Direction Cosine Matrix Estimation from Vector Observations using a Matrix Kalman Filter

D. CHOUKROUN, Member, IEEE
Ben-Gurion University

H. WEISS
Rafael Advanced Defense Systems, Ltd.

I. Y. BAR-ITZHACK, Fellow, IEEE

Y. OSHMAN, Fellow, IEEE
Technion–Israel Institute of Technology

This work presents several algorithms that use vector observations in order to estimate the direction cosine matrix (DCM) as well as three constant biases and three time-varying drifts in body-mounted gyro output errors. All the algorithms use the matrix Kalman filter (MKF) paradigm, which preserves the natural formulation of the DCM state-space model equations. Focusing on the DCM estimation problem, the assumption of white noise in the gyro and in the vector observations errors yields reduced and efficient filter covariance computations. The orthogonality constraint on the DCM is handled via the technique of pseudomeasurement, which is naturally embedded in the MKF. Two additional known "brute-force" procedures are implemented for the sake of comparison. Extensive Monte-Carlo simulations illustrate the performances of the different estimators. When estimating only the DCM, it is shown that all the proposed orthogonalization procedures accelerate the estimation convergence. Nevertheless, the pseudomeasurement technique shows a smoother and shorter transient than the brute-force procedures, which on the other hand yield more accurate steady-states. The reduced covariance computations yield a more accurate steady-state than the full covariance computations but show a slower transient. When estimating the DCM as well as the gyro biases and drifts, enforcing orthogonalization seems to penalize the DCM estimation as long as the biases are not correctly identified. For the sake of computation savings during long duration missions, a mixed estimator, switching between long periods of DCM-only estimation and short periods of DCM-biases estimation, appears to be a promising strategy.

Manuscript received December 12, 2006; revised October 15, 2007; released for publication August 20, 2008.

IEEE Log No. T-AES/46/1/935928.

Refereeing of this contribution was handled by P. Willett.

This work was presented as Paper 2003-5482 at the 43rd AIAA Guidance, Navigation, and Control Conference, Austin, TX, Aug. 2003.

This paper is dedicated to the memory of Itzhack Y. Bar-Itzhack, Professor Emeritus of Aerospace Engineering at the Technion–Israel Institute of Technology, who died 9 May 2007 in Haifa, Israel.

Authors' addresses: D. Choukroun, Dept. of Mechanical Engineering, Ben-Gurion University of the Negev, Israel, E-mail: (danielch@bgu.ac.il); H. Weiss, Rafael Advanced Defense Systems, Ltd., PO Box 2250, Dept. 35, Haifa 31021, Israel; Y. Oshman, Dept. of Aerospace Engineering, Technion–Israel Institute of Technology, Israel.

INTRODUCTION

In the course of spacecraft operation the spacecraft spatial orientation with respect to some reference frame has to be continuously determined. A popular mathematical representation of the spacecraft orientation is the matrix of the rotation between a Cartesian coordinate frame, which is rigidly fixed in the spacecraft body, and a reference Cartesian coordinate frame. This matrix, which is denoted by D, is also known as direction cosine matrix (DCM) [1, p. 411]. Typical observations that are processed on-board spacecraft in order to determine D are vector observations, like the Earth's magnetic field or measurements of lines of sight from the spacecraft to various celestial objects. The DCM is a 3×3 matrix, whose nine elements are not independent. Indeed, like any rotation matrix, D is a proper orthogonal matrix [1, p. 412], which implies six constraints connecting the nine elements. In spite of the redundancy created by the constraints, the DCM D is a convenient rotation representation. This is so because the equations that describe the vector measurement model and the spacecraft kinematical model are linear in D [1, pp. 411, 512]. The latter facts enable the development of efficient estimators of the DCM using vector observations. Notice that two vector observations are necessary and sufficient to fully determine D [2].

Optimal DCM estimators which handle more than two vector observations typically fall into two categories. The first one has its origin in a constrained weighted least-squares problem, known as the Wahba problem [3]. Besides the usual good features of the solutions to Wahba's problem[1] a particular highlight is that they are matrix estimators: they preserve the matrix formulation of the original DCM parameters; their development and analysis, which involve standard matrix decompositions, would have been cumbersome, if not impossible, without maintaining a matrix notation. Nevertheless, a drawback of this family of estimators is the suboptimality with regard to time propagation noises, when estimating a time-varying DCM, and a lack of clear probabilistic meaning.[2] On the other hand, the second approach, namely the minimum-variance or Kalman filtering approach, provides a convenient stochastic framework for developing approximate Kalman filters (KF) of a time-varying DCM. But in order to implement the KF, the usual approach [11] consists in transforming the

[1] They are closed-form solutions, and thus, they need no a priori estimate of D [4–10]. Moreover the orthogonality constraint on D is inherent to the formulation of the problem.

[2] In general, the weights in Wahba's cost function are chosen as scalars equal to the inverse of the variance of the associated vector measurement. Through this choice, the solution to Wahba's problem acquires some statistical ground. Nevertheless, this choice is at most heuristic.

DCM into its vectorized representation, by stacking its columns one under the other. This is done in order to comply with the conventional state-vector formulation of model equations. Following that approach all system equations are vectorized. As a result, the physical insight in the plant equations is lost, rendering the analysis of the related KF cumbersome.

The issue of developing optimal algorithms for estimation and control that operate on matrix systems, while preserving the matrix formulation of the original systems has been given much attention [12–14]. The various solutions to Wahba's problem are such examples. When modeling structural systems, physical parameters are represented as matrices like the stiffness matrix, the inertia matrix or the damping matrix. Reference [12] features a matrix estimator of the stiffness matrix. Matrix calculus tools, such as the gradient matrix [13], were developed in order to design optimal control algorithms that operate on matrix processes. Recently, a generic matrix Kalman filter (MKF) was introduced [14] that operates on plants which are naturally described by matrix equations. Notice that, beyond the notational advantage, using an MKF instead of the associated vectorized KF allows savings in computations (see [14, Table V, Cases A and C for DCM estimation]).

This work is concerned with the development, via the MKF approach, of estimators of the DCM and of additional parameters such as constant biases and time-varying drifts in body-mounted gyro output errors. The matrix formalism adopted throughout this work allows straightforward developments of expressions for the filter noise covariance matrices. This in turn is helpful in designing simplified filters based on specific stochastic assumptions. Focusing on DCM estimation and assuming gyro output, white noise leads to filter covariance computations that can be expressed by reduced 3×3 equations instead of the full 9×9 equations. The more realistic case of constant biases and time-varying drifts in the gyro errors is handled via a state augmentation and the development of the associated matrix state-space model. The traditional issue of orthogonalization of the DCM estimate, which is crucial when it is needed for axis rotation, e.g. in inertial navigation systems, is addressed via the technique of orthogonality pseudomeasurement (OPM). The orthogonality constraint being a quadratic matrix equation is, upon some modeling transformations, recast as a virtual matrix measurement equation of the DCM. Then, this additional matrix measurement yields an additional measurement update stage naturally embedded in the formalism of the MKF. As opposed to other orthogonalization techniques [11, 17], the OPM technique provides the designer with a covariance computation that is associated with the orthogonalization procedure and, furthermore, allows him to tune that constraint. Extensive Monte-Carlo simulations were performed in order to illustrate the performances of the various filters. In particular, the relative advantages of four various orthogonalization schemes are studied, the consequence of using reduced covariance computations is investigated, and the performance of the augmented filter for DCM bias and drift estimation are shown. Furthermore, for long span simulations, a special augmented filter is tested, where the bias and drift estimation processes are turned off from time to time, for filter computational savings, while the overall DCM estimation accuracy stays at an acceptable level.

The remainder of the paper is organized as follows. The next section presents the mathematical model of the DCM system in the case of white noise in the gyro output error. This model is already known in the literature, except for the explicit expression describing the process noise covariance matrix, which is presented here. Two MKFs of the DCM are developed in the following section. The first filter includes a full (9×9) covariance computation algorithm. The second one includes a reduced (3×3) covariance computation. Both filters do not include any orthogonalization procedure. The subsequent section presents the development of the augmented state mathematical model, including constant biases and time-varying drifts in the gyro output error. Four orthogonalization procedures are proposed in the following section. Then, the performance of the various estimators are demonstrated and compared via extensive Monte-Carlo simulations. Finally, conclusions are drawn in the last section.

DCM STATE-SPACE MODEL

Process Model

Let the spacecraft body frame and the reference frame be denoted, respectively, by $\mathcal{B}$ and $\mathcal{R}$. Assuming that $\mathcal{B}$ is rotating with respect to $\mathcal{R}$, we denote by $\boldsymbol{\omega}^o$ the angular velocity vector of this rotation as expressed in $\mathcal{B}$. It is well known that the dynamics of the DCM, D, is governed by the following matrix differential equation [1, p. 512]

$$\frac{d}{dt}D = -[\boldsymbol{\omega}^o\times]D. \tag{1}$$

The matrix $[\boldsymbol{\omega}^o\times]$ in (1) is a 3×3 skew-symmetric matrix and is defined according to the identity $[\boldsymbol{\omega}^o\times]\mathbf{x} = \boldsymbol{\omega}^o \times \mathbf{x}$, where $\times$ denotes the cross-product and $\mathbf{x}$ is any 3×1 column-matrix. The discrete-time version of (1) is the difference equation given by

$$D_{k+1} = \Phi_k^o D_k \tag{2}$$

where Φ_k^o is the transition matrix from time t_k to time t_{k+1}. Taking the time increment $\Delta t = t_{k+1} - t_k$ small enough, we assume that $\boldsymbol{\omega}^o$ is piecewise constant in the intervals of time length Δt, so that Φ_k^o in (2) can

be approximated as

$$\Phi_k^o \simeq e^{-[\omega_k^o\times]\Delta t}. \tag{3}$$

The true angular velocity ω_k^o being unknown, we assume here that a triad of body-mounted gyroscopes measures ω_k^o and that the output of the gyros ω_k is corrupted by an additive error ϵ_k, thus

$$\omega_k = \omega_k^o + \epsilon_k. \tag{4}$$

Then we substitute ω_k for ω_k^o in (3) and use the measured transition matrix, denoted by Φ_k, in (2). Considering the difference equation (2) and using a first-order approximation in the time increment Δt of the Taylor series expansion of Φ_k^o [1, p. 512], yields the following process equation:

$$\begin{aligned} D_{k+1} &= \Phi_k^o D_k \\ &= (I_3 - [\omega_k^o\times]\Delta t)D_k + \text{HOT} \\ &= (I_3 - [\omega_k\times]\Delta t + [\epsilon_k\times]\Delta t)D_k + \text{HOT} \\ &= (I_3 - [\omega_k\times]\Delta t)D_k + ([\epsilon_k\times]\Delta t)D_k + \text{HOT} \\ &= e^{(-[\omega_k\times]\Delta t)}D_k + [\epsilon_k\times]D_k\Delta t + \text{HOT} \\ &= \Phi_k D_k + [\epsilon_k\times]D_k\Delta t + \text{HOT} \\ &= \Phi_k D_k + W_k \end{aligned} \tag{5}$$

where HOT denotes higher order terms (of order Δt^2 and higher). In (5), the third equality results from (4). The definition of Φ_k and W_k is obvious. To summarize,

$$D_{k+1} = \Phi_k D_k + W_k \tag{6}$$

where

$$\Phi_k = e^{(-[\omega_k\times]\Delta t)} \tag{7}$$

$$W_k = [\epsilon_k\times]D_k\Delta t + \text{HOT}. \tag{8}$$

The gyro error ϵ_k is here assumed to be a zero-mean white sequence with known covariance matrix $Q_k^\epsilon/\Delta t$, where the factor $1/\Delta t$ is consistent with the assumption that ϵ_k approximates a continuous-time white noise process. The more realistic case that includes constant biases and time-varying drifts in the gyro output error will be handled in a subsequent section. Notice from (8) that the process noise matrix W_k is state dependent, and that the first-order term is linear in the components of ϵ_k. The latter fact will help at developing an expression for the process noise covariance matrix. The process model for the DCM is described by (6)–(8). It can be seen that these equations feature, in their natural form, a state matrix D_k and a process noise matrix W_k.

Measurement Model

Assume that a physical vector quantity is observed at each epoch time t_k; namely, we simultaneously know $\mathbf{r}_k$, its decomposition in $\mathcal{R}$, and $\mathbf{b}_k$, its measured decomposition in $\mathcal{B}$. Moreover, without loss of generality, we assume that the observed physical vector is of unit length. In general, $\mathbf{r}_k$ is accurately known from tables or almanacs while $\mathbf{b}_k$ is a noisy measurement acquired on-board by sensing devices. Modeling the measurement noise as an additive error term $\mathbf{v}_k$ yields the classical vector measurement equation

$$\mathbf{b}_k = D_k\mathbf{r}_k + \mathbf{v}_k \tag{9}$$

where $\mathbf{v}_k$ is assumed to be a zero-mean white noise sequence with known covariance matrix R_k. Furthermore, we assume that ϵ_k and $\mathbf{v}_k$ are uncorrelated with one another and with the initial attitude matrix D_0. Equation (9) is a linear measurement equation with respect to the state matrix D_k. Combining (6)–(9) yields a matrix state-space model for D_k, on which the MKF can be applied. For the sake of completeness, the generic MKF [14] is reviewed in Appendix A.

DCM FILTER DESIGN

In this section we present two MKFs for the estimation of the DCM. These filters differ by the computational complexity of the covariance computations. The first algorithm performs a full covariance computation, where the estimation error covariance matrix is of size 9×9 while the second algorithm performs a reduced covariance computation, where only a 3×3 matrix is used in order to represent the whole estimation error covariance. For the sake of clarity all the matrices that are involved in the full covariance computation are denoted by capital script letters, like $\mathcal{P}$. We use normal fonts, like P, to denote the matrices involved in the reduced covariance computation.

Full Covariance Filter

In order to obtain an approximate expression for the 9×9 covariance matrix of W_k, which is denoted by $\mathcal{Q}_k^\epsilon$, we neglect the HOT and substitute the best available estimate, $\hat{D}_{k/k}$, for D_k in (8). This yields

$$\begin{aligned} \mathcal{Q}_k^\epsilon &\triangleq \text{cov}\{W_k\} \\ &\triangleq \text{cov}\{\text{vec}\, W_k\} \\ &= \text{cov}\{(\hat{D}_{k/k}^T \otimes I_3)(\text{vec}[\epsilon_k\times])\Delta t\} \\ &= \text{cov}\{(\hat{D}_{k/k}^T \otimes I_3)(\mathcal{L}\epsilon_k)\Delta t\} \\ &= (\hat{D}_{k/k}^T \otimes I_3)\mathcal{L}\,\text{cov}\{\epsilon_k\}\mathcal{L}^T(\hat{D}_{k/k} \otimes I_3)\Delta t^2 \end{aligned} \tag{10}$$

where $\otimes$ denotes the Kronecker product [15, p. 227], I_3 is the three-dimensional identity matrix, $\mathcal{L}$ is a 9×3 matrix defined as

$$\mathcal{L}^T \triangleq [[\mathbf{e}_1\times] \quad [\mathbf{e}_2\times] \quad [\mathbf{e}_3\times]] \tag{11}$$

and $\mathbf{e}_j$, $j = 1,2,3$, is the jth column of I_3. The third equality in (10) is obtained using [16, Lemma 4.3.1, p. 254], and the fourth equality stems from the definition of the cross-product matrix $[\epsilon\times]$. Recalling that the covariance matrix of ϵ_k is given by $Q_k^\epsilon/\Delta t$ yields

$$\mathcal{Q}_k^\epsilon = (\hat{D}_{k/k}^T \otimes I_3)\mathcal{L}Q_k^\epsilon\mathcal{L}^T(\hat{D}_{k/k}\otimes I_3)\Delta t. \tag{12}$$

The full covariance MKF is summarized as follows. Given the matrices $\hat{D}_{0/0}$ and $\mathcal{P}_{0/0}$, compute:

1) Time Update Equations:

$$\Phi_k = e^{(-[\omega_k\times]\Delta t)} \tag{13}$$

$$\hat{D}_{k+1/k} = \Phi_k\hat{D}_{k/k} \tag{14}$$

$$\Psi_k = I_3\otimes\Phi_k \tag{15}$$

$$\mathcal{P}_{k+1/k} = \Psi_k\mathcal{P}_{k/k}\Psi_k^T + \mathcal{Q}_k^\epsilon. \tag{16}$$

2) Measurement Update Equations:

$$\mathcal{H}_{k+1} = \mathbf{r}_{k+1}^T\otimes I_3 \tag{17}$$

$$\mathcal{S}_{k+1} = \mathcal{H}_{k+1}\mathcal{P}_{k+1/k}\mathcal{H}_{k+1}^T + R_{k+1} \tag{18}$$

$$\mathcal{K}_{k+1} = \mathcal{P}_{k+1/k}\mathcal{H}_{k+1}^T\mathcal{S}_{k+1}^{-1} \tag{19}$$

$$\hat{D}_{k+1/k+1} = \hat{D}_{k+1/k} + [\mathcal{K}_{k+1}^1 \quad \mathcal{K}_{k+1}^2 \quad \mathcal{K}_{k+1}^3] \times [I_3\otimes(\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})] \tag{20}$$

$$\mathcal{P}_{k+1/k+1} = (I_9 - \mathcal{K}_{k+1}\mathcal{H}_{k+1})\mathcal{P}_{k+1/k} \times (I_9 - \mathcal{K}_{k+1}\mathcal{H}_{k+1})^T + \mathcal{K}_{k+1}R_{k+1}\mathcal{K}_{k+1}^T \tag{21}$$

where $\mathcal{K}_{k+1}^j$ in (20), $j = 1,2,3$, are 3×3 submatrices of the 9×3 Kalman gain matrix, $\mathcal{K}_{k+1}$, such that $\mathcal{K}^T = [(\mathcal{K}^1)^T(\mathcal{K}^2)^T(\mathcal{K}^3)^T]$. Note from (14) and (20) that the full covariance filter, described by (14) to (21), produces a Kalman filter estimate of the state DCM $\hat{D}_{k/k}$ using the original matrices of the given plant.

Reduced Covariance Filter

Our purpose is here two-fold: it is, first, to achieve a reduction in the computational burden of the filter and, second, to illustrate how the MKF formulation facilitates mathematical manipulations. We assume here that the rows of the process noise matrix W_k are uncorrelated and that they have the same 3×3 covariance matrix Q_k where

$$Q_k = Q_k^\epsilon\Delta t. \tag{22}$$

Therefore, the following expression for $\mathcal{Q}_k^\epsilon$ is assumed:

$$\mathcal{Q}_k^\epsilon = Q_k\otimes I_3. \tag{23}$$

We also assume that the elements of the measurement noise column-matrix $\mathbf{v}_k$ are uncorrelated and that they have the same covariance μ_k, that is,

$$R_{k+1} = \mu_{k+1}I_3. \tag{24}$$

Similarly, we assume that $\mathcal{P}_{0/0}$ is given as follows:

$$\mathcal{P}_{0/0} = P_{0/0}\otimes I_3 \tag{25}$$

where $P_{0/0}$ is known. Then, using basic properties of the Kronecker product, one can show that the nine-dimensional covariance equations can be reduced to three-dimensional equations (see Appendix B). In spite of these strong model simplifications, the reduced estimator performs well as demonstrated through extensive Monte-Carlo simulations. The reduced covariance MKF is summarized as follows. Given the initial estimate $\hat{D}_{0/0}$, and the 3×3 matrix $P_{0/0}$, compute

1) Time Update Equations:

$$\Phi_k = e^{(-[\omega_k\times]\Delta t)} \tag{26}$$

$$\hat{D}_{k+1/k} = \Phi_k\hat{D}_{k/k} \tag{27}$$

$$P_{k+1/k} = P_{k/k} + Q_k. \tag{28}$$

2) Measurement Update Equations:

$$s_{k+1} = \mathbf{r}_{k+1}^T P_{k+1/k}\mathbf{r}_{k+1} + \mu_{k+1} \tag{29}$$

$$\mathbf{g}_{k+1} = P_{k+1}\mathbf{r}_{k+1}/s_{k+1} \tag{30}$$

$$\hat{D}_{k+1/k+1} = \hat{D}_{k+1/k} + (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})\mathbf{g}_{k+1}^T \tag{31}$$

$$P_{k+1/k+1} = (I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)P_{k+1/k}(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)^T + \mu_{k+1}\mathbf{g}_{k+1}\mathbf{g}_{k+1}^T. \tag{32}$$

Notice that the dynamics matrix Φ_k does not appear in the covariance time propagation (28). That is a direct result of the assumed statistical independence among the rows of W_k and of the orthogonality of Φ_k (see Appendix B). The consequence of the assumptions on the noises is that the three rows of the state matrix are estimated based on identical covariance equations, which are of dimension 3. Indeed, all the necessary information is conveniently represented and processed by 3×3 equations. If the full-scale matrices are needed, they are readily computed from the reduced associated matrices as follows (see Appendix B).

$$\mathcal{P}_{k+1/k} = P_{k+1/k}\otimes I_3 \tag{33}$$

$$\mathcal{P}_{k+1/k+1} = P_{k+1/k+1}\otimes I_3 \tag{34}$$

$$\mathcal{S}_{k+1} = s_{k+1}I_3 \tag{35}$$

$$\mathcal{K}_{k+1} = \mathbf{g}_{k+1}\otimes I_3. \tag{36}$$

One realizes that the filter computational burden is thus reduced by a factor of 27 (3^3). Recall that the computational burden of a conventional n-dimensional Kalman filter is dictated by its

covariance time-propagation stage, which is in $\mathcal{O}(n^3)$. In addition and independently of the latter, let us recall here that the implementation of the general MKF is computationally less intensive than that of the corresponding vectorized Kalman filter. As shown in [14], the computational advantage of an MKF resides in the time-propagation stage and was evaluated to be of factor 37 for DCM estimation using vector observations and orthogonalization, where that factor was defined as the ratio between the numbers of FLOP (floating point operation) per cycle.

AUGMENTED MATRIX KALMAN FILTER

Augmented State Matrix Model: In this section we show how to handle the case of constant biases and Markov drifts in the gyro outputs using an MKF. It is assumed that the gyro output is described by

$$\boldsymbol{\omega}_k = \boldsymbol{\omega}_k^o + \mathbf{c}_k + \mathbf{m}_k + \boldsymbol{\epsilon}_k \tag{37}$$

where $\boldsymbol{\omega}_k^o$ is the true angular velocity vector, $\mathbf{c}_k$ is the 3×1 vector of constant biases, $\mathbf{m}_k$ is the 3×1 vector of the gyro Markov drifts, and $\boldsymbol{\epsilon}_k$ is assumed to be a zero-mean white noise sequence with covariance matrix $Q_k^{\epsilon}/\Delta t$. In order to estimate $\mathbf{c}_k$ and $\mathbf{m}_k$ via a KF we model their dynamics as follows

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \boldsymbol{\mu}_k \tag{38}$$

$$\mathbf{m}_{k+1} = \Lambda \mathbf{m}_k + \boldsymbol{\nu}_k \tag{39}$$

where

$$\Lambda \triangleq e^{(-T\Delta t)} \tag{40}$$

and

$$T \triangleq \mathrm{diag}\{1/\tau_x, 1/\tau_y, 1/\tau_z\}. \tag{41}$$

The scalars $1/\tau_x$, $1/\tau_y$, and $1/\tau_z$ are known constants. In (38) and (39), $\boldsymbol{\mu}_k$ and $\boldsymbol{\nu}_k$ are assumed to be zero-mean white noise sequences with covariance matrices $Q_k^{\mu}/\Delta t$ and $Q_k^{\nu}/\Delta t$, respectively. The addition of $\boldsymbol{\mu}_k$ is needed for filter stability reasons: it avoids computations involving a singular process noise covariance matrix. As it is common practice when designing KFs, the values of the process noise covariance matrices, Q_k^{ϵ}, Q_k^{μ} and Q_k^{ν} are used as tuning parameters with the usual trade-off: increasing the filter process noise covariance speeds up the estimation convergence but decreases the steady-state estimation accuracy. In order to develop the DCM process equation we only have to substitute the expression $(-\boldsymbol{\omega}_k + \mathbf{c}_k + \mathbf{m}_k)$ for $-\boldsymbol{\omega}_k$ in (7). This yields

$$D_{k+1} = \Phi_k D_k + \mathcal{E}_k \tag{42}$$

where the matrices Φ_k and $\mathcal{E}_k$ are here defined as

$$\Phi_k \triangleq e^{\{[(-\boldsymbol{\omega}_k + \mathbf{c}_k + \mathbf{m}_k)\times]\Delta t\}} \tag{43}$$

$$\mathcal{E}_k \triangleq [\boldsymbol{\epsilon}_k \times] D_k \Delta t. \tag{44}$$

To complete the state-space model we recall the vector measurement equation (9)

$$\mathbf{b}_{k+1} = D_{k+1}\mathbf{r}_{k+1} + \mathbf{v}_{k+1} \tag{45}$$

where $\mathbf{v}_{k+1}$ is zero-mean white noise sequence with covariance matrix R_{k+1}. It is assumed that the sequences $\boldsymbol{\epsilon}_k$, $\boldsymbol{\mu}_k$, $\boldsymbol{\nu}_k$, and $\mathbf{v}_k$ are uncorrelated with one another and with the initial state variables D_0, $\mathbf{c}_0$, and $\mathbf{m}_0$. We wish to estimate the DCM, and the gyro output parameters $\mathbf{c}_k$ and $\mathbf{m}_k$ via an MKF. To meet this end, we generalize the known technique of state augmentation to the current matrix state-space model (38)–(45).

PROPOSITION 1 *Let X_k denote the augmented 3×5 state matrix that is defined as follows:*

$$X_k \triangleq [D_k \quad \mathbf{c}_k \quad \mathbf{m}_k]. \tag{46}$$

The dynamics model equations of D_k, $\mathbf{c}_k$, and $\mathbf{m}_k$, (38), (39), and (42) can be written as the following augmented 3×5 matrix difference equation

$$X_{k+1} = \sum_{r=1}^{9} \Theta_k^r X_k \Psi_k^r + W_k \tag{47}$$

where

$$\Theta_k^1 \triangleq e^{(-[\boldsymbol{\omega}_k \times]\Delta t)}, \qquad \Psi_k^1 \triangleq E^{11} + E^{22} + E^{33} \tag{48a}$$

$$\Theta_k^2 \triangleq I_3, \qquad \Psi_k^2 \triangleq E^{44} \tag{48b}$$

$$\Theta_k^3 \triangleq -[\mathbf{d}_{k,1} \times]\Delta t, \qquad \Psi_k^3 \triangleq E^{41} \tag{48c}$$

$$\Theta_k^4 \triangleq -[\mathbf{d}_{k,2} \times]\Delta t, \qquad \Psi_k^4 \triangleq E^{42} \tag{48d}$$

$$\Theta_k^5 \triangleq -[\mathbf{d}_{k,3} \times]\Delta t, \qquad \Psi_k^3 \triangleq E^{43} \tag{48e}$$

$$\Theta_k^6 \triangleq \Lambda, \qquad \Psi_k^6 \triangleq E^{55} \tag{48f}$$

$$\Theta_k^7 \triangleq -[\mathbf{d}_{k,1} \times]\Delta t, \qquad \Psi_k^7 \triangleq E^{51} \tag{48g}$$

$$\Theta_k^8 \triangleq -[\mathbf{d}_{k,2} \times]\Delta t, \qquad \Psi_k^8 \triangleq E^{52} \tag{48h}$$

$$\Theta_k^9 \triangleq -[\mathbf{d}_{k,3} \times]\Delta t, \qquad \Psi_k^9 \triangleq E^{53}. \tag{48i}$$

In (48), the vectors $\mathbf{d}_{k,j}$, $j = 1,2,3$, denote the three columns of the state matrix D_k and the matrices E^{ij} denote 5×5 matrices with 1 at position ij and 0 elsewhere. The augmented noise matrix in (47), W_k, is defined as

$$W_k \triangleq [\mathcal{E}_k \quad \boldsymbol{\mu}_k \quad \boldsymbol{\nu}_k] \tag{49}$$

where, to first order in Δt, $\mathcal{E}_k$ is expressed as

$$\mathcal{E}_k = [\boldsymbol{\epsilon}_k \times] D_k \Delta t. \tag{50}$$

PROOF See Appendix C.

Because the sequences $\mathcal{E}_k$, $\boldsymbol{\mu}_k$, and $\boldsymbol{\nu}_k$ are uncorrelated, the covariance matrix of the augmented noise matrix W_k is expressed by the 15×15 diagonal block matrix $\mathcal{Q}_k$ as follows

$$\mathcal{Q}_k = \text{diag}\{\mathcal{Q}_k^{\epsilon}, Q_k^{\mu}\Delta t, Q_k^{\nu}\Delta t\} \tag{51}$$

where $\mathcal{Q}_k^{\epsilon}$ is defined as in (12).

The measurement model equation for D_{k+1} (see (45)) is equivalent to the following measurement equation for the augmented state matrix X_{k+1},

$$\mathbf{b}_{k+1} = X_{k+1}\mathbf{h}_{k+1} + \mathbf{v}_{k+1} \tag{52}$$

where the 5×1 vector $\mathbf{h}_{k+1}$ is defined as

$$\mathbf{h}_{k+1} \stackrel{\Delta}{=} [\mathbf{r}_{k+1} \quad 0 \quad 0]^T. \tag{53}$$

This is readily shown by using (53) in (52), which yields

$$\begin{aligned}\mathbf{b}_{k+1} &= [D_{k+1} \quad \mathbf{c}_{k+1} \quad \mathbf{m}_{k+1}] \begin{bmatrix} \mathbf{r}_{k+1} \\ 0 \\ 0 \end{bmatrix} + \mathbf{v}_{k+1} \\ &= D_{k+1}\mathbf{r}_{k+1} + \mathbf{v}_{k+1}. \end{aligned} \tag{54}$$

Filter Development and Filter Summary: The augmented process equation (47) is not linear in the state variables. This is already seen by looking at (42) and (43) where the matrix Φ_k is a function of the vectors $\mathbf{c}_k$ and $\mathbf{m}_k$. The nonlinearity appears in (47) through the matrices Θ_k^r, $r = 3,4,5,7,8,9$, which are functions of the elements of D_k (48c)–(48e) and (48g)–(48i). In order to overcome this difficulty we substitute for X_k its best available estimate, $\hat{X}_{k/k}$. This substitution yields a pseudolinear process equation. The measurement equation, on the other hand, is linear in the state. The model equations of the developed pseudolinear matrix plant, together with the assumptions on the statistics of the noises, fit the model assumptions of the generic MKF. Thus, a pseudolinear MKF can operate on the developed matrix plant to get an estimate of the state matrix X_k. This filter is summarized in the following.

1) Initialization:

Choose $\hat{X}_{0/0} = [\hat{D}_{0/0}\hat{\mathbf{c}}_{0/0}\hat{\mathbf{m}}_{0/0}]$ and $\mathcal{P}_{0/0}$.

2) Time Update Equations:

Given $\hat{X}_{k/k} = [\hat{D}_{k/k}\hat{\mathbf{c}}_{k/k}\hat{\mathbf{m}}_{k/k}]$ and $\mathcal{P}_{k/k}$, compute

$$\hat{D}_{k+1/k} = e^{\{[(-\omega_k + \hat{\mathbf{c}}_{k/k} + \hat{\mathbf{m}}_{k/k})\times]\Delta t\}}\hat{D}_{k/k} \tag{55}$$

$$\hat{\mathbf{c}}_{k+1/k} = \hat{\mathbf{c}}_{k/k} \tag{56}$$

$$\hat{\mathbf{m}}_{k+1/k} = \Lambda\hat{\mathbf{m}}_{k/k} \tag{57}$$

$$\hat{X}_{k+1/k} = [\hat{D}_{k+1/k}\hat{\mathbf{c}}_{k+1/k}\hat{\mathbf{m}}_{k+1/k}] \tag{58}$$

$$\Psi_k = \begin{bmatrix} I_3 \otimes e^{(-[\omega_k\times]\Delta t)} & (\hat{D}_{k/k}^T \otimes I_3)\mathcal{L}\Delta t & (\hat{D}_{k/k}^T \otimes I_3)\mathcal{L}\Delta t \\ O_{3\times 9} & I_3 & O_3 \\ O_{3\times 9} & O_3 & \Lambda \end{bmatrix} \tag{59}$$

$$\mathcal{Q}_k = \text{diag}\{\mathcal{Q}_k^{\epsilon}, Q_k^{\mu}\Delta t, Q_k^{\nu}\Delta t\} \tag{60}$$

$$\mathcal{P}_{k+1/k} = \Psi_k\mathcal{P}_{k/k}\Psi_k^T + \mathcal{Q}_k \tag{61}$$

where the matrices Λ and $\mathcal{L}$ are defined from (40) and (11).

3) Measurement Update Equations:

Given $\hat{X}_{k+1/k} = [\hat{D}_{k+1/k}\hat{\mathbf{c}}_{k+1/k}\hat{\mathbf{m}}_{k+1/k}]$ and $\mathcal{P}_{k+1/k}$, compute

$$\mathbf{h}_{k+1}^T = [\mathbf{r}_{k+1}^T \quad 0 \quad 0] \tag{62}$$

$$\mathcal{H}_{k+1} = \mathbf{h}_{k+1}^T \otimes I_3 \tag{63}$$

$$S_{k+1} = \mathcal{H}_{k+1}\mathcal{P}_{k+1/k}\mathcal{H}_{k+1}^T + R_{k+1} \tag{64}$$

$$\mathcal{K}_{k+1} = \mathcal{P}_{k+1/k}\mathcal{H}_{k+1}^T S_{k+1}^{-1} \tag{65}$$

$$\begin{aligned}\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} &+ [\mathcal{K}_{k+1}^1 \quad \mathcal{K}_{k+1}^2 \quad \mathcal{K}_{k+1}^3 \quad \mathcal{K}_{k+1}^4 \quad \mathcal{K}_{k+1}^5] \\ &\times (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1}) \end{aligned} \tag{66}$$

$$\begin{aligned}\mathcal{P}_{k+1/k+1} = &(I_{15} - \mathcal{K}_{k+1}\mathcal{H}_{k+1})\mathcal{P}_{k+1/k}(I_{15} - \mathcal{K}_{k+1}\mathcal{H}_{k+1})^T \\ &+ \mathcal{K}_{k+1}R_{k+1}\mathcal{K}_{k+1}^T \end{aligned} \tag{67}$$

where $\mathcal{K}_{k+1}^j$, $j = 1,\ldots,5$, in (66) denote the 3×3 submatrices of the 15×3 gain matrix $\mathcal{K}_{k+1}$, such that $\mathcal{K}_{k+1}^T = [\mathcal{K}^{1T}\mathcal{K}^{2T}\mathcal{K}^{3T}\mathcal{K}^{4T}\mathcal{K}^{5T}]$.

REMARK 1 As is commonly done in nonlinear filters the estimate time update equation (see (55)) is a nonlinear equation with respect to the estimated variables, $\hat{D}_{k+1/k}$, $\hat{\mathbf{c}}_{k+1/k}$, and $\hat{\mathbf{m}}_{k+1/k}$, and directly stems from the process equation (see (42) and (43)) prior to the pseudolinear transformation (47) and (48).

REMARK 2 The pseudolinear transformation of the process equation (47) and (48) is not unique and care must be taken when choosing that transformation. Indeed, looking at the vectorized dynamics matrix in (59), Ψ_k, we realize that the off-diagonal submatrices at locations 1, 2 and 1, 3 create a coupling between the dynamics of $\hat{D}_{k/k}$ and the dynamics of $\hat{\mathbf{c}}_{k/k}$ and $\hat{\mathbf{m}}_{k/k}$, respectively. This fact ensures observability of the augmented system and, thus, enables the estimation of the gyro biases and drifts although they are not directly measured. On the other hand, had we chosen a different pseudolinear transformation of (42) as follows:

$$D_{k+1} = e^{\{[(-\omega_k + \hat{\mathbf{c}}_{k/k} + \hat{\mathbf{m}}_{k/k})\times]\Delta t\}}D_k \tag{68}$$

this would have yield a block-decoupled dynamical system for the estimate variables with the following

15×15 block-diagonal dynamics matrix Ψ'_k:

$$\Psi'_k = \text{diag}\{e^{\{[(-\omega_k + \hat{\mathbf{c}}_{k/k} + \hat{\mathbf{m}}_{k/k}) \times]\Delta t\}}, I_3, \Lambda\} \tag{69}$$

and the resulting pseudolinear model would not be observable.

ORTHOGONALIZATION

The orthogonality property of D is crucial when the matrix is used to perform vector transformation at a high computation rate as it is in inertial navigation [21]. Therefore, a good estimate of D is one that is nearly orthogonal. The KF is not designed to preserve any relationship among the components of the estimated state matrix $\hat{D}_{k/k}$. Even if the a priori estimate $\hat{D}_{k+1/k}$ is orthogonal, the measurement update equation (31) will not necessarily preserve the orthogonality. In this section, four orthogonalization procedures are presented. The first two algorithms use known results whereas the other two are novel applications of the pseudomeasurement (PM) technique for constrained Kalman filtering. Recall that a square matrix, $M \in \mathbb{R}^{n \times n}$, is orthogonal if it satisfies the identities $M^T M = M M^T = I_n$, where I_n is the identity matrix in $\mathbb{R}^{n \times n}$.

Optimal Brute-Force

The optimal brute-force (OBF) procedure, consists in solving the following optimization problem. Given $\hat{D}_{k/k}$, solve:

$$\min_D \|D - \hat{D}_{k/k}\|_F^2$$
$$\text{subject to} \quad D^T D = I_3 \tag{70}$$

where $\|M\|_F$ denotes the Frobenius norm of the matrix M, that is, $\|M\|_F \triangleq \sqrt{\text{tr}(MM^T)}$. This optimization problem, which is a variant of the orthogonal Procrustes problem [17], has a closed-form solution. The optimal solution $D^*_{k/k}$ is the orthogonal polar factor of $\hat{D}_{k/k}$ and can be efficiently computed via the singular value decomposition (SVD) of $\hat{D}_{k/k}$ (see [18, p. 601]). The orthogonal estimate $D^*_{k/k}$ is then substituted for $\hat{D}_{k/k}$ in the time propagation equation (27). This technique is called "brute-force" because the orthogonalization procedure is performed outside the filter.

Iterative Brute-Force

The iterative brute-force (IBF) procedure is as follows [11]

$$\hat{D}_{0,k} = \hat{D}_{k/k}$$
$$\hat{D}_{n+1,k} = \hat{D}_{n,k}(\tfrac{3}{2} I_3 - \tfrac{1}{2}\hat{D}^T_{n,k}\hat{D}_{n,k}) \tag{71}$$

for $n = 1, 2, \ldots$, until a specific convergence condition is satisfied. Then, set

$$D^*_{k/k} = \hat{D}_{n,k} \tag{72}$$

and use $D^*_{k/k}$ as the current estimate. This algorithm is suboptimal but is less computationally intensive than the OBF procedure, which involves the SVD of $\hat{D}_{k/k}$. When applied recursively, this algorithm produces a sequence of estimates that converges to the optimal solution of (70) [19]. That technique was succesfully applied in a previous work on DCM identification [11]. Moreover, as evidenced by simulations, orthogonality is normally reached, for all practical purposes, after one or two iterations of (71).

First Orthogonality Pseudomeasurement (OPM1)

The PM technique is a way of incorporating constraints into the estimation process. Here we apply this technique in order to orthogonalize the DCM estimate. Consider the orthogonality state constraint, which is expressed through the following matrix equation:

$$I_3 = D_k^T D_k. \tag{73}$$

The matrix D_k is invertible since it is orthogonal and we can write the following identity:

$$\tfrac{1}{2}(D_k + D_k^{-T}) = D_k \tag{74}$$

where $(\cdot)^{-T}$ denotes the composition of the matrix inversion and the matrix transposition. Then, substituting $\hat{D}_{k/k}$ for D_k in the left-hand side of (74), and adding a term V^{ort} in the right-hand side in order to compensate for the modeling error yields

$$\tfrac{1}{2}(\hat{D}_{k/k} + \hat{D}^{-T}_{k/k}) = D_k + V^{ort}_k. \tag{75}$$

As opposed to the true DCM D, the estimated DCM $\hat{D}_{k/k}$ is in general not orthogonal and thus $\hat{D}^{-T}_{k/k}$ is different from $\hat{D}_{k/k}$. This is why the correction term V^{ort} was added on the right-hand side of (75). Equation (75) is an orthogonality PM equation, which is denoted by OPM1, where the left-hand side, $\frac{1}{2}(\hat{D}_{k/k} + \hat{D}^{-T}_{k/k})$, is the measurement and V^{ort} is the measurement noise matrix. Notice that $\hat{D}^{-T}_{k/k}$ needs to be numerically computed and that an $n \times n$ matrix inversion involves a number of FLOPS of the order of n^3. The PM equation, (75), is a matrix equation in D_k on which the MKF can operate in a natural way. The measurement update stage that is associated with the measurement equation (75) is written next

$$Y_k = \tfrac{1}{2}(\hat{D}_{k/k} + \hat{D}^{-T}_{k/k}) \tag{76}$$

$$\mathcal{H}_k = [I_9 \quad O_{9\times 3} \quad O_{9\times 3}] \tag{77}$$

$$\mathcal{S}_k = \mathcal{H}_k \mathcal{P}_{k/k} \mathcal{H}_k^T + \mathcal{R}_k^{ort} \tag{78}$$

$$\mathcal{K}_k = \mathcal{P}_{k/k}\mathcal{H}_k^T\mathcal{S}_k^{-1} \tag{79}$$

$$X^*_{k/k} = \hat{X}_{k/k} + \sum_{j=1}^{5}\sum_{l=1}^{3}\mathcal{K}_k^{jl}(Y_k - \hat{D}_{k/k})E^{lj} \tag{80}$$

$$\mathcal{P}^*_{k/k} = (I_{15} - \mathcal{K}_k\mathcal{H}_k)\mathcal{P}_{k/k}(I_{15} - \mathcal{K}_k\mathcal{H}_k)^T + \mathcal{K}_k\mathcal{R}_k^{ort}\mathcal{K}_k^T. \tag{81}$$

The update stage that is described in (76)–(81) corresponds to the full-covariance filter. Notice that because of the non-zero cross-correlation between the estimation errors, the estimates of the gyro biases and drifts are also updated, which constitutes an essential difference with the brute-force procedures. Assuming that the covariance matrix $\mathcal{R}_k^{ort}$ is of the type $\mu^{ort}I_9$, the reduced covariance form can be developed, which yields the following computations instead of (77)–(81)

$$S_k = P_{k/k} + \mu^{ort}I_3 \tag{82}$$

$$K_k = P_{k/k}S_k^{-1} \tag{83}$$

$$D^*_{k/k} = \hat{D}_{k/k} + \tfrac{1}{2}(\hat{D}^{-T}_{k/k} - \hat{D}_{k/k})K_k^T \tag{84}$$

$$P^*_{k/k} = (I_3 - K_k)P_{k/k}(I_3 - K_k)^T + \mu^{ort}K_kK_k^T. \tag{85}$$

The scalar μ^{ort} has the role of a tuning parameter. We choose μ^{ort} according to the weight given to the orthogonality constraint (see (75)). Specifically, if we pick a high value for μ^{ort} the filter will not significantly update $\hat{D}_{k/k}$, and the new estimate $D^*_{k/k}$ will be close to $\hat{D}_{k/k}$. On the other hand, picking a low value for μ^{ort} will yield a value for $D^*_{k/k}$ that is close to the PM namely to $\frac{1}{2}(\hat{D}_{k/k} + \hat{D}^{-T}_{k/k})$. Extensive Monte-Carlo simulations are used to properly tune the value of μ^{ort}. Notice that not all PM models that can be derived from (73) are efficient. Consider for instance $\hat{D}^{-T} = D + V^{ort}$ and assume that the parameter μ^{ort} is very low, then the resulting updated estimate will be very close to the PM $\hat{D}^{-T}$, which is not guaranteed to be even close to orthogonality. In other words, the type of PMs chosen in this paper are such that they ensure orthogonality in the case of very low variance parameter μ^{ort}.

Second Orthogonality Pseudo-Measurement (OPM2)

In this section we present an alternate model for an orthogonality PM, and develop the associated measurement update equations. Consider (71) from the IBF procedure and assume the right-hand side of that equation to be a PM; that is,

$$\hat{D}_{k/k}(\tfrac{3}{2}I_3 - \tfrac{1}{2}\hat{D}^T_{k/k}\hat{D}_{k/k}) = D_k + V^{ort}. \tag{86}$$

Equation (86), which is denoted OPM2, is a legitimate OPM equation since, assuming that there is no estimation error, that is, $\hat{D}_{k/k} = D_k$, and that $V^{ort} = 0$, then, using the orthogonality of D_k, (86) reduces to the trivial identity $D_k = D_k$. The measurement update stage, which is designed using (86), is readily formulated using the MKF algorithm. The measurement update stage of OPM2 in the full-covariance case is identical to that of OPM1 ((76)–(81)) except the expression for Y_k which here becomes

$$Y_k = \hat{D}_{k/k}(\tfrac{3}{2}I_3 - \tfrac{1}{2}\hat{D}^T_{k/k}\hat{D}_{k/k}). \tag{87}$$

Notice that $\hat{D}_{k/k}$ is in general nonorthogonal and thus the measurement Y_k is different from $\hat{D}_{k/k}$. As compared with OPM1, the computational burden associated with OPM2 is less important. Choosing the covariance matrix $\mathcal{R}_k^{ort}$ to be $\mu^{ort}I_9$, the 3×3 simplified covariance filter equations are identical to (82)–(85), except the estimate update equation, which is here:

$$D^*_{k/k} = \hat{D}_{k/k} + \tfrac{1}{2}\hat{D}_{k/k}(I_3 - \hat{D}^T_{k/k}\hat{D}_{k/k})K_k^T. \tag{88}$$

Notice that even when a gain matrix K_{k+1} is close or equal to the identity matrix, the OPM2 orthogonalization procedure is different from a single step of the IBF procedure (71) because it affects the covariances computation while the IBF procedure does not.

REMARK 3 A rotation matrix is a proper orthogonal matrix. Therefore, including the determinant constraint, that is, $\det(D) = 1$, in the estimation process would be mathematically more rigorous. In practice however, as tested via the extensive Monte-Carlo simulations that are shown in the next section, this is less of an issue. This may be explained by the fact that the unconstrained filter fits the measurements to the physical model, which involves rotation. Furthermore, adding orthogonalization ensures that two eigenvalues of the estimated DCM are reciprocal (one is the inverse of the other) and that the third one is real. The latter estimate being close to 1 (corresponding to a rotation) is thus forced to 1.

SIMULATION STUDY

For this simulation study we considered a spacecraft rotating with respect to an inertial reference frame with the following angular velocity expressed in body coordinates:

$$\boldsymbol{\omega}(t) = 0.2\sin(2\pi t/150)[1, -1, 1]^T \quad [\text{rad/s}]. \tag{89}$$

It was assumed that a triad of body-mounted gyros measured this $\boldsymbol{\omega}(t)$ at a sampling rate of 10 Hz. The vector observations rate was chosen to be 10 Hz too.

TABLE I
Filter Labels

Orthogonalization Procedure	DCM-Only Estimator Reduced Covariance	DCM-Only Estimator Full Covariance
None (nominal)	A0	B0
Optimal Brute Force (OBF)	A1	B1
Iterative Brute Force (IBF)	A2	B2
Orthogonal Pseudomeasurement (OPM)	A3	B3

Orthogonalization Procedure	Augmented Estimator Pure	Augmented Estimator Mixed
None (nominal)	C0	D0
Optimal Brute Force (OBF)	C1	D1
Iterative Brute Force (IBF)	C2	D2
Orthogonal Pseudomeasurement (OPM)	C3	D3

The measurement noise was a zero-mean Gaussian white noise. Its equivalent angular standard deviation was $\sigma_b = 100$ arcs (typical of star tracker sensors accuracy). The tuning parameters for the OPM1 and the OPM2 were chosen as $6\sigma_b^2$. Using these data we ran 100 Monte-Carlo simulation runs until the filters reached a steady-state.

DCM-Only Estimation with Reduced Covariance: Comparison of the Various Orthogonalization Procedures: For simplicity we only consider the DCM estimation problem where the gyro errors consist of zero-mean white Gaussian noises with standard deviations of 0.2 deg/hr, which is a large value compared with commonly used gyros on-board spacecrafts. Four orthogonalization procedures were implemented in the DCM filter with reduced covariance computation (filter of type A): the OBF procedure, the IBF procedure, and the two OPM procedures. A fifth filter, which does not implement any orthogonalization, was also tested. (See Table I for the labeling of the various filters.) The performance of the five estimators were compared using two figures of merit, which are defined by

$$J_c(k) \triangleq \|D_k - D^*_{k/k}\|_F \tag{90}$$

$$J_o(k) \triangleq \|I_3 - (D^*_{k/k})^T D^*_{k/k}\|_F \tag{91}$$

where D_k is the true state matrix and $D^*_{k/k}$ is the estimated matrix after orthogonalization. (When orthogonalization was not performed $D^*_{k/k}$ was just $\hat{D}_{k/k}$.) Clearly, $J_c(k)$ is a DCM-estimation convergence criterion, while $J_o(k)$ is a DCM-orthogonality convergence criterion. The results are summarized in Tables II and III and in Figs. 1 and 2. It turned out that the two PM techniques yielded very similar performance and thus they are shown under the same label (A3). Figs. 1(a) and 1(b) present the time histories of the Monte-Carlo means of J_c for all the cases. Fig. 1(a) shows the transient phase, and Fig. 1(b) shows the steady-state phase. Following transients of around 2 s the J_c indices reach their steady-states with very small decay rates. As can be seen in Fig. 1(a), during the transients, filter A3, which implements the PM technique, achieves a better performance (smoother, less overshoot) than all other filters. In particular, filters A1 and A2 show relatively high overshoots. These picks illustrate the nonsmooth characteristics of the brute-force orthogonalization procedures as opposed to the OPM-based procedures, which were designed to be embedded in the MKF. From Fig. 1(b), we see that A1 yields the lower bound of all four solutions while the nominal algorithm, A0, yields the upper bound. As expected, A2 converges asymptotically to A1. The performance of A3 remains on an intermediary steady-state level. Table II shows the Monte-Carlo means of $J_c.10^5$ and of J_o for the five algorithms at the final time. It appears that both filters A1 and A2 reach numerical accuracy for the orthogonality property, which seems to be the reason for which they outperform the filter A3 in steady-state. Fig. 2 depicts the time variations of the Monte-Carlo means of J_o during the simulation time span. The A2 algorithm yields an estimate that is orthogonal for all practical purpose and is computationally less cumbersome than the filter A1. The A3 filters produce a better orthogonal estimate than the nominal filter, which in return explains the improvement in the convergence performance observed in Fig. 1. Although the nominal algorithm A0 yields an estimate that tends to be orthogonal,

TABLE II
Monte-Carlo Means of $(10^5 J_c)$ and of J_o at Final Time for Filters of Type A

Filter	$(10^5 J_c)$	J_o
A0	6.6	5.10^{-4}
A1	3.4	10^{-30}
A2	3.4	10^{-15}
A3	5.4	10^{-4}

TABLE III
Comparative Table of Orthogonalization Procedures for Filters of Type A

	Advantages	Disadvantages
OBF	optimal closed-form	high computation bad transient
IBF	low computation asympt. optimal	bad transient
OPM (1 and 2)	smooth transient covariance computation	suboptimal needs tuning

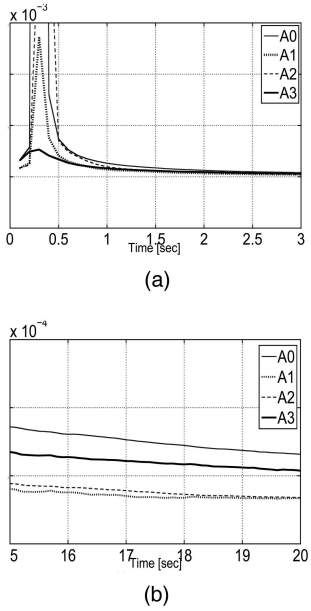


Fig. 1. Monte-Carlo means of DCM estimation convergence index J_c. (a) Transient. (b) Steady-state.

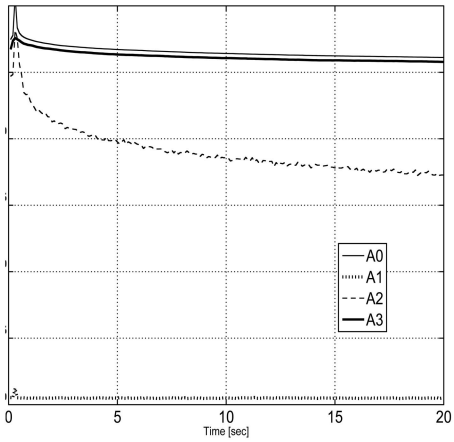


Fig. 2. Monte-Carlo means of orthogonality convergence index J_o.

as it converges to the true DCM, it still remains outperformed by any other filter that implements any orthogonalization procedure. A qualitative comparison between the various orthogonalization procedures is proposed in Table III.

DCM-Only Estimation: Reduced Covariance versus Full-Covariance: In order to investigate the trade-off between the reduction in the dimensions of the covariance computations and the implementation of more accurate noise covariance matrices in the filter, two series of filters were tested: filters A, which feature reduced-covariance computations, and filters B with full-covariance computations (see Table I). The results are summarized in Table IV and Fig. 3.

Table IV shows the ratios between filters A and filters B of the Monte-Carlo means for the convergence indices at the final time. It appears that

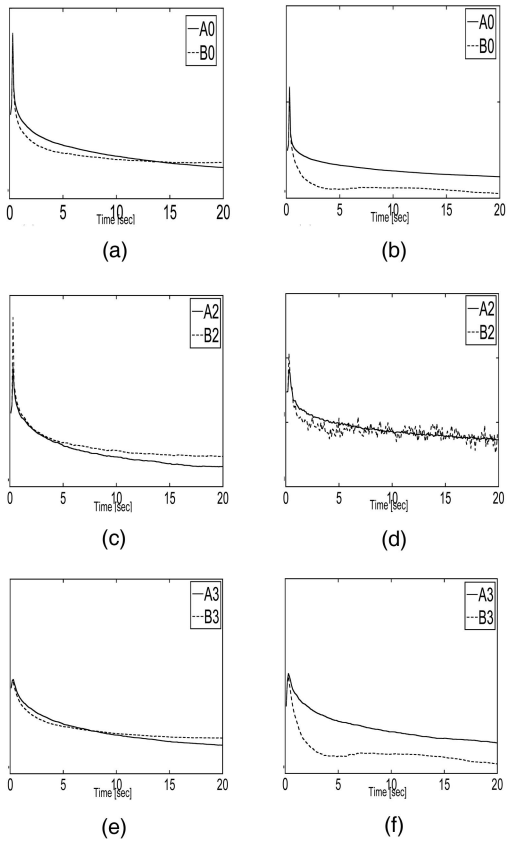


Fig. 3. Monte-Carlo means of DCM estimation and orthogonality convergence indices J_C (left column) and J_O (right column), for cases of reduced-covariance (A) and full-covariance (B) computations in filters of type nominal (0), IBF (2), and OPM (3). (a) Index J_C in nominal filters. (b) Index J_O in nominal filters. (c) Index J_C in IBF filters. (d) Index J_O in IBF filters. (e) Index J_C in OPM filters. (f) Index J_O in OPM filters.

TABLE IV
Comparison of Performances Between Filters of Type A (Reduced Covariance) and Filters of Type B (Full Covariance) for Various Orthogonalization Procedures

Orthogonalization Procedure	$\frac{J_{\text{C, filter A}}}{J_{\text{C, filter B}}}(t_f)$	$\frac{J_{\text{O, filter A}}}{J_{\text{O, filter B}}}(t_f)$
None (nominal)	0.7	7
Optimal Brute Force (OBF)	0.5	1
Iterative Brute Force (IBF)	0.5	1
Orthogonal Pseudomeasurement (OPM)	0.6	3

Note: Table values are Monte-Carlo means of ratios at final time.

the final estimation accuracy is always improved when using the reduced-covariance schemes. On the other hand, as can be seen in Table IV, using the full-covariance scheme enhances the orthogonality of the estimate. Fig. 3 depicts the time-histories of the Monte-Carlo means of the convergence indices for the reduced-covariance filters (plain lines) and the full-covariance filters (dotted lines). Filters implementing the OBF technique were omitted since

they can advantageously be replaced by the IBF-based filters. The first column of Fig. 3 shows that the filters of type B have steeper transient rates than their counterparts of type A, independently of the orthogonalization process. Moreover, as seen from the second column, the B filters clearly outperform their counterparts during the transients. The later facts are the consequences of two causes. The first one is that the full-covariance expression induces a better modeling of the noise statistics, which in turn reflects more accurately the orthogonality relationships between the DCM components, and therefore enhances the orthogonalization of the estimate. The second cause is that implementing a full-covariance computation yields a higher norm for the process noise covariance matrix in the filter (see (12) and (23)). Thus, the filter response becomes quicker. However, the counter-effect appears through a poorer steady-state estimation accuracy. To conclude, the reduced-covariance filters perform advantageously in steady-state while the full-covariance filters have quicker responses.

Augmented Pure Filter: Comparison of Various Orthogonalization Procedures: A series of Monte-Carlo simulations was run (Series C, see Table I) in order to investigate the performance of the augmented filter for various orthogonalization procedures. For each one of the three gyro outputs, the augmented model featured constant biases of 1 deg/hr in each axis, as well as three drifts that converged from initial values of 1 deg/hr at identical rates. These rates stemmed from $\tau_x = 1[\mathrm{s}^{-1}]$, $\tau_y = 1[\mathrm{s}^{-1}]$ and $\tau_z = 1[\mathrm{s}^{-1}]$ (see (41)). The process noises, $\boldsymbol{\mu}_k$ and $\boldsymbol{\nu}_k$, were Gaussian zero-mean white sequences with standard deviations equal to 0.05 deg/hr; that is, $\sigma_\nu/\Delta t = \sigma_\mu/\Delta t = 0.05$ deg/hr. The performances of the bias and drift estimation were measured via the two following indices:

$$J_{\mathrm{BIAS}}(k) = \|\mathbf{c}_k - \hat{\mathbf{c}}_{k/k}\| \tag{92}$$

$$J_{\mathrm{DRIFT}}(k) = \|\mathbf{m}_k - \hat{\mathbf{m}}_{k/k}\|. \tag{93}$$

The comparison simulation tested three filters: filter C0 (nominal), filter C2 (IBF) and filter C3 (OPM). Each Monte-Carlo run lasted 10 hr, which is a typical time for the constant bias estimation to be efficient. The Monte-Carlo means of the convergence indices J_c and J_{BIAS} for the three filters are plotted in Figs. 4(a) and Fig. 4(b), respectively. Fig. 4(a) shows that filter C0 presents the best transient while C2 and C3 have similar ones. In steady-state however, after 4.5 hr, the filters C2 and C3, still performing similarly, clearly show a better accuracy than C0. Regarding the performance in the estimation of the biases, Fig. 4(b) shows that the filter C0 has a better transient than the two others. However, in steady-state, the C0 plot blends with that of C2 while C3 becomes the most accurate estimator. The bias error dropped from an initial value of 1.7 deg/hr to a level of 1 mdeg/hr. It happens that the drift estimation performance reaches the same level of accuracy after only 20 min, which is due to the fact that the drifts are modeled as convergent first-order sequences. Since the results are identical in all three filters, there is no pertinence in showing the J_{DRIFT} plots. One may conclude from the presented results that, as opposed to the DCM-only estimation scheme, the orthogonalization procedures in the augmented filter are penalizing the DCM estimation as long as the biases are not correctly estimated. Thus, during a transient phase, which lasts here approximately 4.5 hr, the nominal filter should be used. Then the estimator should switch to the OPM-based algorithm if the best accuracy in the estimated biases is desired, or to the IBF-based algorithm if the best accuracy in the estimated DCM is desired. However, since the performance between these two filters is quite similar, the IBF seems preferable for computational complexity.

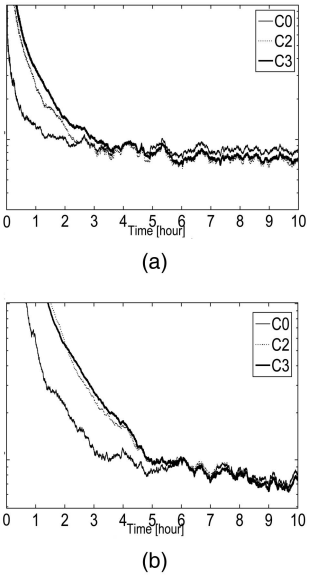


Fig. 4. Monte-Carlo means of estimation convergence indices J_c and J_{BIAS} in augmented pure filter (C) for nominal case (C0), IBF-orthogonalization case (C2), and OPM-orthogonalization case (C3). (a) DCM convergence index. (b) BIAS convergence index.

Augmented Mixed Filter: For long duration missions, it may be feasible for the sake of computation savings, to implement a mixed filter, which would switch between the full-blown augmented version (filter C) and the DCM-only estimator (filter B). A series of Monte-Carlo simulations was run to test this type of filter, referred to as filter D. At the time of switching from C to B, any variable that is related to the bias or drift estimation was "frozen"; the relevant variables for the DCM estimation were extracted and inserted as initial conditions in the B-filter. At the switching from B back to C, the DCM-related variables were inserted back into filter C, and the

bias and drift estimation resumed from their frozen values. As an illustration, we considered the 10 hr span simulation performed previously with the following profile: filter C was implemented during the first 5 hr, then filter B was implemented at $t = 7$ hr and at $t = 9$ hr during 6 min each time. The results are summarized in Figs. 5–7. Fig. 5 shows the Monte-Carlo means of the indices J_c and J_{BIAS} for three types of filters, D0 (nominal), D2 (IBF), and D3 (OPM) (see Table I). It can be seen from Fig. 5(a) that all the filters D converge and stay at a reasonable level of accuracy after $t = 5$ hr. It is interesting to see that both D2 and D3 outperform D0 and that D3 achieves the best accuracy among the three filters. The latter happens despite that the switchings to filter C at $t = 7$ hr and $t = 9$ hr do not necessarily decrease J_{BIAS} (see Fig. 5(b)). For completeness, Fig. 6 shows the worst case realizations (out of a sample of 1000 Monte-Carlo runs), the Monte-Carlo means and the Monte-Carlo 1σ upper bound for the DCM index of convergence J_c for each orthogonalization method (nominal, IBF, and OPM). The lower bound of the 1σ envelope being negative is not shown on the figure since J_c is a nonnegative number. The worst case realization of J_c is defined as the maximizer over the 1000-runs-Monte-Carlo sample of the time maxima of $J_c(t)$ after steady-state is reached (here, after 5 hr); that is $J_c^{\text{worst}} = \max_{i=1,2,\ldots,1000}[\max_{5\le t\le 10\ \text{hr}}[J_c^i(t)]]$. It appears that the worst cases are yielding acceptable performance in the convergence of the DCM estimation. As observed earlier on the Monte-Carlo mean, using the OPM technique slows down the transient but ensures better steady-state accuracy when compared with the nominal or the IBF filters. In order to quantify the latter, Fig. 7 is drawn that depicts the Monte-Carlo means ratios, $J_c(D)/J_c(C)$ and $J_{\text{BIAS}}(D)/J_{\text{BIAS}}(C)$, for the three types of orthogonalization. By plotting these variables, one can compare the performance of the mixed filters D with their pure versions (C), when the full-blown augmented filters are implemented all the time as in the previous subsection. We observe in Fig. 7(a) that all the ratios are in general greater than one, as expected. The worst discrepancy between the performance of filters C and D are in the IBF-based filters with a time-average of 1.5. The least discrepancy takes place in the OPM-based filters with a time-average of 1.1. Clearly, the switchings to filters C increase the accuracy of the DCM estimates in the nominal and IBF-based filters, while the OPM case seems to be less sensitive. Fig. 7(b) shows that the bias error indices tend to be bounded along the 10 hr simulation span, but no clear conclusion can be drawn from this example, which calls for further investigation with different switching times.

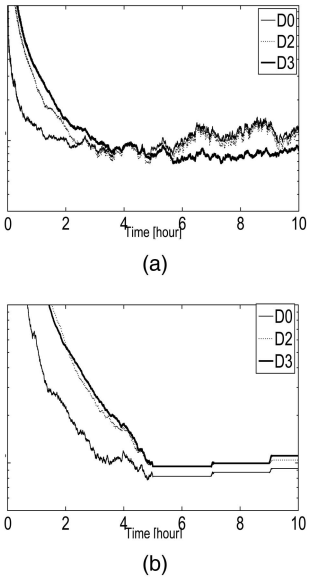
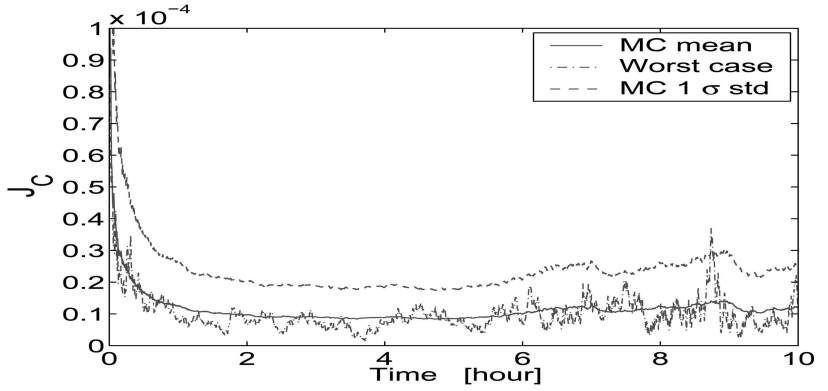


Fig. 5. Monte-Carlo means of estimation convergence indices J_c and J_{BIAS} in augmented mixed filter (D) for nominal case (D0), IBF-orthogonalization case (D2), and OPM-orthogonalization case (D3). (a) DCM convergence index. (b) BIAS convergence index.

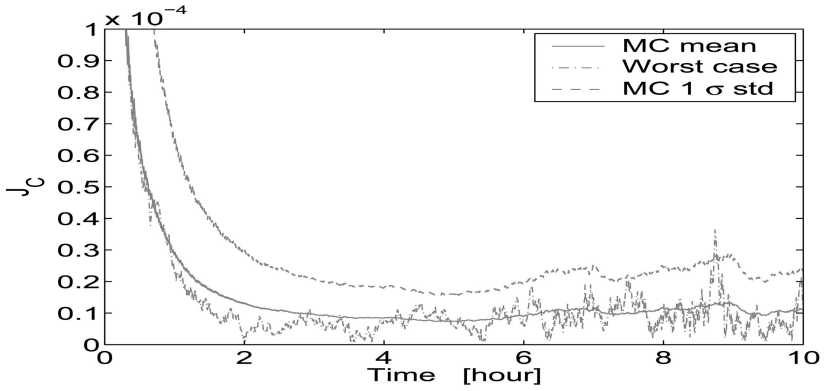
CONCLUSION

This work presented several MKFs for estimation of the DCM using vector observations, where a central common feature was the preservation of the natural matrix notation of the state-space system model on which they operate. Explicit expressions for the system's noise 9×9 covariance matrices were obtained and, under simplifying assumptions on their structure, a reduced (3×3) covariance computation algorithm was developed without impairing much of the estimation accuracy. The case of constant biases and time-varying drifts in the gyro output was handled via state augmentation yielding an augmented MKF. Four different procedures were added to the DCM-only estimators in order to enforce on-line orthogonalization of the state matrix estimate. Two of them were known "brute-force" procedures (IBF and OBF), while the other two featured novel applications of the OPM technique and could be readily implemented in an MKF.

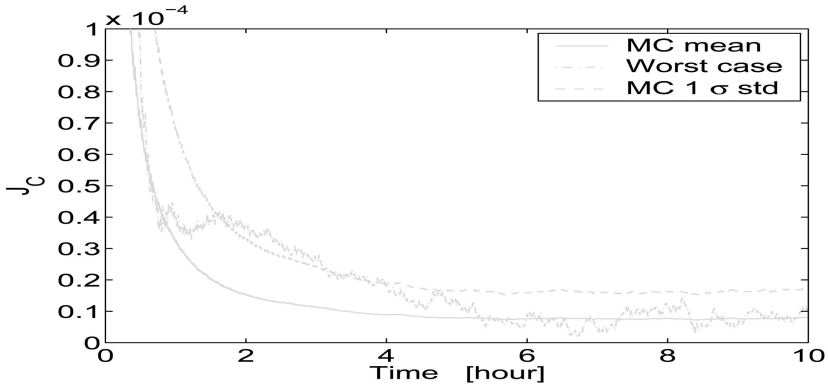
Extensive Monte-Carlo simulations illustrated the performanceã of the various filters. In the case of DCM-only estimation with reduced covariance, any orthogonalization procedure enhanced the estimation process. A good practice would be to start with the OPM filter in order to produce a smooth transient and to have consistent filter covariance computations and, then, to switch to the IBF filter for better steady-state accuracy. In order to create a steeper transient in the DCM estimation error and to enhance the orthogonality of the DCM estimate, it is recommended to implement a full-covariance computation during the transient phase before switching to the reduced-covariance algorithm.



(a)



(b)



(c)

Fig. 6. Worst case realizations, Monte-Carlo means, and Monte-Carlo $\pm\sigma$ envelopes of estimation convergence index J_c in augmented mixed filter (D) for nominal case (D0), IBF-orthogonalization case (D2), and OPM-orthogonalization case (D3). (a) Nominal case D0. (b) IBF case—D2. (c) OPM case—D3.

When high-valued constant biases needed to be estimated, extensive Monte-carlo simulations of the augmented filters showed that any orthogonalization procedure was penalizing the DCM estimation process until the biases were correctly estimated. The orthogonalization-free filter should thus be first implemented during the long transient phase before activating an orthogonalization procedure. For long duration simulations, it turned out that switching off from time to time the biases and drifts estimation in steady-state still yielded a successful DCM estimation. In that case, extensive Monte-Carlo simulations showed that the OPM filter yielded the best steady-state accuracy for the DCM error, and bore the least discrepancy with the case where bias estimates were continuously produced. Nevertheless, further investigation with respect to the switching times and periods should be done.

APPENDIX A. REVIEW OF THE MATRIX KALMAN FILTER

Matrix State-Space Model: Consider a linear discrete-time stochastic dynamic system governed by

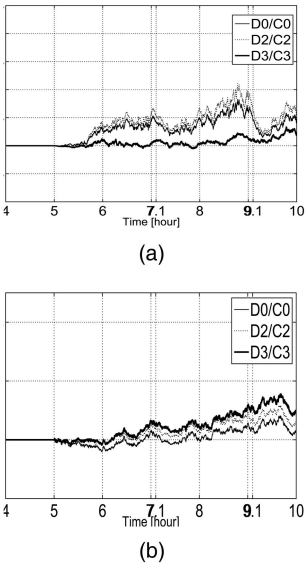


Fig. 7. Monte-Carlo means of ratios of estimation convergence indices J_c and J_{BIAS} between augmented mixed filter (D) and augmented pure filter (C) for nominal case (D0/C0), IBF-orthogonalization case (D2/C2), and OPM-orthogonalization case (D3/C3). (a) Ratio of DCM convergence indices. (b) Ratio of BIAS convergences indices.

the difference equation

$$X_{k+1} = \sum_{r=1}^{\mu} \Theta_k^r X_k \Psi_k^r + W_k \tag{94}$$

where $X_k \in \mathbb{R}^{m\times n}$ is the matrix state variable at time t_k, $\Theta_k^r \in \mathbb{R}^{m\times m}$ and $\Psi_k^r \in \mathbb{R}^{n\times n}$ are "transition matrices," and W_k is an $m \times n$ noise matrix. The matrix measurement equation of the matrix plant is

$$Y_{k+1} = \sum_{s=1}^{\nu} H_{k+1}^s X_{k+1} G_{k+1}^s + V_{k+1} \tag{95}$$

where $Y_{k+1} \in \mathbb{R}^{p\times q}$ is the matrix measurement of X_{k+1} at time t_{k+1}, $H_{k+1}^s \in \mathbb{R}^{p\times m}$, and $G_{k+1}^s \in \mathbb{R}^{n\times q}$ are the observation matrices, and $V_{k+1} \in \mathbb{R}^{p\times q}$ is a noise matrix. Let W denote any $m \times n$ random matrix with generic element w_{ij}, $i = 1,\ldots,m$, $j = 1,\ldots,n$, then the expectation of W is defined as the matrix of the expectations of w_{ij}; that is $E\{W\} \triangleq [E\{w_{ij}\}]$. Let vec denote the mapping from $\mathbb{R}^{m\times n}$ to $\mathbb{R}^{mn}$ which operates on a rectangular matrix by stacking its columns one underneath the other to form a single column-matrix (see [16, p. 244]). The covariance matrix of W is defined as the covariance of its vec-transform [20]; that is $\text{cov}\{W\} \triangleq \text{cov}\{\text{vec}W\}$. Obviously, if $W \in \mathbb{R}^{m\times n}$, then $\text{cov}\{W\} \in \mathbb{R}^{mn\times mn}$. The matrix sequences W_k and V_k are assumed to be zero-mean Gaussian white noise sequences with covariance matrices $Q_k \in \mathbb{R}^{mn\times mn}$ and $R_k \in \mathbb{R}^{pq\times pq}$, respectively. The initial state X_0 is assumed to be Gaussian distributed with mean $\bar{X}_0$ and $mn \times mn$ covariance matrix Π_0. Also, W_k, V_k, and X_0 are uncorrelated with one another. The system and measurement equations of any linear matrix plant are a special case of (94) and (95) [14].

Estimation Problem: The MKF is the unbiased minimum variance estimator of the $m \times n$ matrix state X_k at t_k, given a sequence of $p \times q$ matrix observations up to t_k, $\{Y_l\}$, $l = 1,2,\ldots,k$. Let $\hat{X}_{k/k}$ and $\Delta X_{k/k}$ denote, respectively, the a posteriori state estimate and the a posteriori estimation error; that is, $\Delta X_{k/k} \triangleq X_k - \hat{X}_{k/k}$. Let $P_{k/k}$ denote the a posteriori estimation error covariance matrix; that is $P_{k/k} = \text{cov}(\Delta X_{k/k})$. Then, the filtering problem is equivalent to the following minimization problem

$$\min_{\hat{X}_{k/k}} \{\text{tr}(P_{k/k})\} \tag{96}$$

subject to (94) and (95) and to the stochastic assumptions on the noises and the initial conditions. The MKF is summarized next.

Summary of the MKF:

1) Initialization: $\hat{X}_{0/0} = \bar{X}_0$ and $P_{0/0} = \Pi_0$.
2) Time Update Equations:

$$\hat{X}_{k+1/k} = \sum_{r=1}^{\mu} \Theta_k^r \hat{X}_{k/k} \Psi_k^r \tag{97}$$

$$\Phi_k = \sum_{r=1}^{\mu} [(\Psi_k^r)^T \otimes \Theta_k^r] \tag{98}$$

$$P_{k+1/k} = \Phi_k P_{k/k} \Phi_k^T + Q_k \tag{99}$$

where $\otimes$ denotes the Kronecker product [16].

3) Measurement Update Equations:

$$\tilde{Y}_{k+1} = Y_{k+1} - \sum_{s=1}^{\nu} H_{k+1}^s \hat{X}_{k+1/k} G_{k+1}^s \tag{100}$$

$$\mathcal{H}_{k+1} = \sum_{s=1}^{\nu} [(G_{k+1}^s)^T \otimes H_{k+1}^s] \tag{101}$$

$$S_{k+1} = \mathcal{H}_{k+1} P_{k+1/k} \mathcal{H}_{k+1}^T + R_{k+1} \tag{102}$$

$$K_{k+1} = P_{k+1/k} \mathcal{H}_{k+1}^T S_{k+1}^{-1} \tag{103}$$

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{n} \sum_{l=1}^{q} K_{k+1}^{jl} \tilde{Y}_{k+1} E^{lj} \tag{104}$$

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

$$K_{k+1} = \underbrace{\left.\begin{bmatrix} K_{k+1}^{11} & \cdots & K_{k+1}^{1l} & \cdots \\ \vdots & \ddots & \vdots & \ddots \\ K_{k+1}^{j1} & \cdots & K_{k+1}^{jl} & \cdots \\ \vdots & \ddots & \vdots & \ddots \end{bmatrix}\right\} n \text{ matrices}}_{q \text{ matrices}} \tag{105}$$

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

$$P_{k+1/k+1} = (I_{mn} - K_{k+1}\mathcal{H}_{k+1})P_{k+1/k}(I_{mn} - K_{k+1}\mathcal{H}_{k+1})^T + K_{k+1}R_{k+1}K_{k+1}^T. \quad (106)$$

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

$$\text{var}\{\Delta X[i,j]\} = P[(j-1)m+i,(j-1)m+i] \quad (107a)$$

$$\text{cov}\{\Delta X[i,j],\Delta X[k,l]\} = P[(j-1)m+i,(l-1)m+k] \quad (107b)$$

where $i,k = 1,2,\ldots,m$, and $j,l = 1,2,\ldots,n$. The variable ΔX denotes either the a posteriori or the a priori estimation error as applicable, and P is the associated covariance matrix. The MKF, described in (97)–(107b), behaves like the conventional linear KF. It is a time-varying algorithm for optimal recursive estimation of a matrix state process using a sequence of matrix measurements. The MKF is a natural extension of the conventional KF. The two new dimensions are the number of columns in the state matrix n and the number of columns in the measurement matrix q. Conversely, the ordinary vector KF is a special case of the MKF (by taking $n = q = 1$).

APPENDIX B. REDUCED COVARIANCE FILTER

Let $\otimes$ denote the Kronecker product. The model assumptions are as follows

$$\mathcal{Q}_k^\epsilon = Q_k \otimes I_3 \quad (108a)$$

$$R_k = \mu_k I_3 \quad (108b)$$

$$\mathcal{P}_{0/0} = P_{0/0} \otimes I_3. \quad (108c)$$

Time Update Stage: Assuming that $\mathcal{P}_{k/k} = P_{k/k} \otimes I_3$, the 9×9 covariance time-update equation is formulated as (see (16))

$$\begin{aligned}\mathcal{P}_{k+1/k} &= \Psi_k \mathcal{P}_{k/k}\Psi_k^T + \mathcal{Q}_k^\epsilon \\ &= (I_3 \otimes \Phi_k)(P_{k/k} \otimes I_3)(I_3 \otimes \Phi_k)^T + \mathcal{Q}_k^\epsilon \\ &= (P_{k/k} \otimes \Phi_k\Phi_k^T) + \mathcal{Q}_k^\epsilon \\ &= (P_{k/k} \otimes I_3) + (Q_k \otimes I_3) \\ &= (P_{k/k} + Q_k) \otimes I_3. \end{aligned} \quad (109)$$

The third equality in (109) results from the mixed-product property of the Kronecker product (see [16, p. 244]). The fourth equality is due to the orthogonality of Φ_k. The last equality is obtained using Lemma 4.2.7, [16, p. 213]. Thus, we get (33) by defining

$$P_{k+1/k} \triangleq P_{k/k} + Q_k. \quad (110)$$

Measurement Update Stage: Using (109) and (110), the 9×3 Kalman gain matrix is computed as follows (see (19))

$$\begin{aligned}\mathcal{K}_{k+1} &= \mathcal{P}_{k+1/k}\mathcal{H}_{k+1}^T\mathcal{S}_{k+1}^{-1} \\ &= (P_{k+1/k} \otimes I_3)(\mathbf{r}_{k+1}^T \otimes I_3)^T \\ &\quad \times [(\mathbf{r}_{k+1}^T \otimes I_3)(P_{k+1/k} \otimes I_3) \times (\mathbf{r}_{k+1}^T \otimes I_3)^T + \mu_{k+1}I_3]^{-1} \\ &= (P_{k+1/k}\mathbf{r}_{k+1} \otimes I_3)[(\mathbf{r}_{k+1}^T P_{k+1/k}\mathbf{r}_{k+1} + \mu_{k+1})I_3]^{-1} \\ &= [P_{k+1/k}\mathbf{r}_{k+1}/(\mathbf{r}_{k+1}^T P_{k+1/k}\mathbf{r}_{k+1} + \mu_{k+1})] \otimes I_3 \end{aligned} \quad (111)$$

where the third and fourth equalities stem from the mixed-product property of the Kronecker product. Thus, we get (36) by defining

$$\mathbf{g}_{k+1} \triangleq P_{k+1/k}\mathbf{r}_{k+1}/s_{k+1} \quad (112)$$

where

$$s_{k+1} \triangleq \mathbf{r}_{k+1}^T P_{k+1/k}\mathbf{r}_{k+1} + \mu_{k+1}. \quad (113)$$

The measurement update equation for the covariance matrix is simplified as follows.

$$\begin{aligned}\mathcal{P}_{k+1/k+1} &= (I_9 - \mathcal{K}_{k+1}\mathcal{H}_{k+1})\mathcal{P}_{k+1/k}(I_9 - \mathcal{K}_{k+1}\mathcal{H}_{k+1})^T \\ &\quad + \mathcal{K}_{k+1}R_{k+1}\mathcal{K}_{k+1}^T \\ &= [(I_3 \otimes I_3) - (\mathbf{g}_{k+1} \otimes I_3)(\mathbf{r}_{k+1}^T \otimes I_3)](P_{k+1/k} \otimes I_3) \\ &\quad \times [(I_3 \otimes I_3) - (\mathbf{g}_{k+1} \otimes I_3)(\mathbf{r}_{k+1}^T \otimes I_3)]^T \\ &\quad + (\mathbf{g}_{k+1} \otimes I_3)(\mu_{k+1}I_3)(\mathbf{g}_{k+1} \otimes I_3)^T \\ &= [(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T) \otimes I_3](P_{k+1/k} \otimes I_3) \\ &\quad \times [(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T) \otimes I_3]^T \\ &\quad + (\mu_{k+1}\mathbf{g}_{k+1}\mathbf{g}_{k+1}^T) \otimes I_3 \end{aligned} \quad (114)$$

$$= [(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)P_{k+1/k}(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)^T + \mu_{k+1}\mathbf{g}_{k+1}\mathbf{g}_{k+1}^T] \otimes I_3 \quad (115)$$

where we repeatedly used the mixed-product property of the Kronecker product to obtain the third and fourth equalities. We get (34) by defining

$$P_{k+1/k+1} \triangleq (I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)P_{k+1/k}(I_3 - \mathbf{g}_{k+1}\mathbf{r}_{k+1}^T)^T + \mu_{k+1}\mathbf{g}_{k+1}\mathbf{g}_{k+1}^T. \quad (116)$$

The measurement update equation for the state estimate (20) is simplified as follows. Using (36) in

(20) yields

$$\begin{aligned}\hat{D}_{k+1/k+1} &= \hat{D}_{k+1/k} + [\mathcal{K}^1_{k+1} \quad \mathcal{K}^2_{k+1} \quad \mathcal{K}^3_{k+1}] \\ &\quad \times [I_3 \otimes (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})] \\ &= \hat{D}_{k+1/k} + [g_{k+1}[1]I_3 \quad g_{k+1}[2]I_3 \quad g_{k+1}[3]I_3] \\ &\quad \times [I_3 \otimes (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})] \\ &= \hat{D}_{k+1/k} + (\mathbf{g}^T_{k+1} \otimes I_3)[I_3 \otimes (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})] \\ &= \hat{D}_{k+1/k} + [\mathbf{g}^T_{k+1} \otimes (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})] \\ &= \hat{D}_{k+1/k} + (\mathbf{b}_{k+1} - \hat{D}_{k+1/k}\mathbf{r}_{k+1})\mathbf{g}^T_{k+1} \end{aligned} \tag{117}$$

where $g_{k+1}[j]$, $j = 1,2,3$, denotes the jth component of the 3×1 vector $\mathbf{g}_{k+1}$. The passage form the third to the fourth equality in (117) results from the mixed-product property and the last equality is a direct result of the definition of the Kronecker product for two column-matrices.

APPENDIX C. PROOF OF PROPOSITION 1

PROOF The proof proceeds by direct computation of the terms in (47). In the development we use the symbol $O_{i\times j}$ to denote the $i \times j$ matrix of zeros. Straightforward computations yield to

$$\Theta^1_k X_k \Psi^1_k = [e^{(-[\omega_k\times]\Delta t)} D_k \quad O_{3\times 2}] \tag{118a}$$

$$\Theta^2_k X_k \Psi^2_k = [O_{3\times 3} \quad \mathbf{c}_k \quad \mathbf{0}_{3\times 1}] \tag{118b}$$

$$\Theta^3_k X_k \Psi^3_k = [[\mathbf{c}_k\times]\mathbf{d}_{k,1} \quad O_{3\times 4}]\Delta t \tag{118c}$$

$$\Theta^4_k X_k \Psi^4_k = [\mathbf{0}_{3\times 1} \quad [\mathbf{c}_k\times]\mathbf{d}_{k,2} \quad O_{3\times 3}]\Delta t \tag{118d}$$

$$\Theta^5_k X_k \Psi^5_k = [O_{3\times 2} \quad [\mathbf{c}_k\times]\mathbf{d}_{k,3} \quad O_{3\times 2}]\Delta t \tag{118e}$$

$$\Theta^6_k X_k \Psi^6_k = [O_{3\times 3} \quad \mathbf{0}_{3\times 1} \quad \Lambda\mathbf{m}_k] \tag{118f}$$

$$\Theta^7_k X_k \Psi^7_k = [[\mathbf{m}_k\times]\mathbf{d}_{k,1} \quad O_{3\times 4}]\Delta t \tag{118g}$$

$$\Theta^8_k X_k \Psi^8_k = [\mathbf{0}_{3\times 1} \quad [\mathbf{m}_k\times]\mathbf{d}_{k,2} \quad O_{3\times 3}]\Delta t \tag{118h}$$

$$\Theta^9_k X_k \Psi^9_k = [O_{3\times 2} \quad [\mathbf{m}_k\times]\mathbf{d}_{k,3} \quad O_{3\times 2}]\Delta t. \tag{118i}$$

Summing (118a) and (118i) yields

$$\begin{aligned}\sum_{r=1}^{9} \Theta^r_k X_k \Psi^r_k = [&(e^{(-[\omega_k\times]\Delta t)} D_k + [\mathbf{c}_k\times]D_k\Delta t \\ &+ [\mathbf{m}_k\times]D_k\Delta t) \quad \mathbf{c}_k \quad \Lambda\mathbf{m}_k].\end{aligned} \tag{119}$$

Recalling that the 3×3 submatrix in the right-hand side of (119) is, to first order in Δt, equivalent to the matrix $e^{\{[(-\omega_k+\mathbf{c}_k+\mathbf{m}_k)\times]\Delta t\}} D_k$, and using the definition of the noise matrix W_k (see (49)) yields

$$\begin{aligned}&\sum_{r=1}^{9} \Theta^r_k X_k \Psi^r_k + W_k \\ &\quad = [e^{\{[(-\omega_k+\mathbf{c}_k+\mathbf{m}_k)\times]\Delta t\}} D_k \quad \mathbf{c}_k \quad \Lambda\mathbf{m}_k] + [\mathcal{E}_k \quad \boldsymbol{\mu}_k \quad \boldsymbol{\nu}_k] \\ &\quad = [e^{\{[(-\omega_k+\mathbf{c}_k+\mathbf{m}_k)\times]\Delta t\}} D_k + \mathcal{E}_k \quad \mathbf{c}_k + \boldsymbol{\mu}_k \quad \Lambda\mathbf{m}_k + \boldsymbol{\nu}_k] \\ &\quad = [D_{k+1} \quad \mathbf{c}_{k+1} \quad \mathbf{m}_{k+1}] \\ &\quad = X_{k+1}\end{aligned} \tag{120}$$

where the last two equalities stem from the model equations (38)–(44).

ACKNOWLEDGMENT

We wish to dedicate this paper to the memory of Itzhack Y. Bar-Itzhack who passed away between the time that this paper was accepted and the time it was printed. At the time of writing, Prof. Bar-Itzhack held the Sophie and William Shamban Chair in Aeronautical Engineering and part of this work was performed while he was holding a National Research Council Research Associateship Award at NASA-Goddard Space Flight Center. Itzhack Y. Bar-Itzhack was an eminent and world-renowned scientist, an enlightening advisor, a caring mentor, a kind colleague and a sincere friend. He is deeply missed.

REFERENCES

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

[2] Black, H. D.
A passive system for determining the attitude of a satellite.
AIAA Journal, **2**, 7 (1964), 1350–1351.

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

[4] Farrel, J. L., and Stuelpnagel, J. C. (eds.)
A least squares estimate of spacecraft attitude.
SIAM Review, **8**, 3 (1966), 384–386.

[5] Wessner, R. H.
In: A least squares estimate of spacecraft attitude.
SIAM Review, **8**, 3 (1966), 384–386.

[6] Velman, J. R.
In: A least squares estimate of spacecraft attitude.
SIAM Review, **8**, 3 (1966), 384–386.

[7] Brock, J. E.
In: A least squares estimate of spacecraft attitude.
SIAM Review, **8**, 3 (1966), 384–386.

[8] Markley, F. L.
Attitude determination using vector observations: A fast optimal matrix algorithm.
Journal of the Astronautical Sciences, **41**, 2 (Apr.–June 1993), 261–280.

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

[10] Bar-Itzhack, I. Y., and Markley, F. L.
Unconstrained optimal transformation matrix.
IEEE Transactions on Aerospace and Electronic Systems, **34**, 1 (Jan. 1998), 338–340.

[11] Bar-Itzhack, I. Y. and Reiner, J.
Recursive attitude determination from vector observations: DCM identification.
Journal of Guidance, Control, and Dynamics, **7** (Jan.–Feb. 1984), 51–56.

[12] Baruch, M., and Bar-Itzhack, I. Y.
Optimal weighted orthogonalization of measured modes.
AIAA Journal, **16**, 4 (1978), 346–351.

[13] Athans, M., and Tse, E.
A direct derivation of the optimal linear filter using the maximum principle.
IEEE Transactions on Automatic Control, **AC-12**, 6 (1967), 690–698.

[14] Choukroun, D., Weiss, H., Bar-Itzhack, I. Y., and Oshman, Y.
Kalman filtering for matrix estimation.
IEEE Transactions on Aerospace and Electronic Systems, **42**, 1 (Jan. 2006), 147–159.

[15] Bellman, R.
Introduction to Matrix Analysis.
New-York: McGraw-Hill, 1960.

[16] Horn, R. A., and Johnson, C. R.
Topics in Matrix Analysis.
Cambridge, U.K.: Cambridge University Press, 1991.

[17] Markley, F. L., and Mortari, D.
How to compute attitude from vector observations.
Presented at the Astrodynamics Specialist Conference, Girdwood, AK, Aug. 15–19, 1999; AIAA/AAS Paper 99-427.

[18] Golub, G. H., and Van Loan, C. F.
Matrix Computations.
Baltimore, MD: The Johns Hopkins University Press, 1983.

[19] Bar-Itzhack, I. Y., and Meyer, J.
On the convergence of iterative orthogonalization processes.
IEEE Transactions on Aerospace and Electronic Systems, **AES-12** (Jan. 1976), 146–151.

[20] Nissen, D. H.
A note on the variance of a matrix.
Econometrica, **36** (1968), 603–604.

[21] Bar-Itzhack, I. Y.
Iterative optimal orthogonalization of the strapdown matrix.
IEEE Transactions on Aerospace and Electronic Systems, **AES-11**, 11 (Jan. 1975), 30–37.

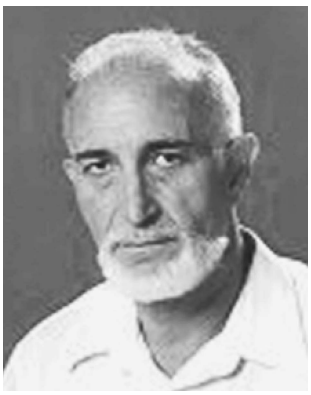
Daniel Choukroun received the B.Sc. (summa cum laude), M.Sc., and Ph.D. degrees in 2000, and 2003, respectively. He also received the title Engineer under Instruction from the Ecole Nationale de l'Aviation Civile, France, in 1994.

From 1998 to 2003, he was a teaching and research assistant in the field of automatic control at The Technion–Israel Institute of Technology. From 2003 to 2006, he has been a postdoctoral fellow and a lecturer at the University of California at Los Angeles in the Department of Mechanical and Aerospace Engineering. In 2006, he joined the Department of Mechanical Engineering at Ben-Gurion University of the Negev, Israel, where he is now lecturer and a member of the Pearlstone Center for Aeronautical Engineering Studies. He received the Miriam and Aaron Gutwirth Special Excellency Award for achievement in research from The Technion–Israel Institute of Technology.

His research interests are in optimal estimation and control theory with applications to aerospace systems. He is a member of the AIAA Technical Committee on Guidance Navigation and Control.

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 has 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) received his B.Sc. and M.Sc. degrees in electrical engineering, in 1961 and 1964, respectively, both from Technion–Israel Institute of Technology. In 1968 he received the Ph.D. degree from the University of Pennsylvania, also in electrical engineering.

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

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

Yaakov Oshman (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 of the State University of New York at Buffalo, where he was, in 1988, a visiting professor. Since 1989 he has been with the Department of Aerospace Engineering at the Technion–Israel Institute of Technology, where he is 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.

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. 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. 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 on the program committees of over a dozen international conferences.

Dr. Oshman serves as technical editor for Guidance and Control Systems for the *IEEE Transactions on Aerospace and Electronic Systems* since 2005, and as a member of the Board of Governors of the IEEE Aerospace and Electronic Systems Society since 2008. Dr. Oshman is a Fellow of the AIAA.