf{e}_2 \\ -[\mathbf{e}_3\times] \quad -\mathbf{e}_3 \quad I_3 \quad \mathbf{0}] \end{aligned} \tag{31}$$

$$\Gamma_k = \tfrac{1}{2}[(I_4 \otimes X_k) - (X_k^T \otimes I_4)]M. \tag{32}$$

Then,

i) The 16×16 *covariance matrix of* W_k, *denoted by* Q_k, *satisfies:*

$$Q_k = E\{\Gamma_k \boldsymbol{\epsilon}_k \boldsymbol{\epsilon}_k^T \Gamma_k^T\}\Delta t^2. \tag{33}$$

ii) If, furthermore, the components of $\boldsymbol{\epsilon}_k$ *are independent and identically distributed with covariance parameter* $\sigma^2/\Delta t$, *and* N_k *denotes the second-order moment of the state* X_k, *then* Q_k *is computed as follows:*

$$Q_k = \sum_{i=1}^{3} \Upsilon_i N_k \Upsilon_i^T \sigma^2 \Delta t \tag{34}$$

where for $i = 1,2,3$

$$\Upsilon_i = \tfrac{1}{2}(A_i \oplus A_i) \tag{35}$$

with $\oplus$ denoting the Kronecker sum [24] *and*

$$A_1 = \begin{bmatrix} \mathbf{0} & -\mathbf{e}_3 & \mathbf{e}_2 & \mathbf{e}_1 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$
$$A_2 = \begin{bmatrix} \mathbf{e}_3 & -\mathbf{0} & -\mathbf{e}_1 & \mathbf{e}_2 \\ 0 & -1 & 0 & 0 \end{bmatrix} \tag{36}$$
$$A_3 = \begin{bmatrix} -\mathbf{e}_2 & \mathbf{e}_1 & \mathbf{0} & \mathbf{e}_3 \\ 0 & 0 & -1 & 0 \end{bmatrix}.$$

The proof is provided in Appendix II. The usefulness and implementation of both results from Proposition 1 is discussed in the next section.

Measurement Equation

The measurement model equation is the following simple matrix equation [9]

$$Y_{k+1} = X_{k+1} + V_{k+1} \tag{37}$$

where Y_{k+1} is the measured K-matrix constructed using the noisy vector observations acquired at time t_{k+1}, and V_{k+1} denotes the measurement noise. Looking at the general measurement model described in (2) one realizes that (37) is a particular case of (2) where $H^s_{k+1} = G^s_{k+1} = I_4$ and $\nu = 1$. Thus (37) is readily implementable in the MKF.

Given a batch of vector observations, $\mathbf{b}_i$, $\mathbf{r}_i$, $i = 1,2,\ldots,m$ ($m > 1$), acquired at t_{k+1}, the expression for the noise matrix V_{k+1} in the measurement equation (37) is

$$V_{k+1} = \sum_{i=1}^{m} \alpha_i V^i_{k+1} \tag{38}$$

where $\alpha_i \triangleq a_i / \sum_{i=1}^{M} a_i$, and the a_is are positive weighting scalars associated with each vector measurement. Each matrix V^i_{k+1} in (38) is expressed using a pair, $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1}), \mathbf{r}_i(t_{k+1})$ as follows,

$$V_{k+1} = \begin{bmatrix} \mathcal{S}_b - \kappa_b I_3 & \mathbf{z}_b \\ \mathbf{z}_b^T & \kappa_b \end{bmatrix} \tag{39}$$

where

$$\begin{aligned} B_b &= \boldsymbol{\delta}\mathbf{b}_{k+1}\mathbf{r}^T_{k+1}, & \mathcal{S}_b &= B_b + B_b^T \\ \mathbf{z}_b &= \boldsymbol{\delta}\mathbf{b}_{k+1} \times \mathbf{r}_{k+1}, & \kappa_b &= \mathrm{tr}(B_b) \end{aligned} \tag{40}$$

and where the 3×1 vector $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$ is the error in the ith vector observation. It is assumed that the sequence $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$, $k = 1,\ldots,m$, is a zero-mean, white sequence with a known covariance matrix, $R^{\mathbf{b}_i}_{k+1}$. Moreover, the vector observations acquired at the same time are assumed to be uncorrelated with one another. That is,

$$E\{\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})\} = \mathbf{0} \tag{41a}$$
$$E\{\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})\boldsymbol{\delta}\mathbf{b}_i(t_l)^T\} = R^{\mathbf{b}_i}_{k+1}\delta_{k+1,l} \tag{41b}$$
$$E\{\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})\boldsymbol{\delta}\mathbf{b}_j(t_{k+1})^T\} = R^{\mathbf{b}_i}_{k+1}\delta_{ij} \tag{41c}$$

where $i,j = 1,2,\ldots m$, $k,l = 1,2\ldots$ and $\delta_{k+1,l}, \delta_{ij}$ denote Kronecker deltas. In addition, it is assumed that the vector observations are uncorrelated with the process noise $\boldsymbol{\epsilon}_k$ and with the initial state X_0. From (39) and (40) it appears that the elements of V^i_{k+1} are linear combinations of the elements of $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$. Since $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$ is zero-mean, then the sequence V^i_{k+1} is also zero-mean. Since V_{k+1} is a weighted sum of zero-mean sequences V^i_{k+1} (see (38)), then V_{k+1} is also zero-mean. Using (41) an analytic expression for the covariance matrix of V_{k+1}, denoted by R_{k+1}, can be derived. A summary of the computation of the 16×16 matrix R_{k+1} is given next. Its proof is lengthy but straightforward and is provided in Appendix III. For the sake of clarity the symbol t_{k+1} is dropped from the following equations; however, it should be remembered that all the computations carry the time tag t_{k+1}. Given $\mathbf{r}_i$, $R^{\mathbf{b}_i}$, and a_i, $i = 1,2,\ldots m$, compute

$$\tilde{R}_i = \begin{bmatrix} [\mathbf{r}_i \times] & \mathbf{r}_i \\ -\mathbf{r}_i^T & 0 \end{bmatrix} \tag{42a}$$
$$\Lambda_i = (\tilde{R}_i \otimes I_4)M \tag{42b}$$
$$R_i = \Lambda_i R^{\mathbf{b}_i} \Lambda_i^T \tag{42c}$$
$$\alpha_i = \frac{a_i}{\sum_{i=1}^{m} a_i} \tag{42d}$$
$$R_{k+1} = \sum_{i=1}^{m} \alpha_i^2 R_i \tag{42e}$$

where the matrix M is defined in (31).

FILTER IMPLEMENTATION

State-Dependent Process Noise

1) As mentioned earlier, and as seen from (29), the process noise W_k is a function of the state X_k. Fortunately, this dependence is linear which allows an exact computation of the covariance of W_k provided that the second-order moment of X_k, N_k, is known. It is well known that N_k can be propagated via a Lyapunov difference equation. The computation of Q_k is thus done as follows:

$$Q_k = \sum_{i=1}^{3} \Upsilon_i N_k \Upsilon_i^T \sigma^2 \Delta t \tag{43a}$$
$$\mathcal{F}_k = \Phi_k \otimes \Phi_k \tag{43b}$$
$$N_{k+1} = \mathcal{F}_k N_k \mathcal{F}_k^T + Q_k \tag{43c}$$

where the matrices Υ_i are defined in (35) and (36). The above equations are implemented in the time-propagation stage of the Kalman filter.

2) While the above algorithm for computing Q_k has the merit of being exact, it implies an additional computational burden in the filter, which should be weighed against the performance improvements. On the other hand, the practitioner

may consider implementing the following approximate computations. In order to avoid computing the matrix N_k, Q_k can be approximated by simply replacing the state X_k by its best available estimate, $\hat{X}_{k/k}$, in (32), (33). This step is similar to what is done in an extended Kalman filter (EKF). Then, assuming independent and identically distributed components in ϵ_k, a first-order approximation in Δt for the covariance matrix of W_k yields

$$Q_k \simeq \tfrac{1}{4}[(I_4 \otimes \hat{X}_{k/k}) - (\hat{X}_{k/k}^T \otimes I_4)]MM^T \times [(I_4 \otimes \hat{X}_{k/k}) - (\hat{X}_{k/k}^T \otimes I_4)]^T \sigma^2 \Delta t.$$

3) A third option for approximating Q_k is

$$Q_k = \bar{Q}_k \otimes I_4. \tag{44}$$

Under additional conditions, (44) provides a drastic reduction in the covariance computation while maintaining satisfactory performance. This case is discussed in the subsection on reduced covariance.

Singularity of the Measurement Noise Covariance Matrix

The covariance matrices R_i as computed in (42c) are singular. Indeed, assuming that the covariance matrix in each vector-measurement error $R^{\mathbf{b}_i}$ is of full rank (rank three), then every 16×16 matrix R_i is at most of rank three, and is, thus, singular. Therefore, if there is a single vector measurement ($m = 1$) at t_{k+1} the rank of R_{k+1} is three and R_{k+1} is singular. Simulations showed that if two noncollinear vector observations are acquired at t_{k+1}, the rank of R_{k+1} increases to six, and, in the case of three or more than three noncollinear measurements ($m \geq 3$), the rank equals nine. This is consistent with the properties of the error matrix V_{k+1} (see (37)). Since Y_{k+1} and X_{k+1} are symmetric matrices with a trace equal to zero, then V_{k+1} has necessarily the same properties. These properties introduce seven linear constraints among the elements of V_{k+1}; namely, six constraints for the symmetry and one for the trace, which lowers the rank of V_{k+1} from 16 down to 9. There are several techniques to cope with the issue of a singular measurement covariance matrix (see e.g. [10, p. 354]). To circumvent the problem here, we add small values to the main diagonal of R_{k+1}, which has a stabilizing effect on the numerics of the Kalman filter. This is done by choosing a small β relatively to the assumed level of the noise, and computing R_{k+1} as follows

$$R_{k+1} = \sum_{i=1}^{m} \alpha_i^2 R_i + \beta I_{16} \tag{45}$$

where I_{16} denotes the 16×16 identity matrix.

Algorithm Summary

The MKF of the K-matrix is summarized in the following.

1) Initialization:

$$\hat{X}_{0/0} = Y_0, \qquad P_{0/0} = R_0. \tag{46}$$

2) Time Update:

$$\hat{X}_{k+1/k} = \Phi_k \hat{X}_{k/k} \Phi_k^T \tag{47}$$

$$\mathcal{F}_k = \Phi_k \otimes \Phi_k \tag{48}$$

$$P_{k+1/k} = \mathcal{F}_k P_{k/k} \mathcal{F}_k^T + Q_k. \tag{49}$$

3) Measurement Update:

$$\tilde{Y}_{k+1} = Y_{k+1} - \hat{X}_{k+1/k} \tag{50}$$

$$S_{k+1} = P_{k+1/k} + R_{k+1} \tag{51}$$

$$K_{k+1} = P_{k+1/k} S_{k+1}^{-1} \tag{52}$$

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{4}\sum_{l=1}^{4} K_{k+1}^{jl} \tilde{Y}_{k+1} E^{lj} \tag{53}$$

where K_{k+1}^{jl} is a 4×4 submatrix of the 16×16 matrix K_{k+1} defined by

$$K_{k+1} = \underbrace{\begin{bmatrix} K_{k+1}^{11} & \cdots & K_{k+1}^{14} \\ \vdots & \ddots & \vdots \\ K_{k+1}^{41} & \cdots & K_{k+1}^{44} \end{bmatrix}}_{4 \text{ submatrices}} \left.\vphantom{\begin{bmatrix} K \\ \vdots \\ K \end{bmatrix}}\right\} 4 \text{ submatrices} \tag{54}$$

$$P_{k+1/k+1} = (I_{16} - K_{k+1}) P_{k+1/k} (I_{16} - K_{k+1})^T + K_{k+1} R_{k+1} K_{k+1}^T. \tag{55}$$

The MKF that is described in (46)–(55) produces a sequence of K-matrix estimates $\hat{X}_{k/k}$. This constitutes the first stage of the attitude estimation process. The second stage consists of the computation of the eigenvector of $\hat{X}_{k/k}$ that belongs to the largest eigenvalue.

Reduced Covariance Filter

In this section we show how to reduce the computational complexity of the filter by making further reasonable assumptions on the statistics of the system's noises. Namely, we assume that the rows of W_k are independently identically distributed with 4×4 covariance matrix $\bar{Q}_k$. Similar assumptions are made with respect to the matrix V_k and to the initial estimation error matrix $\Delta X_{0/0}$. These assumptions are expressed as

$$Q_k = \bar{Q}_k \otimes I_4 \tag{56a}$$

$$R_k = \bar{R}_k \otimes I_4 \tag{56b}$$

$$P_{0/0} = \bar{P}_{0/0} \otimes I_4. \tag{56c}$$

As shown in Appendix IV, the use of (56) in the equations of the filter ((46)–(55)) reduces the covariance computation from 16×16 to 4×4 matrix equations. Moreover, it enables a more compact notation in the state measurement update stage. The reduced covariance filter is summarized next.

1) Initialization:

$$\hat{X}_{0/0} = Y_0, \qquad \bar{P}_{0/0} = \bar{R}_0. \tag{57}$$

2) Time Update:

$$\hat{X}_{k+1/k} = \Phi_k \hat{X}_{k/k} \Phi_k^T \tag{58}$$

$$\bar{P}_{k+1/k} = \Phi_k \bar{P}_{k/k} \Phi_k^T + \bar{Q}_k \tag{59}$$

where $\bar{P}_{k/k}$ and $\bar{P}_{k+1/k}$ are 4×4 matrices.

3) Measurement Update:

$$\tilde{Y}_{k+1} = Y_{k+1} - \hat{X}_{k+1/k} \tag{60}$$

$$\bar{S}_{k+1} = \bar{P}_{k+1/k} + \bar{R}_{k+1} \tag{61}$$

$$\bar{K}_{k+1} = \bar{P}_{k+1/k} (\bar{S}_{k+1})^{-1} \tag{62}$$

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \tilde{Y}_{k+1} \bar{K}_{k+1}^T \tag{63}$$

$$\begin{aligned} \bar{P}_{k+1/k+1} &= (I_4 - \bar{K}_{k+1}) \bar{P}_{k+1/k} (I_4 - \bar{K}_{k+1})^T \\ &\quad + \bar{K}_{k+1} \bar{R}_{k+1} \bar{K}_{k+1}^T \end{aligned} \tag{64}$$

where $\bar{P}_{k+1/k+1}$, $\bar{S}_{k+1}$, and $\bar{K}_{k+1}$ are 4×4 matrices.

As a result of assumptions (56) the filter covariance computations are identical for each 1×4 row of the estimation error matrix. This is why a single 4×4 covariance computation is needed. If the full covariance matrices are needed they are readily computed as shown in Appendix IV by

$$P_{k/k} = \bar{P}_{k/k} \otimes I_4 \tag{65a}$$

$$P_{k+1/k} = \bar{P}_{k+1/k} \otimes I_4 \tag{65b}$$

$$S_{k+1} = \bar{S}_{k+1} \otimes I_4 \tag{65c}$$

$$\mathcal{K}_{k+1} = \bar{K}_{k+1} \otimes I_4. \tag{65d}$$

Covariance Matrix of the Quaternion Estimation Error

In this section we use and extend a previous result, as presented in [3] in order to evaluate the covariance matrix of the quaternion estimation error. It is shown how this matrix can be extracted from the covariance matrix $P_{k/k}$, as computed in (55).

Let $\boldsymbol{\delta z}$ denote the measurement error in the 3×1 vector $\mathbf{z}$ of a measured K-matrix, which is constructed from m vector measurements and let P_{zz} denote the covariance matrix of $\boldsymbol{\delta z}$. Let $\boldsymbol{\delta q}$ denote the quaternion multiplicative estimation error; that is

$$\boldsymbol{\delta q} = \hat{\mathbf{q}} \star \mathbf{q}^{-1} \tag{66}$$

where $\mathbf{q}$ is the true quaternion, $\hat{\mathbf{q}}$ is the estimate, $\star$ and $\hat{\mathbf{q}}^{-1}$ denote the operations of quaternion product and quaternion inverse, respectively [1, p. 758]. Thus, $\boldsymbol{\delta q}$ itself is a quaternion; it has a vector part $\boldsymbol{\delta e}$ and a scalar part δq. This error-quaternion is related to the rotation that brings the true body frame onto the estimated body frame. Assuming that the angle of this rotation is small, δq is approximated by 1, as done in [3]. The uncertainty is thus concentrated in $\boldsymbol{\delta e}$. It can be shown [3] that a good approximation P_{ee} to the covariance matrix of $\boldsymbol{\delta e}$ is computed as

$$P_{ee} = N P_{zz} N \tag{67}$$

where

$$N = \left\{ 2 \sum_{i=1}^{m} \alpha_i (I_3 - \mathbf{r}_i \mathbf{r}_i^T) \right\}^{-1} \tag{68}$$

and $\mathbf{r}_i$, $i = 1, 2, \ldots, m$, is the batch of observed LOS vectors, as resolved in the reference frame, which are acquired at a particular epoch time. In order to use (67) and (68) in our algorithm, consider $\boldsymbol{\delta z}$ as an estimation error rather than a measurement error. The 3×3 covariance matrix of $\boldsymbol{\delta z}$, P_{zz}, is easily extracted from the 16×16 covariance matrix $P_{k/k}$. Since

$$\boldsymbol{\delta z} = \begin{bmatrix} \Delta X(1,4) \\ \Delta X(2,4) \\ \Delta X(3,4) \end{bmatrix} \tag{69}$$

then, using (14) with $m = 4$ and $n = 1$, yields

$$P_{zz} = \begin{bmatrix} P(13,13) & P(13,14) & P(13,15) \\ P(14,13) & P(14,14) & P(14,15) \\ P(15,13) & P(15,14) & P(15,15) \end{bmatrix} \tag{70}$$

where $P(13,13)$ denotes the element of $P_{k/k}$ at location $(13,13)$. Thus, one can evaluate the covariance matrix P_{ee} by first computing $P_{k/k}$ from (55), then by extracting the submatrix P_{zz} (70) and finally by using (67) and (68). Notice that the matrix N in (68) only depends on the chosen reference directions. If the reference frame is inertial, N is computed only once at the start of the algorithm. Otherwise N is propagated using the dynamics of the reference frame, which is assumed to be accurately known.

CONSTRAINED ESTIMATION

In this section a constrained estimation algorithm is developed which enforces the properties of symmetry and zero-trace on the K-matrix estimate. For this purpose, one possible approach is to develop a reduced-order model according to the number of constraints among the state variables. This would however destroy the matrix formulation of the estimator. In order to preserve the original matrix formulation, the PM approach for constrained estimation is adopted. Next, we briefly explain the main idea of the constrained estimation approach via PM measurement. Assume that the true state variable X satisfies the following constraint

$$f(X) = 0 \tag{71}$$

where $f(X)$ may be a scalar, a vector, or a matrix mapping and assume that an imaginary device measures $f(X)$ with some small error. The associated PM equation is thus

$$f(X) = v \tag{72}$$

where v is the associated PM noise with appropriate dimension. In order to fit in the Kalman filtering framework, this noise is typically modeled as a zero-mean white noise with a given covariance matrix R_v. Based on the PM model equation (72), a PM update stage is developed in the Kalman filtering framework and sequentially implemented after a nominal unconstrained update stage. The PM noise covariance matrix R_v is used as a tuning parameter in the filter to "strengthen" or "soften" the constraint. For instance, if the value of R_v is very low, the constraint will be strongly enforced via the PM measurement update stage. In the following, PM model equations are developed for the symmetry and trace constraints.

Notice that the constrained estimation approach via PM consists of relaxing a hard constraint in the underlying optimization problem, and varying the degree of enforcement of the constraint by changing some parameter—here the covariance of a virtual noise. Thus, what matters about v is only the value given to its covariance, because that value will impact the PM update stage and will thus enforce the constraint on the estimate. That technique was successfully applied, e.g., to quaternion normalization in [17] and to direction cosine matrix orthogonalization in [18]. Reference [19] presents a comprehensive survey on constrained Kalman filtering via PM (called there pseudoobservations) and projections, develops a successful combination of these approaches, and illustrates it in a constrained quaternion estimation problem.

SYMMETRY CONSTRAINT

Symmetry Pseudomeasurement: Out of the many formulations of the symmetry constraint on a matrix X, we consider the following one:

$$\tfrac{1}{2}(X + X^T) = X. \tag{73}$$

It is assumed that an imaginary device is measuring the 4×4 state matrix X_{k+1} with some small zero-mean white noise, denoted by V_{k+1}^{sym}, and is giving as output the symmetric part of its best available estimate $\hat{X}_{k+1/k+1}$, where $\hat{X}_{k+1/k+1}$ is obtained from (53). Thus, the symmetry PM equation is

$$\tfrac{1}{2}(\hat{X}_{k+1/k+1} + \hat{X}_{k+1/k+1}^T) = X_{k+1} + V_{k+1}^{\text{sym}}. \tag{74}$$

Note that if $\hat{X}_{k+1/k+1}$ is error free and if we drop the matrix noise V_{k+1}^{sym} from (73), we recover (73), which is the desired basic property. Since, as is evident from (2), (74) has the standard structure of a linear matrix measurement, it can be incorporated into the system mathematical model. Let R_{k+1}^{sym} denote the 16×16 covariance matrix of V_{k+1}^{sym}. Using (74) and R_{k+1}^{sym}, the symmetry-measurement update stage is formulated as

$$\tilde{Y}_{k+1} = \tfrac{1}{2}(\hat{X}_{k+1/k+1}^T - \hat{X}_{k+1/k+1}) \tag{75a}$$

$$S_{k+1} = P_{k+1/k+1} + R_{k+1}^{\text{sym}} \tag{75b}$$

$$\mathcal{K}_{k+1} = P_{k+1/k+1} S_{k+1}^{-1} \tag{75c}$$

$$\hat{X}_{k+1/k+1}^{+} = \hat{X}_{k+1/k+1} + \sum_{j=1}^{4}\sum_{l=1}^{4} \mathcal{K}_{k+1}^{jl}\tilde{Y}_{k+1}E^{lj} \tag{75d}$$

$$\begin{aligned} P_{k+1/k+1}^{+} &= (I_{16} - \mathcal{K}_{k+1})P_{k+1/k+1}(I_{16} - \mathcal{K}_{k+1})^T \\ &\quad + \mathcal{K}_{k+1}R_{k+1}^{\text{sym}}\mathcal{K}_{k+1}^T \end{aligned} \tag{75e}$$

where the 4×4 matrices $\mathcal{K}_{k+1}^{jl}$ in (75d) are submatrices of the 16×16 matrix $\mathcal{K}_{k+1}$; they are defined according to the partition described in (54). The covariance matrix R_{k+1}^{sym} is a filter tuning matrix parameter, according to which one can weight the symmetry constraint in the estimation process. For example, taking R_{k+1}^{sym} to zero yields a unity gain matrix, i.e., $\mathcal{K}_{k+1} = I_{16}$, which then produces the following update stage:

$$\hat{X}_{k+1/k+1}^{+} = \tfrac{1}{2}(\hat{X}_{k+1/k+1} + \hat{X}_{k+1/k+1}^T). \tag{76}$$

In that case the symmetry update stage (76) produces the symmetric part of the previous estimate. This intuitively appealing result can also be drawn from a deterministic optimization approach; indeed, it can be shown that the symmetric part of any square matrix is its closest symmetric matrix (with respect to the Frobenius norm [25]). More generally, the latter result can be seen as a particular case of a recursive least-squares estimate converging to an orthogonal projection estimate for a vanishingly small variance parameter (see [19] for a discussion between these two approaches).

Trace Constraint

The trace constraint is handled using the following PM model:

$$0 = \text{tr}X_{k+1} + v_{k+1}^{\text{tr}} \tag{77}$$

where "tr" denotes the trace operator, v_{k+1}^{tr} is a scalar zero-mean white noise with covariance r_{k+1}^{tr}, and the value of the PM is 0. Using the definition of the trace operator

$$\begin{aligned} \text{tr}X_{k+1} &= \sum_{i=1}^{4} X_{k+1}(i,i) \\ &= \sum_{i=1}^{4} \mathbf{e}_i{}^T X_{k+1}\mathbf{e}_i \end{aligned} \tag{78}$$

where $\mathbf{e}_i$, $i = 1,2,3,4$, are the standard unit vectors in $\mathbb{R}^4$. Equation (78) has the standard form of a linear matrix measurement model. Using (78) the trace update stage is formulated as follows

$$\mathbf{h}^{\mathrm{tr}} = \sum_{i=1}^{4} (\mathbf{e}_i \otimes \mathbf{e}_i) \tag{79a}$$

$$s_{k+1} = \mathbf{h}^{\mathrm{tr}T} P_{k+1/k+1} \mathbf{h}^{\mathrm{tr}} + r^{\mathrm{tr}}_{k+1} \tag{79b}$$

$$\mathbf{k}_{k+1} = P_{k+1/k+1} \mathbf{h}^{\mathrm{tr}} / s_{k+1} \tag{79c}$$

$$\bar{K}_{k+1} = \sum_{j=1}^{4} \mathbf{k}^j_{k+1} \mathbf{e}_j{}^T \tag{79d}$$

$$\hat{X}^+_{k+1/k+1} = \hat{X}_{k+1/k+1} - (\mathrm{tr}\hat{X}_{k+1/k+1})\bar{K}_{k+1} \tag{79e}$$

$$P^+_{k+1/k+1} = (I_{16} - \mathbf{k}_{k+1}\mathbf{h}^{\mathrm{tr}T}) P_{k+1/k+1} (I_{16} - \mathbf{k}_{k+1}\mathbf{h}^{\mathrm{tr}T})^T + r^{\mathrm{tr}}_{k+1} \mathbf{k}_{k+1} \mathbf{k}^T_{k+1} \tag{79f}$$

where each vector $\mathbf{k}^j_{k+1}$ in the computation of $\bar{K}_{k+1}$ (79d) is the 4×1 column-vector at position j, $j = 1,2,3,4$, in the 16×1 gain column-vector, $\mathbf{k}_{k+1}$. Equation (79e) is proved as follows:

$$\begin{aligned} \hat{X}^+_{k+1/k+1} &= \hat{X}_{k+1/k+1} + \sum_{j=1}^{4} [(-\mathrm{tr}\hat{X}_{k+1/k+1})(\mathbf{k}^j_{k+1}\mathbf{e}_j{}^T)] \\ &= \hat{X}_{k+1/k+1} - (\mathrm{tr}\hat{X}_{k+1/k+1}) \left(\sum_{j=1}^{4} \mathbf{k}^j_{k+1} \mathbf{e}_j{}^T \right) \\ &= \hat{X}_{k+1/k+1} - (\mathrm{tr}\hat{X}_{k+1/k+1})\bar{K}_{k+1}. \end{aligned} \tag{80}$$

The first equality in (80) is the general formulation of the state measurement update stage in an MKF. The second equality stems from the fact that $\mathrm{tr}\hat{X}_{k+1/k+1}$ is independent of j, and the last equality is obtained by using (79d). Similar to the symmetry constraint case, the covariance r^{tr}_{k+1} is used as a tuning parameter in order to enforce the zero-trace property along the estimation process. In the limiting case where $r^{\mathrm{tr}}_{k+1} = 0$, straightforward computations yield

$$\bar{K}_{k+1} = \tfrac{1}{4} I_4. \tag{81}$$

Using (81) into (79e) yields

$$\hat{X}^+_{k+1/k+1} = \hat{X}_{k+1/k+1} - \tfrac{1}{4}(\mathrm{tr}\hat{X}_{k+1/k+1}) I_4. \tag{82}$$

Computing the trace of $\hat{X}_{k+1/k+1}$, as given in (82), and using the fact that $\mathrm{tr}I_4 = 4$, yields $\mathrm{tr}\hat{X}_{k+1/k+1} = 0$; that is, the updated $\hat{X}_{k+1/k+1}$ exactly satisfies the zero-trace constraint.

The constrained MKF consists of the algorithm described earlier, (46)–(55), to which the symmetry and trace update stages are added. Thus, each update stage operates on the preceding state estimate and estimation-error covariance. Like in an iterative Kalman filter, the three stages are performed sequentially at the same epoch time.

NUMERICAL STUDY

In this section the unconstrained MKF, the constrained MKF (CMKF) and the Optimal-REQUEST (OPREQ) filter [9] are tested and compared via extensive Monte-Carlo simulations. In the present case study we consider an SC with the same kinematics model as the Microwave Anisotropy Probe (MAP) satellite [22], which was launched on June 30, 2001. The attitude measurement devices simulated here are composed of a digital Sun sensor (DSS), an autonomous star-tracker (AST), and a triad of rate gyroscopes. Two Cartesian coordinate frames are considered; namely, the Sun frame, which is assumed to be inertial, and the body frame. The rotation of the body frame with respect to the Sun frame is composed of a spin rotation and of a nutation; the spin and the nutation rates are 0.464 rev/min and 1 rev/hr, respectively; the constant nutation angle, which is defined between the SC spinning axis and the anti-Sun LOS vector, is equal to 22.5 deg.

It is assumed that the AST observes the same star during the whole simulation. Therefore, two identical inertial LOS vectors are observed at each sampling time; namely, the Sun–SC LOS vector, and the star–SC LOS vector. These LOS vectors are represented in the Sun frame by the unit vectors $\mathbf{r}_1$ and $\mathbf{r}_2$, respectively. The Sun frame is assumed to have its third axis coinciding with the LOS between the SC and the Sun, thus $\mathbf{r}_1 = [0\ \ 0\ \ 1]^T$. For the sake of the example, it is assumed that the AST can find a star along a direction perpendicular to $\mathbf{r}_1$, for instance, $\mathbf{r}_2 = [1\ \ 0\ \ 0]^T$. The unit vector measurements, $\mathbf{b}_i$, $i = 1,2$, are simulated by adding a small zero-mean white Gaussian noise to the ideal observed directions, and by normalizing the result; that is

$$\mathbf{b}_i = \frac{A\mathbf{r}_i + \delta\mathbf{b}_i}{\|A\mathbf{r}_i + \delta\mathbf{b}_i\|} \tag{83}$$

where A is the correct transformation matrix from the Sun to the body coordinates and

$$\delta\mathbf{b}_i \sim \mathcal{N}\{\mathbf{0}, \sigma_i^2 I_3\} \tag{84}$$

for $i = 1,2$. Here, σ_1 equals 1 arc-mn ($\simeq$ 17 mdeg), and σ_2 equals 10 arc-sec ($\simeq$ 2.8 mdeg). The vector measurement sampling period is 10 s. The output of a triad of gyroscopes is contaminated by a zero-mean Gaussian white noise with a covariance matrix $\sigma_\epsilon^2 I_3$, where $\sigma_\epsilon = 100$ mdeg/hr. The gyros sampling period is 0.5 s. It is assumed that the initial attitude is completely unknown. Each simulation run lasts 10000 s.

Comparison between Optimal-REQUEST and the MKF

The results of the Monte-Carlo simulation (100 runs) are summarized in Figs. 5–8. In each figure the solid lines are used to plot the variables of the unconstrained MKF, that is, the full covariance filter as described in (46)–(55). The dashed lines are associated with the OPREQ variables. Let ΔX_{11}, ΔX_{13}, and ΔX_{14} denote the elements $(1,1)$, $(1,3)$, and $(1,4)$, respectively, of the updated estimation error matrix $\Delta X_{k/k}$. Extensive simulations show that the behavior of these three elements represents the behavior of all other elements of the matrix $\Delta X_{k/k}$. Figure 5 shows the time history of the means and of the $\pm 1\sigma$-envelopes of ΔX_{11}, ΔX_{13}, and ΔX_{14}. Figure 5(a) shows that the plots of $\Delta X_{11}^{\text{OPREQ}}$ and $\Delta X_{11}^{\text{MKF}}$ are practically undistinguishable. The means oscillate around zero with two time periods; namely, a short period of about 2 min, which corresponds to the spin rotation of the SC, and a long period of 1 hr, which is due to the SC nutation. The order of magnitude of the standard deviations is 5×10^{-6}. Figure 5(b) shows that, similar to ΔX_{11}, the filters OPREQ and MKF yield very close variations in ΔX_{13}. However, the oscillations in ΔX_{13} are less sensitive to the short period than in ΔX_{11}. In Fig. 5(c), however, we notice that the oscillations in the mean of $\Delta X_{14}^{\text{OPREQ}}$ are much less damped than the oscillations in the mean of $\Delta X_{14}^{\text{MKF}}$, and the ratio between the amplitudes of those oscillations reaches 8. Furthermore, the $\pm 1\sigma$-envelope of $\Delta X_{14}^{\text{MKF}}$ constitutes a lower-bound for that of $\Delta X_{14}^{\text{OPREQ}}$. The dc level of the oscillations in the standard deviation is about 4×10^{-7} for the MKF, and twice as large (8×10^{-7}) for OPREQ. Thus, for X_{14}, MKF clearly outperforms OPREQ. The advantage of MKF becomes even more obvious when analyzing the angular and quaternion estimation errors.

Next, the gains of MKF and OPREQ are compared. For this purpose we compute the scalar ρ_{MKF} as the Euclidean norm of the 16×16 gain matrix $\mathcal{K}_{k+1}$ of MKF, and plot its Monte-Carlo mean versus the mean of the gain of OPREQ ρ_{OPREQ}. Figure 6 shows the variations of the Monte-Carlo means of ρ_{MKF} (solid) and of ρ_{OPREQ} (dashed). During the transition phase, the first 1500 s, the two quantities are very close to one another. Then ρ_{MKF} reaches a steady-state value around 0.015, while ρ_{OPREQ} oscillates at the spin and nutation frequencies. The maxima of ρ_{OPREQ} are about 0.025 and the dc level of the oscillations is around 0.020. This result gives some insight into the result described earlier in Fig. 5(c); because of the higher gain, OPREQ filter weighs the new incoming observations more heavily than MKF; therefore, the update estimate in OPREQ is noisier than that in the MKF.

As mentioned earlier the quaternion estimation error, denoted by $\boldsymbol{\delta q}$, is defined as the quaternion of the small rotation that brings the estimated body frame onto the true body frame (see (66)). It has four components, which are denoted by δe_1, δe_2, δe_3, and δq. The variations of the means and $\pm 1\sigma$-envelopes of the four components of $\boldsymbol{\delta q}$ are depicted in Fig. 7. As can be seen from Figs. 7(a) and (b), the errors δe_1 and δe_2 have very similar variations. The oscillations in the means are less damped in OPREQ than in the MKF; the ratio between the oscillation peaks reaches 8. The dc level of the $\pm 1\sigma$-envelopes in OPREQ is twice that of the MKF (1.2×10^{-5} as compared with 6×10^{-6}). The same analysis applies to δe_3 in Fig. 7(c) except that the variations of $\delta e_3^{\text{OPREQ}}$ are much noisier. As opposed to $\delta e_3^{\text{OPREQ}}$, the variations of δe_3^{MKF} have a regular oscillating pattern, essentially modulated by the nutation frequency. Instead of plotting the variations of δq, we plot those of $(1 - \delta q)$ in Fig. 7(d); indeed, this quantity is the one that becomes small when the quaternion estimation error becomes small (a quaternion expressing a zero-rotation is equal to $[0\ \ 0\ \ 0\ \ 1]^T$). After a transition phase of about 1500 s, the plots of OPREQ and the MKF clearly separate. The variations of $1 - \delta q^{\text{MKF}}$ are smooth, with a mean of 10^{-10} and a standard deviation of 10^{-10}; on the other hand, the mean of $1 - \delta q^{\text{MKF}}$ oscillates above 10^{-10}, and the standard deviation is of the order of 2×10^{-10}.

Let $\delta\phi$ denote the angle of the small rotation that is represented by $\boldsymbol{\delta q}$. This angle is extracted from δq using the known relation, $\delta q = \cos(\delta\phi/2)$. Figure 8 presents the time histories of the mean and of the $\pm 1\sigma$-envelope of $\delta\phi$. The mean of the MKF is stabilized at 1.2 mdeg, while the mean of OPREQ oscillates above it, around a dc level of 1.7 mdeg, i.e., the mean error of OPREQ is about 1.4 times larger than that of MKF. The standard deviation in the MKF is about 0.8 mdeg, about 1.5 times smaller than that in OPREQ, which is 1.2 mdeg.

Overall, we see that MKF outperforms OPREQ. We also deduce from Fig. 7(d) that there is a small bias in the estimated quaternion. Considering the order of magnitude in $\delta\phi$, both algorithms perform well since the estimation angular error is at a level of 1.2 mdeg (in MKF) and of 1.7 mdeg (in OPREQ), which is less than the measurement angular errors, i.e., about 3 mdeg in the most accurate measurements, and about 17 mdeg in the least accurate ones.

Discussion

We saw in Fig. 5(c) that the Monte-Carlo standard deviation (Monte-Carlo STD) in ΔX_{14} were twice as large in OPREQ as in MKF; the same ratio in favor of MKF appeared in Figs. 7(a) and (b) with respect to the Monte-Carlo STD in δe_1 and δe_2. Fig.7(c)

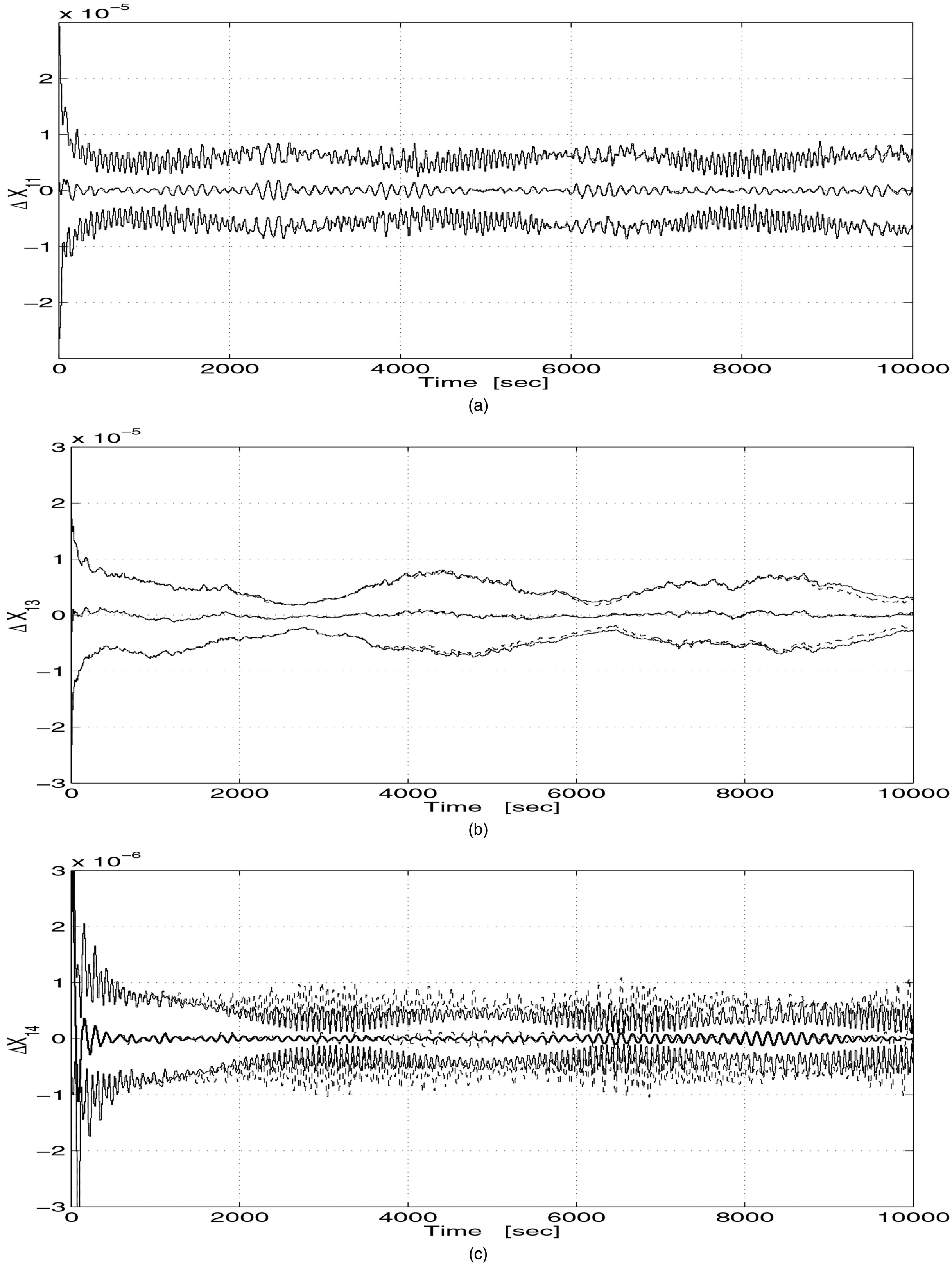


Fig. 5. Monte-Carlo means and $\pm 1\sigma$ envelopes of the estimation errors ΔX_{11}, ΔX_{13} and ΔX_{14} in MKF (solid) and in the OPREQ filter (dashed). (a) ΔX_{11}. (b) ΔX_{13}. (c) ΔX_{14}.

features a ratio of 1.5 between the Monte-Carlo STD of δe_3 in OPREQ and MKF. The ratio between the Monte-Carlo STD (and means) of δq in OPREQ and MKF is also 1.5, according to Fig. 7(d). Finally, Fig. 8 shows a ratio of 1.4 between the Monte-Carlo means of $\delta\phi$ in both filters, and the same holds

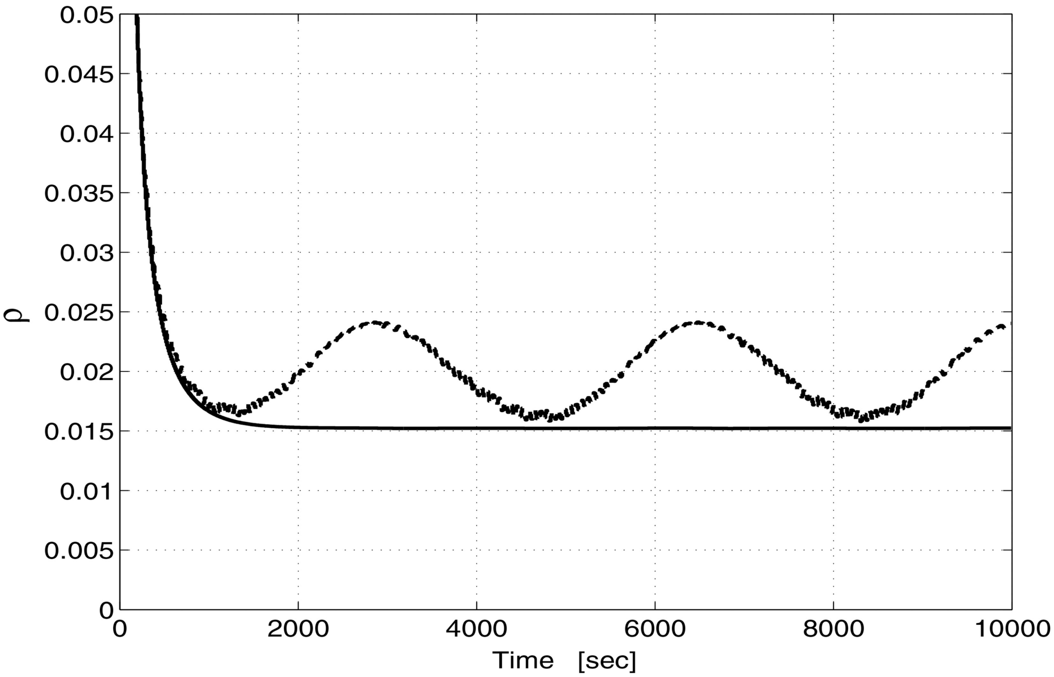


Fig. 6. Monte-Carlo means of the gains ρ_{OPREQ} (dashed) and ρ_{MKF} (solid).

for the Monte-Carlo STD of $\delta\phi$. We can therefore conclude from these results that the MKF algorithm outperforms the OPREQ algorithm, and that the increase in performance can be quantified by ratios between 1.4 and 2.

The present discussion is concerned with an analysis of the ratio in the performances increase. We first derive conditions under which the MKF algorithm reduces to OPREQ. Assume that W_k, V_k, and $P_{0/0}$ have independent identically distributed rows with 4×4 covariance matrices, $\bar{Q}_k$, $\bar{R}_k$, and $\bar{P}_{0/0}$, respectively.[2] In addition, assume that the gain matrix in (52), $\mathcal{K}_{k+1}$, is a scalar matrix, i.e.,

$$\mathcal{K}_{k+1} = \rho_{k+1} I_{16} \tag{85}$$

then the filter update equations (52) and (53) become

$$\hat{X}_{k+1/k+1} = (1-\rho_{k+1})\hat{X}_{k+1/k} + \rho_{k+1} Y_{k+1} \tag{86a}$$

$$\bar{P}_{k+1/k+1} = (1-\rho_{k+1})^2 \bar{P}_{k+1/k} + \rho_{k+1}^2 \bar{R}_{k+1} \tag{86b}$$

where $\hat{X}$ denotes the estimate of the K-matrix, Y_{k+1} is the matrix measurement constructed using the vector measurement acquired at t_{k+1}, $\bar{P}$ denotes the 4×4 estimation error covariance matrix for each row of the estimation error matrix, and $\bar{R}$ denotes the covariance matrix of the effective measurement noise used in MKF. Equations (86) are obtained by using (85) in (63) and (64). The update equations in OPREQ are written here for convenience [9]

$$K_{k+1/k+1} = (1-\rho^*_{k+1})K_{k+1/k} + \rho^*_{k+1}\delta K_{k+1} \tag{87a}$$

$$P_{k+1/k+1} = (1-\rho^*_{k+1})^2 P_{k+1/k} + {\rho^*_{k+1}}^2 \mathcal{R}_{k+1} \tag{87b}$$

where $K_{k+1/k+1}$ denotes the updated estimate of the K-matrix, δK_{k+1} is the matrix measurement at t_{k+1}, ρ^*_{k+1} is the optimized fading memory factor, $P_{k+1/k+1}$ and $\mathcal{R}_{k+1}$ denote the "uncertainty" matrices for the updated estimation error and for the effective measurement noise used in OPREQ, respectively. Comparing (86) and (87) we realize that they have similar structures. They differ, however, because the matrices $\bar{P}_{k+1/k+1}$ and $P_{k+1/k+1}$ are different, and so are the matrices $\bar{R}_{k+1}$ and $\mathcal{R}_{k+1}$. In fact, these matrices are related as follows

$$\mathcal{R}_{k+1} = 4\bar{R}_{k+1} \tag{88}$$

$$P_{k+1/k+1} = 4\bar{P}_{k+1/k+1}. \tag{89}$$

Equation (88) is easily shown by recalling the definition of $\mathcal{R}_{k+1}$ [9], which yields

$$\begin{aligned}
\mathcal{R}_{k+1} &\triangleq E[V_{k+1}V_{k+1}^T] \\
&= E\left[\sum_{i=1}^{4} \mathbf{v}_i^c \mathbf{v}_i^{cT}\right] \\
&= \sum_{i=1}^{4} E[\mathbf{v}_i^c \mathbf{v}_i^{cT}] \\
&= 4\bar{R}_{k+1}
\end{aligned}$$

[2]These assumptions lead to to the reduced covariance MKF as given in (57)–(64).

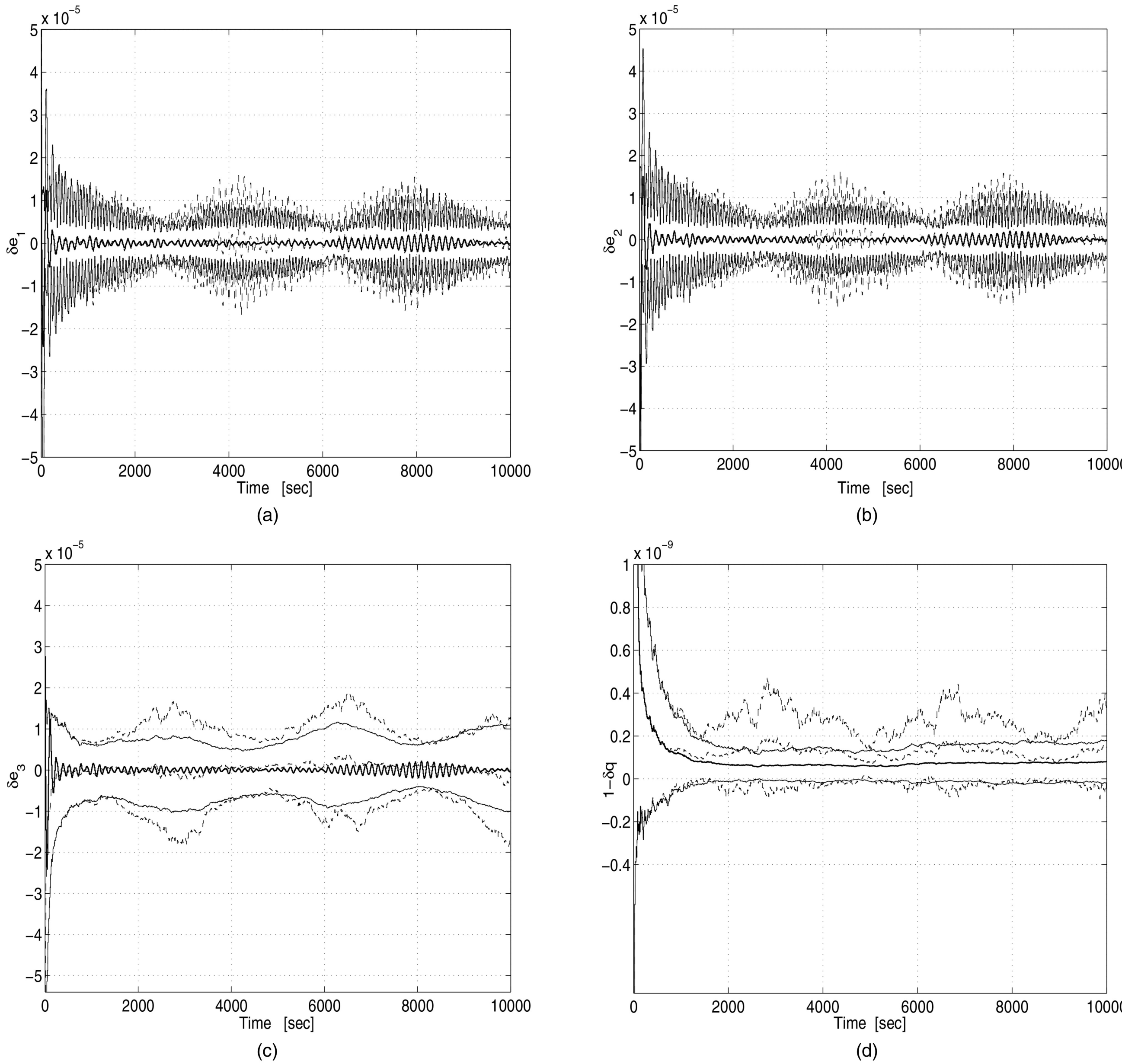


Fig. 7. Monte-Carlo means and $\pm 1\sigma$ envelopes of the quaternion estimation errors in MKF (solid) and in OPREQ filter (dashed). (a) δe_1. (b) δe_2. (c) δe_3. (d) $1-\delta q$.

where $\mathbf{v}_i^c$, $i = 1,2,3,4$, denote the 4×1 column-vectors of the matrix V_{k+1}. Notice that these vectors are identical to the rows of V_{k+1} since V_{k+1} is a symmetric matrix (see (39) and (40)). The third equality is due to the linearity of the expectation operator. The last equality stems from the assumption that the rows of V_{k+1} (and thus also its columns) are independent and identically distributed, with covariance matrix $\bar{R}_{k+1}$. The same argument is readily used for $P_{k+1/k+1}$. As a result, the covariance update equations of the MKF and of OPREQ are identical, up to a multiplication by a constant.

While finding the conditions under which the general MKF algorithm reduces to the OPREQ algorithm, we have quantified the difference between the effective measurement noise levels in the filters, and found that it is four times greater in OPREQ than in the matrix filter. It is believed that this is the principal cause of the discrepancy between the filters' performance. The latter is illustrated by a simple example. Consider the following scalar state-space equations

$$x_{k+1} = x_k + w_k \tag{90}$$

$$y_{k+1} = x_{k+1} + v_{k+1} \tag{91}$$

where w_k and v_{k+1} are the process and measurement noise sequences, respectively, which satisfy the usual stochastic assumptions of the basic state-space model. Let q and r denote the covariances of w_k and v_k, respectively. The scalar algebraic Riccati equation is

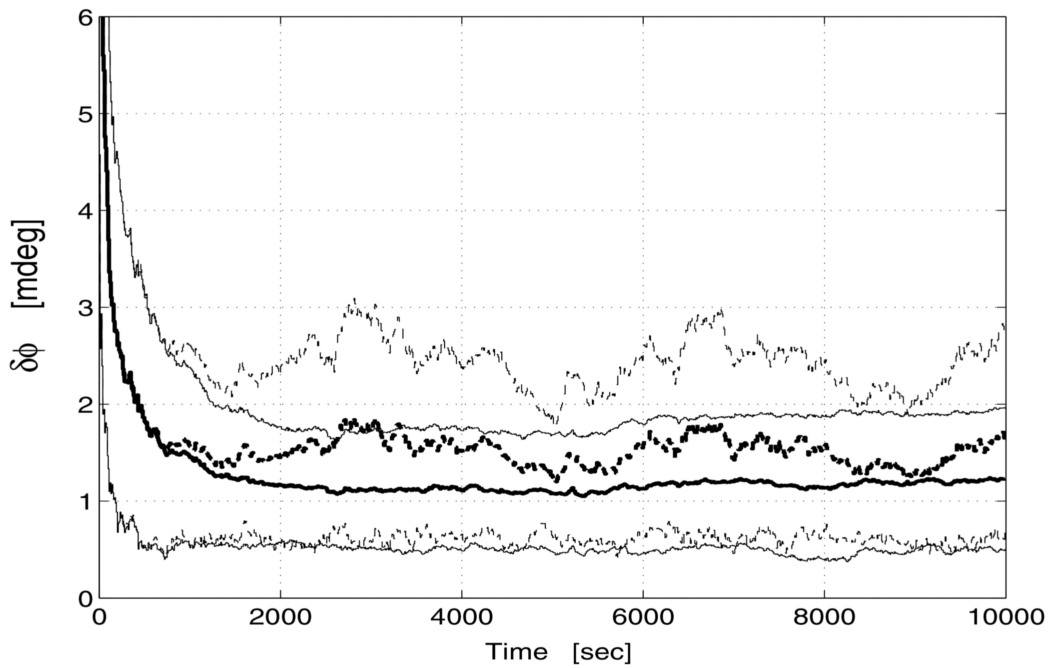


Fig. 8. Monte-Carlo means and $\pm 1\sigma$ envelopes of the angular estimation error $\delta\phi$ in MKF (solid) and in OPREQ filter (dashed).

readily formulated as follows:

$$p_\infty^2 + qp_\infty - qr = 0 \tag{92}$$

where p_∞ denotes the steady-state estimation error covariance. Solving for p_∞ in (92) and assuming[3] that $q/r \ll 1$ yields the following approximation for p_∞:

$$p_\infty \simeq \sqrt{qr}. \tag{93}$$

It is clear from (93) that multiplying the measurement noise covariance r by 4 will deteriorate the estimation error standard deviation $\sqrt{p_\infty}$ by factor of $\sqrt{2} \simeq 1.4$, which is close to the Monte-Carlo simulations results.

Constrained Matrix Kalman Filter

We test here the performance of the CMKF, which is the MKF that embeds the symmetry and trace update stages. We compare the response of the CMKF and the MKF to an initial perturbation in the symmetry and trace properties of the estimate matrix. The initial perturbations in the elements of the initial estimate are zero-mean uniformly distributed random variables with standard deviation $\sigma_{dst} = 0.05$. The covariance matrix of the symmetry PM R_{k+1}^{sym} is chosen as $R_{k+1}^{\text{sym}} = (\sigma_{dst})^2 I_{16}$, and the covariance of the trace PM is $r_{k+1}^{\text{tr}} = \sigma_{dst}^2$. All the other simulation conditions are identical to those of the preceding simulation. The results of a 100-run Monte-Carlo simulation are presented in Figs. 9 and 10. For both MKF and CMKF, Fig. 9 shows plots of the MC-means for $\delta\phi$ and Fig. 10 shows the plots of the MC-means of the quaternion estimation errors, δe_1, δe_2, δe_3, and $1 - \delta q$. An inspection of Fig. 9 reveals that CMKF performs better than MKF during the transient phase and the steady-state phase. Figure 10 further illustrates the fact that constraining the estimation process speeds up the error transient response. Furthermore, it appears that the initial symmetry and trace perturbation yields an estimation performance degradation, as compared with the MKF without initial perturbation, by a factor of about 50.

CONCLUSION

A novel recursive estimator of the quaternion-of-rotation from sequential vector observations is presented. The proposed estimation algorithm is an enhanced OPREQ filter, where the first step consists of denoising the elements of a time-varying K-matrix via Kalman filtering techniques. The K-matrix estimator is developed using the MKF paradigm. Explicit expressions for the covariance matrices of the process and measurement matrix noises are developed. An exact treatment of the state-multiplicative process noise in the Kalman filtering framework is provided. A reduced estimator is developed under special assumptions on the noise stochastic models. Constraining the symmetry and zero-trace properties in the matrix estimate is done in the MKF framework via PM techniques. Extensive Monte-Carlo simulations are used to compare the performance of the unconstrained MKF with that of OPREQ, and to illustrate the advantage of constraining the estimation process. Although both algorithms exhibit, in general, similar transition

[3]This assumption is fully justified in the application case where the process noise comes from gyro outputs and the measurement noise comes from vector measurement sensing devices. In this case, $q \simeq 10^{-14}$ [rad^2/s^2] and $r \simeq 10^{-10}$.

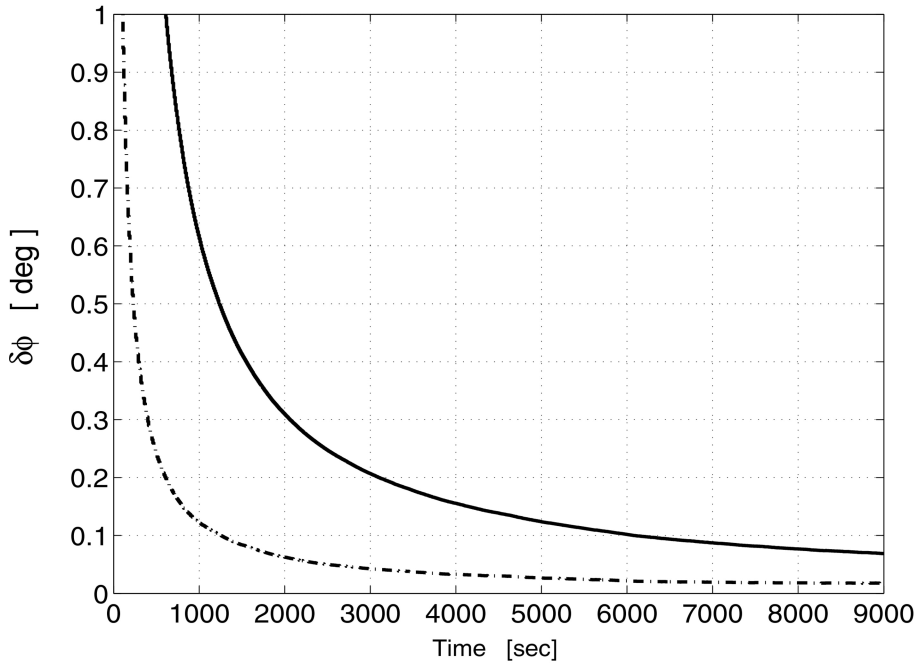


Fig. 9. Monte-Carlo mean of the angular estimation error $\delta\phi$ in the MKF (solid) and in the CMKF (dashed).

phases, the MKF clearly outperforms OPREQ in steady-state. The Monte-Carlo means of all the estimation errors in MKF are much more damped, and the Monte-Carlo estimation error standard deviations are between 1.4 times and twice as small as those of OPREQ. The Monte-Carlo results show that the CMKF achieves a better accuracy in the case of initial perturbations in the desired estimate properties.

APPENDIX I. DERIVATION OF (29)

We begin by presenting some known matrix identities. For any vectors $\mathbf{u}$, $\mathbf{v}$ in $\mathbb{R}^3$ and any general matrix M in $\mathbb{R}^{3\times 3}$ the following identities hold:

$$[\mathbf{u}\times]\mathbf{v} = -[\mathbf{v}\times]\mathbf{u} \tag{94a}$$

$$[\mathbf{u}\times][\mathbf{v}\times] = \mathbf{v}\mathbf{u}^T - \mathbf{v}^T\mathbf{u}I_3 \tag{94b}$$

$$[(\mathbf{u}\times\mathbf{v})\times] = \mathbf{v}\mathbf{u}^T - \mathbf{u}\mathbf{v}^T \tag{94c}$$

$$\mathbf{v} = [\mathrm{tr}(M)I_3 - M]\mathbf{u} \quad \Rightarrow \quad [\mathbf{v}\times] = M^T[\mathbf{u}\times] + [\mathbf{u}\times]M \tag{94d}$$

$$[\mathbf{u}\times] = M^T - M \quad \Rightarrow \quad \mathbf{u}^T\mathbf{v} = \mathrm{tr}([\mathbf{v}\times]M). \tag{94e}$$

All these results arise from the definition of the cross-product matrix and can be easily established by direct computation. Equations (94a) to (94d) correspond to [5, eqs. (A15)–(A18)]. Next, we recall that W_k can be approximated to first-order in $\boldsymbol{\epsilon}_k$ and Δt by [9]

$$W_k \simeq \begin{bmatrix} S_\epsilon - \kappa_\epsilon I_3 & \mathbf{z}_\epsilon \\ \mathbf{z}_\epsilon^T & \kappa_\epsilon \end{bmatrix} \Delta t \tag{95}$$

where

$$\begin{aligned} B_\epsilon &= [\boldsymbol{\epsilon}_k\times]B_k, \qquad S_\epsilon = B_\epsilon + B_\epsilon^T \\ [\mathbf{z}_\epsilon\times] &= B_\epsilon^T - B_\epsilon, \qquad \kappa_\epsilon = \mathrm{tr}(B_\epsilon). \end{aligned} \tag{96}$$

This form is valid for both high and low angular velocities. The expression for W_k, as given in (95), results from a Taylor expansion of the discrete-time dynamics matrix Φ_k in (28) to first-order in the gyro error $\boldsymbol{\epsilon}_k$, and in the time increment Δt. The angular velocity components indeed enter the neglected second-order terms in the expansion of Φ_k as follows (see [26, Appendix B] for the proof):

$$\Delta\Phi = \mathcal{E}\Delta t + \tfrac{1}{2}(\Omega\mathcal{E} + \mathcal{E}\Omega)\Delta t^2 - \tfrac{1}{2}\mathcal{E}^2\Delta t^2 + \mathcal{O}(\Delta t^3) \tag{97}$$

where Ω and $\mathcal{E}$ are defined in (30) and (27), respectively. Equation (97) features the second-order terms of the quaternion transition matrix approximation presented in [27] (see (43) there), under the assumption of a zero-hold assumption on the integrated gyro measured rates and gyro noises. From (97), the ratio between the norms of the first-order term in Δt and the second-order term in Δt, which involves Ω, is of the order $(\|\boldsymbol{\omega}\|\Delta t)$ where $\|\boldsymbol{\omega}\|$ is the angular velocity norm. Even for a very high velocity of 1 rad/s, a time increment of 1 ms is sufficient to make the ratio of the order of 10^{-3}, which proves the validity of the first-order approximation.

The matrix B_k in (96) is associated with the ideal noise-free matrix K_k. The vector $\boldsymbol{\epsilon}_k$ denotes an additive gyro output white noise and Δt is the incremental time between two gyro readings. It is shown in the following that (95) and (29) are

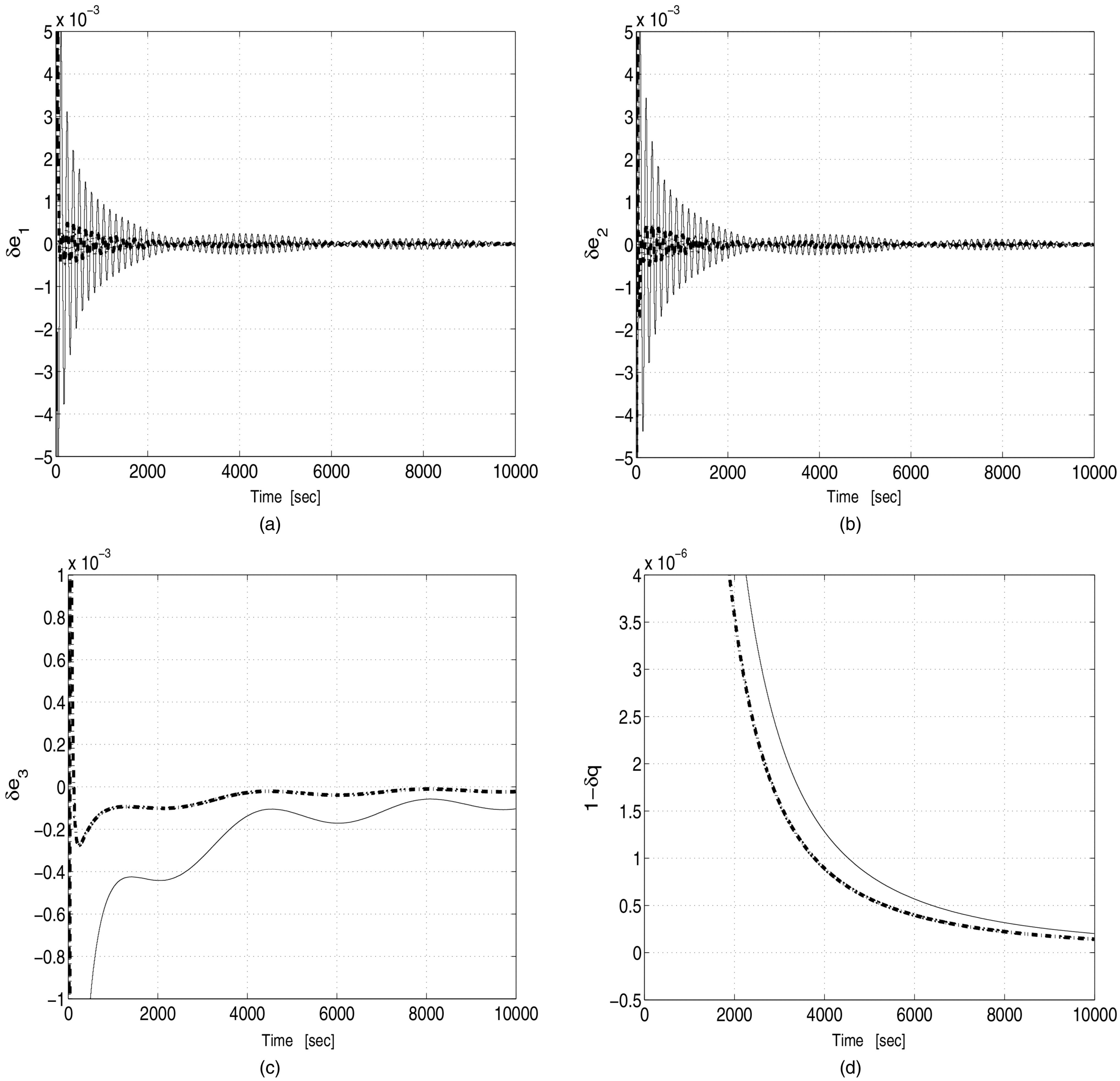


Fig. 10. Monte-Carlo means of the quaternion estimation errors in the matrix Kalman filter MKF (solid) and in the Constrained matrix Kalman filter (CMKF) (dashed). (a) δe_1. (b) δe_2. (c) δe_3. (d) $1-\delta q$.

equivalent. The time subscripts are omitted for the sake of clarity. Thus X_k and $\mathcal{E}_k$ are denoted by X and $\mathcal{E}$. Since X is symmetric and $\mathcal{E}$ is skew-symmetric,

$$X\mathcal{E} - \mathcal{E}X = X\mathcal{E} + (X\mathcal{E})^T. \tag{98}$$

Using (30) and exploiting the structure of X, which is a K-matrix, the first term on the right-hand side of (98) can be rewritten as follows

$$\begin{aligned} X\mathcal{E} &= \frac{1}{2}\begin{bmatrix} S-\sigma I_3 & \mathbf{z} \\ \mathbf{z}^T & \sigma \end{bmatrix}\begin{bmatrix} -[\boldsymbol{\epsilon}\times] & \boldsymbol{\epsilon} \\ -\boldsymbol{\epsilon}^T & 0 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} -S[\boldsymbol{\epsilon}\times] + \sigma[\boldsymbol{\epsilon}\times] - \mathbf{z}\boldsymbol{\epsilon}^T & S\boldsymbol{\epsilon} - \sigma\boldsymbol{\epsilon} \\ -\mathbf{z}^T[\boldsymbol{\epsilon}\times] - \sigma\boldsymbol{\epsilon}^T & \mathbf{z}^T\boldsymbol{\epsilon} \end{bmatrix}. \end{aligned} \tag{99}$$

Using (99) in the right-hand side of (98) yields

$$\begin{aligned} & X\mathcal{E} - \mathcal{E}X \\ &\quad = \frac{1}{2}\begin{bmatrix} [\boldsymbol{\epsilon}\times]S - S[\boldsymbol{\epsilon}\times] - (\mathbf{z}\boldsymbol{\epsilon}^T + \boldsymbol{\epsilon}\mathbf{z}^T) & S\boldsymbol{\epsilon} + [\boldsymbol{\epsilon}\times]\mathbf{z} - 2\sigma\boldsymbol{\epsilon} \\ \boldsymbol{\epsilon}^T S - \mathbf{z}^T[\boldsymbol{\epsilon}\times] - 2\sigma\boldsymbol{\epsilon}^T & 2\mathbf{z}^T\boldsymbol{\epsilon} \end{bmatrix} \\ &\quad = \begin{bmatrix} M_{11} & \mathbf{m}_{12} \\ \mathbf{m}_{12}^T & m_{22} \end{bmatrix} \end{aligned} \tag{100}$$

which constitutes an implicit definition of the 3×3 submatrix M_{11}, the 3×1 vector $\mathbf{m}_{12}$, and the scalar m_{22}. Using (94b), with $\mathbf{u} = \mathbf{z}$ and $\mathbf{v} = \boldsymbol{\epsilon}$, yields

$$[\mathbf{z}\times][\boldsymbol{\epsilon}\times] = \boldsymbol{\epsilon}\mathbf{z}^T - \mathbf{z}^T\boldsymbol{\epsilon} I_3 \tag{101}$$

where we use the fact that $\epsilon^T \mathbf{z} = \mathbf{z}^T \epsilon$. Using the fact that [1, p. 428]

$$[\mathbf{z}\times] = B^T - B \tag{102}$$

in the left-hand side of (101) yields

$$\epsilon \mathbf{z}^T = (B^T - B)[\epsilon\times] + \mathbf{z}^T \epsilon I_3. \tag{103}$$

Summing (103) with its transpose, and utilizing the skew-symmetry of $[\epsilon\times]$, yields

$$\epsilon \mathbf{z}^T + \mathbf{z}\epsilon^T = (B^T - B)[\epsilon\times] - [\epsilon\times](B - B^T) + 2\mathbf{z}^T \epsilon I_3. \tag{104}$$

Inserting (104) into the expression for M_{11}, using $S = B + B^T$, yields

$$\begin{aligned} M_{11} &= \tfrac{1}{2}\{[\epsilon\times](B + B^T) - (B + B^T)[\epsilon\times] - (B^T - B)[\epsilon\times] \\ &\qquad + [\epsilon\times](B - B^T) - 2\mathbf{z}^T \epsilon I_3\} \\ &= ([\epsilon\times]B - B^T[\epsilon\times]) - \mathbf{z}^T \epsilon I_3. \end{aligned} \tag{105}$$

Let the 3×3 matrix B_ϵ be defined by

$$B_\epsilon \stackrel{\Delta}{=} [\epsilon\times]B. \tag{106}$$

Using (94e) with $\mathbf{u} = \mathbf{z}$, $M = B$, and $\mathbf{v} = \epsilon$, yields

$$\mathbf{z}^T \epsilon = \mathrm{tr}([\epsilon\times]B). \tag{107}$$

Define the matrix S_ϵ and the scalar κ_ϵ by

$$S_\epsilon \stackrel{\Delta}{=} B_\epsilon + B_\epsilon^T \tag{108}$$

$$\kappa_\epsilon \stackrel{\Delta}{=} \mathrm{tr}(B_\epsilon). \tag{109}$$

Using (107), (108), and (109) in (105) yields the following expression for M_{11}:

$$M_{11} = S_\epsilon - \kappa_\epsilon I_3. \tag{110}$$

The expression for m_{22} immediately stems from (107) and (108):

$$m_{22} = \kappa_\epsilon. \tag{111}$$

In order to express $\mathbf{m}_{12}$, we start from its definition as given in (100):

$$\begin{aligned} \mathbf{m}_{12} &= \tfrac{1}{2}(S\epsilon + [\epsilon\times]\mathbf{z} - 2\sigma\epsilon) \\ &= \tfrac{1}{2}(S\epsilon - [\mathbf{z}\times]\epsilon - 2\sigma\epsilon) \\ &= \tfrac{1}{2}(S - [\mathbf{z}\times] - 2\sigma I_3)\epsilon \\ &= \tfrac{1}{2}[(B + B^T) - (B^T - B) - 2\sigma I_3]\epsilon \\ &= (B - \sigma I_3)\epsilon \end{aligned} \tag{112}$$

where the second line stems from (94a) (with $\mathbf{u} = \epsilon$ and $\mathbf{v} = \mathbf{z}$), and the fourth line is obtained using $S = B + B^T$ and (102). In order to compute the cross-product matrix $[\mathbf{m}_{12}\times]$, we use the fact that $\sigma = \mathrm{tr}(B)$, and we apply the proposition (94d), where $\mathbf{v} = \mathbf{m}_{12}$, $M = -B$, and $\mathbf{u} = \epsilon$; thus

$$\begin{aligned} [\mathbf{m}_{12}\times] &= -B^T[\epsilon\times] + [\epsilon\times](-B) \\ &= ([\epsilon\times]B)^T - [\epsilon\times]B \\ &= B_\epsilon^T - B_\epsilon \end{aligned} \tag{113}$$

where the third equality stems from (106). Therefore, denoting the vector $\mathbf{m}_{12}$ by $\mathbf{z}_\epsilon$, we can write

$$[\mathbf{z}_\epsilon\times] = B_\epsilon^T - B_\epsilon. \tag{114}$$

To conclude, using (110), (111), and (114) in (100) yields

$$X\mathcal{E} - \mathcal{E}X = \begin{bmatrix} S_\epsilon - \kappa_\epsilon I_3 & \mathbf{z}_\epsilon \\ \mathbf{z}_\epsilon^T & \kappa_\epsilon \end{bmatrix} \tag{115}$$

where

$$\begin{aligned} B_\epsilon &\stackrel{\Delta}{=} [\epsilon\times]B, \qquad S_\epsilon \stackrel{\Delta}{=} B_\epsilon + B_\epsilon^T \\ [\mathbf{z}_\epsilon\times] &\stackrel{\Delta}{=} B_\epsilon^T - B_\epsilon, \qquad \kappa_\epsilon \stackrel{\Delta}{=} \mathrm{tr}(B_\epsilon). \end{aligned} \tag{116}$$

APPENDIX II. PROOF OF PROPOSITION 1

Preliminaries

The process equation (26) is rewritten here for convenience

$$X_{k+1} = \Phi_k X_k \Phi_k^T + W_k \tag{117}$$

where the noise matrix W_k is expressed as

$$W_k = (X_k \mathcal{E}_k - \mathcal{E}_k X_k)\Delta t \tag{118}$$

and $\mathcal{E}_k$ is the 4×4 skew-symmetric matrix defined as

$$\mathcal{E}_k = \frac{1}{2}\begin{bmatrix} -[\epsilon_k\times] & \epsilon_k \\ -\epsilon_k^T & 0 \end{bmatrix} \tag{119}$$

with ϵ_k denoting the additive noise error in the gyro outputs. As mentioned earlier, the covariance matrix of W_k, denoted by Q_k, is defined as the covariance of $\mathrm{vec}W_k$, where $\mathrm{vec}W_k$ is the 16×1 vector obtained by applying the vec-operator on the 4×4 matrix W_k. The column-vector $\mathrm{vec}W_k$ is denoted by $\mathbf{w}_k$. Thus,

$$Q_k = \mathrm{cov}\{W_k\} \stackrel{\Delta}{=} \mathrm{cov}\{\mathbf{w}_k\}. \tag{120}$$

In order to derive the expression for Q_k, as given in (33), the vec-operator is applied to (118), which yields a linear relation between $\mathbf{w}_k$ and ϵ_k. Then, the covariance matrix of $\mathbf{w}_k$ is expressed as a function of the covariance matrix of ϵ_k, $Q_k^\epsilon/\Delta t$. It is assumed that the matrix Q_k^ϵ is known.

$$\begin{aligned} \mathbf{w}_k &= [\mathrm{vec}(X_k\mathcal{E}_k - \mathcal{E}_k X_k)]\Delta t \\ &= [\mathrm{vec}(X_k\mathcal{E}_k) - \mathrm{vec}(\mathcal{E}_k X_k)]\Delta t \\ &= [(I_4 \otimes X_k)\mathrm{vec}\mathcal{E}_k - (X_k^T \otimes I_4)\mathrm{vec}\mathcal{E}_k]\Delta t \\ &= [(I_4 \otimes X_k) - (X_k^T \otimes I_4)]\mathrm{vec}\mathcal{E}_k \Delta t \end{aligned} \tag{121}$$

where the second equality is obtained using the linearity property of the vec-operator, the third equality is derived using a basic property of the Kronecker product (see [24, p. 255]), and the last equality stems from a factorization with respect to $\mathrm{vec}\mathcal{E}_k$. Equation (121) shows that $\mathbf{w}_k$ is a state-dependent linear function of $\mathrm{vec}\mathcal{E}_k$.

PROOF OF i. Applying the vec-operator to (119) yields

$$
\begin{aligned}
(\mathrm{vec}\mathcal{E}_k)^T &= \left\{\mathrm{vec}\left(\frac{1}{2}\begin{bmatrix} -[\boldsymbol{\epsilon}_k\times] & \boldsymbol{\epsilon}_k \\ -\boldsymbol{\epsilon}_k^T & 0 \end{bmatrix}\right)\right\}^T \\
&= \frac{1}{2}\left\{\mathrm{vec}\left(\begin{bmatrix} 0 & \epsilon_3 & -\epsilon_2 & \epsilon_1 \\ -\epsilon_3 & 0 & \epsilon_1 & \epsilon_2 \\ \epsilon_2 & -\epsilon_1 & 0 & \epsilon_3 \\ -\epsilon_1 & -\epsilon_2 & -\epsilon_3 & 0 \end{bmatrix}\right)\right\}^T \\
&= \frac{1}{2}[0 \quad -\epsilon_3 \quad \epsilon_2 \quad -\epsilon_1 \quad \epsilon_3 \quad 0 \quad -\epsilon_1 \quad \epsilon_2 \quad -\epsilon_2 \quad \epsilon_1 \quad 0 \quad -\epsilon_3 \quad \epsilon_1 \quad \epsilon_2 \quad \epsilon_3 \quad 0] \\
&= \frac{1}{2}[\epsilon_1 \quad \epsilon_2 \quad \epsilon_3]\begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix} \\
&= \tfrac{1}{2}[\epsilon_1 \quad \epsilon_2 \quad \epsilon_3][-[\mathbf{e}_1\times] \quad -\mathbf{e}_1 \quad -[\mathbf{e}_2\times] \quad -\mathbf{e}_2 \quad -[\mathbf{e}_3\times] \quad -\mathbf{e}_3 \quad I_3 \quad \mathbf{0}] \\
&= \tfrac{1}{2}(M\boldsymbol{\epsilon}_k)^T
\end{aligned} \tag{122}
$$

where $\boldsymbol{\epsilon}_k = [\epsilon_1 \ \ \epsilon_2 \ \ \epsilon_3]^T$, and the 16×3 matrix M is defined in (122). Taking the transpose of (122) yields

$$\mathrm{vec}\mathcal{E}_k = \tfrac{1}{2}M\boldsymbol{\epsilon}_k. \tag{123}$$

Using (123) in (121) yields

$$\mathbf{w}_k = \tfrac{1}{2}[(I_4\otimes X_k) - (X_k^T\otimes I_4)]M\boldsymbol{\epsilon}_k\Delta t. \tag{124}$$

Defining the 16×3 matrix Γ_k as

$$\Gamma_k \triangleq \tfrac{1}{2}[(I_4\otimes X_k) - (X_k^T\otimes I_4)]M \tag{125}$$

and using (125) in (124) yields

$$\mathbf{w}_k = \Gamma_k\boldsymbol{\epsilon}_k\Delta t. \tag{126}$$

Squaring (126) and applying the expectation operator yields the sought result.

PROOF OF ii. Let Γ_{C_i} denote the three 16-dimensional columns of Γ_k, then (126) can be rewritten as follows:

$$\mathbf{w}_k = \sum_{i=1}^{3}\Gamma_{C_i}\epsilon_i\Delta t. \tag{127}$$

Let the three matrices A_i, $i = 1,2,3$, which are given in (36), be defined by

$$A_i \triangleq \mathrm{vec}^{-1}(M\mathbf{e}_i) \tag{128}$$

where vec^{-1} denotes the inverse of the vec operator. Then, using basic properties of the Kronecker product yields the following identities

$$
\begin{aligned}
\Gamma_{C_i} &= \Gamma_k\mathbf{e}_i \\
&= \tfrac{1}{2}[(A_i^T\otimes I_4) - (I_4\otimes A_i)]\mathrm{vec}X_k \\
&= \Upsilon_i\mathrm{vec}X_k.
\end{aligned} \tag{129}
$$

Substituting (129) into (127) yields

$$\mathbf{w}_k = \sum_{i=1}^{3}\Upsilon_i\mathrm{vec}X_k\epsilon_i\Delta t. \tag{130}$$

Finally, squaring (130), applying the expectation operator, and using the statistical assumptions on $\boldsymbol{\epsilon}_k$ yields the sought result.

APPENDIX III. MEASUREMENT NOISE COVARIANCE MATRIX R_{k+1}

In the case of m simultaneous vector observations at time t_{k+1}, the measurement noise term in (37), V_{k+1}, is given by (38), which is rewritten here for convenience

$$v_{k+1} = \sum_{i=1}^{m}\alpha_i V_{k+1}^i \tag{131}$$

where

$$V_{k+1}^i = \begin{bmatrix} \mathcal{S}_{b_i} - \kappa_{b_i}I_3 & \mathbf{z}_{b_i} \\ \mathbf{z}_{b_i}^T & \kappa_{b_i} \end{bmatrix} \tag{132a}$$

$$B_{b_i} = \boldsymbol{\delta}\mathbf{b}_i\mathbf{r}_i^T, \qquad \mathcal{S}_{b_i} = B_{b_i}^T + B_{b_i} \tag{132b}$$

$$\mathbf{z}_{b_i} = \boldsymbol{\delta}\mathbf{b}_i\times\mathbf{r}_i, \qquad \kappa_{b_i} = \mathrm{tr}(B_{b_i}) \tag{132c}$$

$$\alpha_i = a_i\Big/\sum_{i=1}^{m}a_i \tag{132d}$$

and a_i, $i = 1,2\ldots,m$, are positive weights. Let R_{k+1} and $R_{k+1}^{\mathbf{b}_i}$ denote, respectively, the covariance matrices of V_{k+1} and $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$, $i = 1,2\ldots,m$. The matrix V_{k+1} is expressed as a linear function of $\boldsymbol{\delta}\mathbf{b}_i(t_{k+1})$, $i = 1,2\ldots,m$, and the matrix R_{k+1} is expressed as a function of the matrices $R^{\mathbf{b}_i}_{k+1}$. Let $\mathbf{v}_{k+1}$ and $\mathbf{v}_i(t_{k+1})$ denote the vec-transforms of V_{k+1} and V^i_{k+1}, $i = 1,2\ldots,m$, respectively. Applying the vec-operator to (131) yields

$$\mathbf{v}_{k+1} = \mathrm{vec}\left(\sum_{i=1}^{m}\alpha_i V^i_{k+1}\right) = \sum_{i=1}^{m}\alpha_i \mathbf{v}_i(t_{k+1}). \tag{133}$$

PROPOSITION 2 *The matrix V^i_{k+1} given in (132) can be factorized as follows*

$$V^i_{k+1} = \tilde{R}_i^T \delta\mathcal{B}_i \tag{134}$$

where

$$\tilde{R}_i = \begin{bmatrix} [\mathbf{r}_i\times] & \mathbf{r}_i \\ -\mathbf{r}_i^T & 0 \end{bmatrix} \tag{135a}$$

$$\delta\mathcal{B}_i = \begin{bmatrix} -[\boldsymbol{\delta}\mathbf{b}_i i\times] & \boldsymbol{\delta}\mathbf{b}_i \\ -\boldsymbol{\delta}\mathbf{b}_i^T & 0 \end{bmatrix}. \tag{135b}$$

PROOF Proposition 2 is proven by direct computation, and using (94b), with $\mathbf{u} = \mathbf{r}$, and $\mathbf{v} = \boldsymbol{\delta}\mathbf{b}_i$.

Applying the vec-operator on (134) yields

$$\begin{aligned}\mathbf{v}_i(t_{k+1}) &= \mathrm{vec}(\tilde{R}_i^T\delta\mathcal{B}_i)\\ &= (I_4\otimes\tilde{R}_i^T)\mathrm{vec}(\delta\mathcal{B}_i)\\ &= -(I_4\otimes\tilde{R}_i)\mathrm{vec}(\delta\mathcal{B}_i)\end{aligned} \tag{136}$$

where the second equality is obtained using a basic property of the Kronecker product (see [24, p. 255]), and the third equality is due to the skew-symmetry of $\tilde{R}_i$ (see 135a). The 16×1 vector $\mathrm{vec}(\delta\mathcal{B}_i)$ is expressed as a function of the 3×1 vector $\boldsymbol{\delta}\mathbf{b}_i$ as follows.

$$\begin{aligned}[\mathrm{vec}(\delta\mathcal{B}_i)]^T &= \left\{\mathrm{vec}\left(\begin{bmatrix} 0 & \delta b_3 & -\delta b_2 & \delta b_1 \\ -\delta b_3 & 0 & \delta b_1 & \delta b_2 \\ \delta b_2 & -\delta b_1 & 0 & \delta b_3 \\ -\delta b_1 & -\delta b_2 & -\delta b_3 & 0\end{bmatrix}\right)\right\}^T\\
&= [0 \quad -\delta b_3 \quad \delta b_2 \quad -\delta b_1 \quad \delta b_3 \quad 0 \quad -\delta b_1 \quad -\delta b_2 \quad -\delta b_2 \quad \delta b_1 \quad 0 \quad -\delta b_3 \quad \delta b_1 \quad \delta b_2 \quad \delta b_3 \quad 0]\\
&= [\delta b_1 \quad \delta b_2 \quad \delta b_3]\begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \end{bmatrix}\\
&= [\delta b_1 \quad \delta b_2 \quad \delta b_3][-[\mathbf{e}_1\times] \quad -\mathbf{e}_1 \quad -[\mathbf{e}_2\times] \quad -\mathbf{e}_2 \quad -[\mathbf{e}_3\times] \quad -\mathbf{e}_3 \quad I_3 \quad \mathbf{0}]\\
&= (M\boldsymbol{\delta}\mathbf{b}_i)^T\end{aligned} \tag{137}$$

where the 16×3 matrix M is defined is defined as in (31). Taking the transpose of (137) yields

$$\mathrm{vec}(\delta\mathcal{B}_i) = M\boldsymbol{\delta}\mathbf{b}_i. \tag{138}$$

Using (138) in (136) yields

$$\mathbf{v}_i(t_{k+1}) = -(I_4\otimes\tilde{R}_i)M\boldsymbol{\delta}\mathbf{b}_i. \tag{139}$$

Defining the 16×16 matrix Λ_i as

$$\Lambda_i \triangleq (I_4\otimes\tilde{R}_i)M \tag{140}$$

and using (140) in (139) yields

$$\mathbf{v}_i(t_{k+1}) = -\Lambda_i\boldsymbol{\delta}\mathbf{b}_i. \tag{141}$$

Substituting (141) to $\mathbf{v}_i(t_{k+1})$ in (133) yields

$$\mathbf{v}_{k+1} = -\sum_{i=1}^{m}\alpha_i\Lambda_i\boldsymbol{\delta}\mathbf{b}_i. \tag{142}$$

Since $\mathbf{v}_{k+1}$ is a linear combination of zero-mean white-noise processes (see (41)), it is a zero-mean process itself, and its covariance R_{k+1} is readily computed as

$$R_{k+1} = \sum_{i=1}^{m}\alpha_i^2\Lambda_i R^{\mathbf{b}_i}_{k+1}\Lambda_i^T \tag{143}$$

where $R^{\mathbf{b}_i}_{k+1}$ is the covariance of the measurement error in $\mathbf{b}_i$, $i = 1,2\ldots,m$.

APPENDIX IV. DERIVATION OF THE REDUCED COVARIANCE FILTER

Time Update

PROOF Assume that $P_{k/k} = \bar{P}_{k/k}\otimes I_4$ and $Q_k = \bar{Q}_k\otimes I_4$, then the covariance time update is formulated as (see (48) and (49))

$$\begin{aligned}P_{k+1/k} &= \mathcal{F}_k P_{k/k}\mathcal{F}_k^T + Q_k\\ &= (\Phi_k\otimes\Phi_k)(\bar{P}_{k/k}\otimes I_4)(\Phi_k\otimes\Phi_k)^T + (\bar{Q}_k\otimes I_4)\\ &= (\Phi_k\bar{P}_{k/k}\Phi_k^T\otimes\Phi_k\Phi_k^T) + (\bar{Q}_k\otimes I_4)\\ &= (\Phi_k\bar{P}_{k/k}\Phi_k^T\otimes I_4) + (\bar{Q}_k\otimes I_4)\\ &= (\Phi_k\bar{P}_{k/k}\Phi_k^T + \bar{Q}_k)\otimes I_4.\end{aligned} \tag{144}$$

The second equality in (144) is obtained using (48) and the above assumptions. The third equality is obtained using the mixed product property of the Kronecker product [24, Lemma 4.2.10, p. 243] and [24, Lemma 4.2.4, p. 244]. The fourth equality results from the orthogonal nature of Φ_k. The last equality is obtained using [24, Lemma 4.2.7, p. 243]. Looking at the last equality in (144) we define the 4×4 matrix $\bar{P}_{k+1/k}$ as

$$\bar{P}_{k+1/k} \triangleq \Phi_k \bar{P}_{k+1/k} \Phi_k^T + \bar{Q}_k$$

which yields

$$P_{k+1/k} = \bar{P}_{k+1/k} \otimes I_4.$$

MEASUREMENT UPDATE

PROOF Assume that $P_{k+1/k} = \bar{P}_{k+1/k} \otimes I_4$ and $R_{k+1} = \bar{R}_{k+1} \otimes I_4$, then the innovation covariance matrix is computed as follows (see (51))

$$\begin{aligned} S_{k+1} &= P_{k+1/k} + R_{k+1} \\ &= (\bar{P}_{k+1/k} \otimes I_4)) + (\bar{R}_{k+1} \otimes I_4) \\ &= (\bar{P}_{k+1/k} + R_{k+1}) \otimes I_4. \end{aligned} \tag{145}$$

The third equality in (145) is obtained using [24, Lemma 4.2.7, p. 243]. Let the 4×4 matrix $\bar{S}_{k+1}$ be defined as follows

$$\bar{S}_{k+1} \triangleq \bar{P}_{k+1/k} + R_{k+1}$$

yields

$$S_{k+1} = \bar{S}_{k+1} \otimes I_4.$$

The 16×16 Kalman gain matrix, $\mathcal{K}_{k+1}$, is expressed as given by (52)

$$\begin{aligned} \mathcal{K}_{k+1} &= P_{k+1/k} S_{k+1}^{-1} \\ &= (\bar{P}_{k+1/k} \otimes I_4)(\bar{S}_{k+1} \otimes I_4)^{-1} \\ &= (\bar{P}_{k+1/k} \otimes I_4)[(\bar{S}_{k+1})^{-1} \otimes I_4] \\ &= [\bar{P}_{k+1/k}(\bar{S}_{k+1})^{-1}] \otimes I_4. \end{aligned} \tag{146}$$

Lemma 4.2.5 and Lemma 4.2.10 in [24] are used in order to obtain the third and fourth equalities, respectively, in (146). Then, defining the 4×4 matrix $\bar{K}_{k+1}$ as

$$\bar{K}_{k+1} \triangleq \bar{P}_{k+1/k}(\bar{S}_{k+1})^{-1}$$

yields

$$\mathcal{K}_{k+1} = \bar{K}_{k+1} \otimes I_4. \tag{147}$$

The estimation error covariance measurement update is formulated as in (55)

$$\begin{aligned} P_{k+1/k+1} &= (I_{16} - \mathcal{K}_{k+1}) P_{k+1/k} (I_{16} - \mathcal{K}_{k+1})^T + \mathcal{K}_{k+1} R_{k+1} \mathcal{K}_{k+1}^T \\ &= [(I_4 \otimes I_4) - (\bar{K}_{k+1} \otimes I_4)](\bar{P}_{k+1/k} \otimes I_4) \\ &\quad \times [(I_4 \otimes I_4) - (\bar{K}_{k+1} \otimes I_4)]^T \\ &\quad + (\bar{K}_{k+1} \otimes I_4)(\bar{R}_{k+1} \otimes I_4)(\bar{K}_{k+1} \otimes I_4)^T \\ &= [(I_4 - \bar{K}_{k+1}) \otimes I_4](\bar{P}_{k+1/k} \otimes I_4)[(I_4 - \bar{K}_{k+1}) \otimes I_4]^T \\ &\quad + (\bar{K}_{k+1} \bar{R}_{k+1} \bar{K}_{k+1}^T \otimes I_4) \\ &= [(I_4 - \bar{K}_{k+1}) \bar{P}_{k+1/k} (I_4 - \bar{K}_{k+1})^T + \bar{K}_{k+1} \bar{R}_{k+1} \bar{K}_{k+1}^T \otimes I_4]. \end{aligned} \tag{148}$$

Thus, defining the 4×4 matrix $\bar{P}_{k+1/k+1}$ as

$$\bar{P}_{k+1/k+1} \triangleq (I_4 - \bar{K}_{k+1}) \bar{P}_{k+1/k} (I_4 - \bar{K}_{k+1})^T + \bar{K}_{k+1} \bar{R}_{k+1} \bar{K}_{k+1}^T$$

yields

$$P_{k+1/k+1} = \bar{P}_{k+1/k+1} \otimes I_4.$$

The estimate measurement update, as given in (53), is rewritten here

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{4} \sum_{l=1}^{4} \mathcal{K}_{k+1}^{jl} \tilde{Y}_{k+1} E^{lj} \tag{149}$$

where each 4×4 block-matrix $\mathcal{K}_{k+1}^{jl}$ is, in view of (147), a scalar matrix of the form

$$\mathcal{K}_{k+1}^{jl} = \bar{K}_{k+1}[j,l] I_4 \tag{150}$$

and $\bar{K}_{k+1}[j,l]$ denotes the element $[j,l]$ in $\bar{K}_{k+1}$. Using (150) in (149) yields

$$\begin{aligned} \hat{X}_{k+1/k+1} &= \hat{X}_{k+1/k} + \sum_{j=1}^{4} \sum_{l=1}^{4} \tilde{Y}_{k+1} \bar{K}_{k+1}[j,l] E^{lj} \\ &= \hat{X}_{k+1/k} + \tilde{Y}_{k+1} \left(\sum_{j=1}^{4} \sum_{l=1}^{4} \bar{K}_{k+1}[j,l] E^{lj} \right) \\ &= \hat{X}_{k+1/k} + \tilde{Y}_{k+1} \bar{K}_{k+1}^T. \end{aligned} \tag{151}$$

The first equality in (151) is obtained using the fact that $\bar{K}_{k+1}[j,l]$ is a scalar. In the second equality we use the fact that $\tilde{Y}_{k+1}$ is independent from the summing indices j and l. The third equality comes from the canonical decomposition of a matrix in $\mathbb{R}^{4\times4}$.

REFERENCES

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

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

[3] Shuster, M. D. and Oh, S. D.
Three-axis attitude determination from vector observations.
Journal of Guidance and Control, **4**, 1 (1981), 70–77.

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

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

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

[7] Shuster, M. D.
A simple Kalman filter and smoother for spacecraft attitude.
Journal of the Astronautical Sciences, **37**, 1 (Jan.–Mar. 1989), 89–106.

[8] 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.

[9] Choukroun, D., Oshman, Y., and Bar-Itzhack, I. Y.
Optimal-REQUEST algorithm for attitude determination.
Journal of Guidance, Control, and Dynamics, **27**, 3 (May–June 2004), 418–425.

[10] Mendel, J. M.
Lessons in Estimation Theory for Signal Processing, Communications, and Control.
Upper Saddle River, NJ: Prentice-Hall, 1988.

[11] 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.

[12] 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.

[13] Choukroun, D., et al.
Kalman filtering for matrix estimation.
IEEE Transactions on Aerospace and Electronic Systems, **42**, 1 (Jan. 2006), 147–159.

[14] Gustafson, D. and Speyer, J. L.
Linear minimum variance filters applied to carrier tracking.
IEEE Transactions on Automatic Control, **AC-21**, 1 (Feb. 1976), 65–73.

[15] Gershon, E., Shaked, U., and Yaesh, I.
H_∞ Control and Estimation of State-multiplicative Linear Systems.
LNCIS–Lecture Notes in Control and Information Sciences, vol. 318. New York: Springer, 2005.

[16] Choukroun, D., Oshman, Y., and Bar-Itzhack, I. Y.
A novel quaternion Kalman filter.
IEEE Transactions on Aerospace and Electronic Systems, **42**, 1 (Jan. 2006), 174–190.

[17] 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.

[18] Choukroun, D., et al.
Direction cosine matrix estimation from vector observations using a matrix Kalman filter.
IEEE Transactions on Aerospace and Electronic Systems, **46**, 1 (Jan. 2010), 61–79.

[19] Julier, S. J. and LaViola, Jr., J. J.
On Kalman filtering with nonlinear equality constraints.
IEEE Transactions on Signal Processing, **55**, 6, Part 2 (June 2007), 2774–2784.

[20] http://ltpwww.gsfc.nasa.gov/IAS/handbook/handbook htmls/chapter2/.

[21] http://sunland.gsfc.nasa.gov/smex/sampex/.

[22] Beasley, D., Neal, N., and Diller, G.
Microwave Anisotropy Probe (MAP) Mission.
NASA Press Kit, June 2001. www.gsfc.nasa.gov/gsfc/spacesci/pictures/2001/0612map/map-press-kit.pdf.

[23] Henderson, H. V. and Searle, S. R.
The vec-permutation matrix, the vec-operator and Kronecker products: A review.
Linear Multilinear Algebra, **9** (1981), 271–288.

[24] Horn, R. A. and Johnson, C. R.
Topics in Matrix Analysis.
New York: Cambridge University Press, 1991.

[25] Bar-Itzhack, I. Y.
On matrix symmetrization.
Journal of Guidance, Control, and Dynamics, **21**, 1 (Jan.–Feb. 1998), 178–179.

[26] Choukroun, D.
Novel methods for attitude determination using vector observations.
Ph.D. Thesis, Dept. Aerospace Engineering, Technion–Israel Institute of Technology, Haifa, Israel, May 2003.

[27] Ronen, M. and Oshman, Y.
Third-order, minimal-parameter solution of the orthogonal matrix differential equation.
Journal of Guidance, Control, and Dynamics, **20**, 3 (May–June 1997), 516–521.

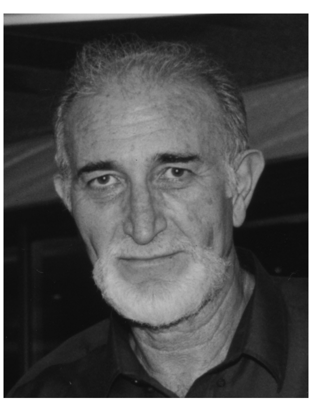
Daniel Choukroun (M'06) received the B.Sc. (summa cum laude), M.Sc., and Ph.D. degrees in 1997, 2000, and 2003, respectively, from the Technion–Israel Institute of Technology, Faculty of Aerospace Engineering. He also received the title Engineer under Instruction from the Ecole Nationale de l'Aviation Civile, France, in 1994, and a Master in Space Systems Engineering from Delft University of Technology, Netherlands, in 2011.

From 1998 to 2003, he was a teaching and research assistant in the field of automatic control at the Technion–Israel Institute of Technology. From 2003 to 2006 he was a postdoctoral fellow and a lecturer at the University of California at Los Angeles in the Department of Mechanical and Aerospace Engineering. From 2006 to 2010 he was with the Department of Mechanical Engineering, Ben-Gurion University of the Negev, Israel. In 2010, he joined the Faculty of Aerospace Engineering at Delft University of Technology, The Netherlands, where he is currently assistant professor in Space Systems Engineering. He was invited lecturer at MAST (Master in Aerospace Science and Technology), Universitat Politecnica de Catalunya, Barcelona, Spain (2009), and visiting assistant professor at Université Paul Sabatier and LAAS/CNRS–Laboratoire d'Architecture et d'Analyse de Systèmes, Toulouse, France (2010). 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. He is a member of the AIAA Technical Committee on Guidance Navigation and Control. He serves as technical editor for guidance and control systems for the *IEEE Transactions on Aerospace and Electronic Systems.*

Haim Weiss (M'02) received his B.Sc. and M.Sc. degrees in electrical engineering, in 1966 and 1971, respectively, both from Technion–Israel Institute of Technology. In 1979 he received the Ph.D. degree from the University of Newcastle, Australia, also in electrical engineering.

From 1966 to 1970 he served as an electronic engineer in the Israeli Air Force. In 1972 he joined RAFAEL where he is now a research fellow. From 1999–2001 he was a Lady Davis visiting scientist at the Faculty of Aerospace Engineering, Technion, Israel. Currently he holds the position of adjunct senior teaching fellow. His research interests include estimation, guidance, spacecraft attitude determination and control, and aided inertial navigation systems.



Itzhack Y. Bar-Itzhack (M'73—SM'84—F'95) (deceased 2007) 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 Technion where he was a Sophie and William Shambam Professor of Aeronautical Engineering, and a member of the Technion Asher Space Research Institute. During the 1977–1978 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 1987–1989 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. He was back at NASA–Goddard in the 1995–1996, 2000–2001 and 2005–2006 academic years, again as a national research council research associate working on attitude, angular velocity, and orbit determination problems.

Bar-Itzhack 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 was an IEEE Aerospace and Electronic Systems Society Distinguished Lecturer in the distinguished lectures series. He was 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. He was a member of Sigma Xi, and AIAA Fellow.

Yaakov Oshman (A'96—SM'97—F'07) 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, 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 presently a professor, and holder of the Louis and Helen Rogow Chair in Aeronautical Engineering. He is a member of the Technion's Asher Space Research Institute. He headed the Technion's Philadelphia Flight Control Laboratory (1993–1996). During the 1996/7 and 1997/8 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 Advanced Defense Systems Ltd., Israeli Ministry of Defense, and Israel Aerospace Industries (MBT Missile Division). His research has been supported by the Israeli Aircraft Industries, the Israeli Ministry of Defense, the U.S. Air Force Office of Scientific Research (AFOSR), RAFAEL Advanced Defense Systems Ltd., and the Israeli Science Foundation (ISF), as well as by various Technion research grants. His research interests are in advanced estimation, tracking, information fusion and control methods and their applications in aerospace guidance, navigation, and control systems, including structural estimation and control, and health monitoring/fault detection and isolation (FDI) systems.

Dr. Oshman served as the President of the Israeli Association for Automatic Control, a national member organization of the International Federation of Automatic Control (IFAC), between 2003–2008. He was a member of the national board of the Israeli Society of Aeronautics and Astronautics between 2004–2009. He was a member of the AIAA Guidance, Navigation, and Control Technical Committee between 2002–2008, and an International Advisor (member of the editorial board) of the *AIAA Journal of Guidance, Control and Dynamics* between 2002–2009. He has published over 150 journal and conference papers and book chapters, and numerous technical reports in these areas. 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 was awarded the Best Paper Award of the 2004 AIAA Guidance, Navigation and Control Conference. He 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 awarded multiple Technion awards for outstanding teaching. He has been on the program committees of over a dozen international conferences. He serves as technical editor for guidance and control systems for the *IEEE Transactions on Aerospace and Electronic Systems* between 2005–2011. He was a member of the Board of Governors of the IEEE Aerospace and Electronic Systems Society between 2008–2010, and serves as Chair of the Israeli Chapter of IEEE/AESS since 2010. He is a Fellow of the AIAA.