Title
Kalman filtering for matrix estimation

Journal
IEEE Transactions on Aerospace and Electronic Systems, 42(1)

ISSN
0018-9251

Authors
Choukroun, Daniel
Weiss, Haim
Bar-Itzhack, Itzhack Y
et al.

Publication Date
2006

Peer reviewed

Kalman Filtering for Matrix Estimation

D. CHOUKROUN, Member, IEEE
UCLA

H. WEISS, Member, IEEE

I. Y. BAR-ITZHACK, Fellow, IEEE

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

A general discrete-time Kalman filter (KF) for state matrix estimation using matrix measurements is presented. The new algorithm evaluates the state matrix estimate and the estimation error covariance matrix in terms of the original system matrices. The proposed algorithm naturally fits systems which are most conveniently described by matrix process and measurement equations. Its formulation uses a compact notation for aiding both intuition and mathematical manipulation. It is a straightforward extension of the classical KF, and includes as special cases other matrix filters that were developed in the past. Beyond the analytical value of the matrix filter, it is shown through various examples arising in engineering problems that this filter can be computationally more efficient than its vectorized version.

Manuscript received March 9, 2004; revised April 18, 2005; released for publication July 19, 2005.

IEEE Log No. T-AES/42/1/870597.

Refereeing of this contribution was handled by C. Jauffret.

Authors' current addresses: D. Choukroun, Dept. of Mechanical and Aerospace Engineering, UCLA, Los Angeles, CA, E-mail: (dach@ucla.edu); H. Weiss, RAFAEL, Ministry of Defense, PO Box 2250, Dept. 35, Haifa 31021, Israel; I. Y. Bar-Itzhack and Y. Oshman, Asher Space Research Institute, Technion–Israel Institute of Technology, Haifa, Israel, E-mail: (ibaritz@techunix.technion.ac.il).

I. INTRODUCTION

The classical Kalman filter (KF) was developed to operate on plants which are governed by vector differential/difference equations [1, 2]. In addition, the measured variable was assumed to be a vector. The description of plant dynamics and related measurement models by vector equations is a very common one; however, there are problems in which the variables are most naturally described by means of matrices, like the inertia, stiffness, and damping matrices of a given structural system. In addition, when the matrix variables are time varying, their dynamics are governed by matrix differential/difference equations; this is the case, for instance, of the direction cosine matrix (DCM) in rigid body kinematics problems [3, p. 512], and of the estimation-error covariance matrix in a Kalman-Bucy filter. Consider the problem of estimating the state of a time-varying matrix plant, where the state matrix is observed through matrix measurements. In principle, the tools for tackling this estimation problem are already available. After all, one can decompose the matrix plant into a set of vector equations and proceed with the application of the conventional KF (see [4] for the estimation of the 3×3 DCM using a 9×1 KF). One drawback of that approach, however, is that the loss of the original structure may induce a loss of physical insight into the vectorized estimation algorithm. Moreover, for a high-dimensional model, an excessive number of equations results, and it may be difficult to determine any structure and properties of the solution. These facts will be illustrated in the body of the paper through various examples taken from engineering problems.

Continuous research efforts have been conducted to develop estimation and control algorithms that naturally operate on matrix plant models. Considering the Riccati differential equation of the estimation error covariance matrix in a linear filter, the Kalman-Bucy filter gain was computed as an optimal control gain for that plant [5]. Special tools in matrix calculus, such as the gradient matrix were developed for that purpose (see [6]). Such tools were utilized to develop estimators of matrix parameters in a mechanical system [7]. In statistics, the "total least-squares" method handles the problem of matrix estimation ([8, p. 595] and [9]). This is a method of data fitting, which is based on the standard linear model (see e.g. [10, p. 9]), and which is appropriate when there are errors in both the observation vector and in the observation matrix. The issue of estimating a matrix of deterministic parameters was also addressed in [11]–[14] using, however, a probabilistic approach. A batch maximum-likelihood estimator of unknown matrices of parameters is presented in [11]. The problem of obtaining a maximum-likelihood estimate of the covariance matrix of a multivariate normal distribution was considered in [12]. A matrix version of the best linear unbiased estimator (BLUE, see [10, p. 121]) was derived in [13]. That estimator is a batch algorithm that processes a single matrix measurement, whose rows are assumed to be identically independently distributed. A special recursive least-squares algorithm, called simplified multivariate least-squares algorithm, was derived in [15, p. 96]. For special conditions, that algorithm features an elegant extension of the classical least-squares algorithm to the matrix case.

The present work is a continuation of the aforementioned research efforts. Here, we consider a stochastic linear time-varying discrete-time plant with a state matrix observed by matrix measurements. The state dynamics is described by a stochastic difference equation with an additive zero-mean Gaussian white noise matrix. The measurement is modeled as a linear matrix-valued function of the state matrix with an additive zero-mean white noise matrix. All matrices are defined over the real field. We propose a general state matrix Kalman filter (MKF) for this plant; that is, a KF that provides a sequence of optimal estimates in a matrix format and performs a covariance analysis of the estimation error. The MKF has the statistical properties of the ordinary KF while retaining the advantages of a compact matrix notation by expressing the estimated matrix in terms of the original plant parameters. Along with the development of the MKF, several examples will be provided that illustrate the analytical and numerical advantages of the MKF over its vectorized version.

The structure of the paper is as follows. Section II presents the mathematical formulation of the estimation problem. Section III contains the derivation of the MKF, followed by a summary of the algorithm and a discussion of its structure. Section IV presents several comparative examples taken from engineering problems, which illustrate the notational advantage of the MKF. Section V discusses the numerical advantage of the MKF over the vectorized filter. In Section VI, an MKF is designed in the context of spacecraft attitude determination from vector measurements. Conclusions are presented in Section VII.

II. PROBLEM FORMULATION

A. Matrix State-Space Model

Consider the general linear discrete-time stochastic dynamic system (the plant) governed by the difference equation

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

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

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

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 [16, p. 244]. As an example, if $W \in \mathbb{R}^{3\times 3}$, such that $W = [\mathbf{w}_{c1}\ \ \mathbf{w}_{c2}\ \ \mathbf{w}_{c3}]$ then

$$\operatorname{vec} W \triangleq \begin{bmatrix} \mathbf{w}_{c1} \\ \mathbf{w}_{c2} \\ \mathbf{w}_{c3} \end{bmatrix} \in \mathbb{R}^9. \tag{3}$$

The covariance matrix of W is defined as the covariance of its vec-transform [13]; that is $\operatorname{cov}\{W\} \triangleq \operatorname{cov}\{\operatorname{vec} W\}$. Obviously, if $W \in \mathbb{R}^{m\times n}$, then $\operatorname{cov}\{W\} \in \mathbb{R}^{mn\times mn}$.

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 equation of any linear matrix plant is a special case of (1). This is so because the summation operation, together with the left and right multiplication on X_k in (1), enable to write each element of X_{k+1} as a linear combination of all the elements of X_k (plus an additive scalar noise). The proof, although straightforward, is lengthy, and is omitted here for the sake of brevity. A detailed proof is provided in [17, pp. 250–253]. In general μ might be equal to $(mn)^2$. However, certain plants may need fewer terms in the system equation (1). Similarly, (2) can be shown to be a standard form for a linear matrix measurement model of the matrix plant. In the general case $\nu = mnpq$, but certain measurement models may include fewer terms.

B. 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, $\{Y_l\}$, $l = 1\ldots k$. Let $\hat{X}_{k/k}$ and $\tilde{X}_{k/k}$ denote, respectively, the a posteriori state estimate and the a posteriori estimation error; that is, $\tilde{X}_{k/k} \triangleq X_k - \hat{X}_{k/k}$. Let $P_{k/k}$ denote the a posteriori estimation error covariance matrix; that is $P_{k/k} = \text{cov}(\tilde{X}kk)$. Then, the filtering problem is equivalent to the following unbiased minimization problem

$$\min_{X_k}\{\text{tr}(P_{k/k})\} \tag{4}$$

subject to (1) and (2) and to the stochastic assumptions on the noises and the initial conditions. When the state and the measurement are vectors, the solution of the above minimization problem leads to the standard KF.

III. STATE-MATRIX KALMAN FILTER

A. Derivation Approach

The approach to the filter derivation consists of three principal steps. Note that these steps are not the computation steps of the filter itself, which are summarized later in this section. In the first step we apply the vec-operator on the matrix plant described by (1) and (2). Thus, the matrix equations (1) and (2) are transformed into equivalent vector equations. The result of this step is a standard state-space model, the vec-system, where the state is an mn-dimensional vector, and the measurement is a pq-dimensional vector.

In the second step, KF theory is applied to the vec-system. The time update and measurement update stages are developed using the classical KF algorithm. As a result, we obtain two sets of equations: the first set computes the mn-dimensional state estimate of the vec-system and the second set computes the $mn \times mn$ estimation error covariance matrix.

The third step aims at retrieving the matrix form of the original problem. This is done by applying the inverse of the vec-operator to the KF of the vec-system. Thus, the state estimate equations, which are mn-dimensional vector equations, are transformed into equivalent $m \times n$ matrix equations. As a result, a matrix innovations sequence is naturally defined and the time update and measurement update stages are expressed in terms of matrices. By definition, the covariance computations of the vec-system and of the matrix system are identical. Thus the covariance analysis performed in the second step is left in its current extended form. These manipulations yield a state-matrix KF in a compact notation.

B. Filter Derivation

It is emphasized that the manipulations involved in computing the various vec-transforms and their reverse are part of the MKF derivation but not of the filter implementation. In the following we consistently denote the vec-transform of a matrix W by the associated bold lower case symbol $\mathbf{w}$; that is, $\mathbf{w} \triangleq \text{vec}\,W$. The following property of the vec-operator will be intensively used. Let $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{n\times p}$, and $B \in \mathbb{R}^{p\times q}$, and let $\otimes$ denote the Kronecker product, also called direct product, or tensor product of two matrices [16, p. 243], then [16, p. 255]

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}\,X. \tag{5}$$

1) *Time Update*: Assume that $\hat{X}_{k/k}$ and $P_{k/k}$ have been computed at time t_k. Applying the vec-operator to (1), using its linear property and (5), yields

$$\text{vec}\,X_{k+1} = \left\{\sum_{r=1}^{\mu}[(\Psi_k^r)^T \otimes \Theta_k^r]\right\}\text{vec}\,X_k + \text{vec}\,W_k \tag{6}$$

which is equivalent to

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k \tag{7}$$

where the terms in (7) are obviously defined from (6). Equation (7) is the process equation of the vec-system. Applying the time update stage of the KF [10, p. 228] to the vec-system yields

$$\hat{\mathbf{x}}_{k+1/k} = \Phi_k \hat{\mathbf{x}}_{k/k}. \tag{8}$$

Thus,

$$\text{vec}\,\hat{X}_{k+1/k} = \left\{\sum_{r=1}^{\mu}[(\Psi_k^r)^T \otimes \Theta_k^r]\right\}\text{vec}\,\hat{X}_{k/k} \tag{9}$$

$$\hat{X}_{k+1/k} = \sum_{r=1}^{\mu}\Theta_k^r \hat{X}_{k/k}\Psi_k^r \tag{10}$$

where (10) is obtained using again the linear property of the vec-operator and (5) (but this time in the inverse direction). Notice from (10) that the a priori estimate at time t_{k+1} is computed using only the matrix-plant parameters, Θ_k^r and Ψ_k^r, (see (1)), and the a posteriori matrix estimate at time t_k. Recalling that, by definition, the covariance matrix of a matrix random variable is the covariance matrix of its vec-transform, the time update stage for the estimation error covariance matrix $P_{k/k} \in \mathbb{R}^{mn\times mn}$ is computed from [10, p. 229]

$$P_{k+1/k} = \Phi_k P_{k/k}\Phi_k^T + Q_k \tag{11}$$

where $Q_k = \text{cov}\{W_k\}$.

2) *Measurement Update*: Assume that $\hat{X}_{k+1/k}$ and $P_{k+1/k}$ have been computed, and that a new matrix measurement Y_{k+1} is performed at time t_{k+1}. Proceeding similarly to the previous subsection, it is straightforward to show that the three following

equations are equivalent

$$\tilde{\mathbf{y}}_{k+1} = \mathbf{y}_{k+1} - \mathcal{H}_{k+1}\hat{\mathbf{x}}_{k+1/k} \tag{12}$$

$$\mathrm{vec}\,\tilde{Y}_{k+1} = \mathrm{vec}\,Y_{k+1} - \left\{\sum_{s=1}^{\nu}[(G_{k+1}^{s})^{T} \otimes H_{k+1}^{s}]\right\}\mathrm{vec}\,\hat{X}_{k+1/k} \tag{13}$$

$$\tilde{Y}_{k+1} = Y_{k+1} - \sum_{s=1}^{\nu} H_{k+1}^{s}\hat{X}_{k+1/k}G_{k+1}^{s} \tag{14}$$

where $\tilde{\mathbf{y}}_{k+1} \in \mathbb{R}^{pq}$ and $\tilde{Y}_{k+1} \in \mathbb{R}^{p\times q}$, respectively, denote the innovations sequences in the vec-KF and in the MKF, and $\mathcal{H}_{k+1}$ is defined from (13). The point here is that a natural expression for the matrix innovations sequence $\tilde{Y}_{k+1}$ is obtained as a function of the original plant coefficients H_{k+1}^{s} and G_{k+1}^{s} (see (2)). According to KF theory, the $pq \times pq$ covariance matrix of $\tilde{Y}_{k+1}$, denoted by S_{k+1}, the $mn \times pq$ KF gain matrix, and the a posteriori vec-estimate, $\hat{\mathbf{x}}_{k+1/k+1}$ are computed as [10, p. 246]

$$S_{k+1} = \mathcal{H}_{k+1}P_{k+1/k}\mathcal{H}_{k+1}^{T} + R_{k+1} \tag{15}$$

$$K_{k+1} = P_{k+1/k}\mathcal{H}_{k+1}S_{k+1}^{-1} \tag{16}$$

$$\hat{\mathbf{x}}_{k+1/k+1} = \hat{\mathbf{x}}_{k+1/k} + K_{k+1}\tilde{\mathbf{y}}_{k+1} \tag{17}$$

where $R_k = \mathrm{cov}\{V_k\}$. In order to recover the matrix format for the state estimate update equation (17), we use the following proposition.

PROPOSITION 1 *Let $X \in \mathbb{R}^{i_3\times i_4}$, $Z \in \mathbb{R}^{i_1\times i_2}$, and $A \in \mathbb{R}^{i_1 i_2\times i_3 i_4}$ be given matrices. Let $\mathbf{z} \in \mathbb{R}^{i_1 i_2}$ and $\mathbf{x} \in \mathbb{R}^{i_3 i_4}$ be defined by $\mathbf{z} = \mathrm{vec}\,Z$ and $\mathbf{x} = \mathrm{vec}\,X$, respectively. Let the block-matrix A be partitioned as*

$$A = \begin{bmatrix} A^{11} & A^{12} & \cdots & A^{1l} & \cdots & A^{1i_4} \\ A^{21} & A^{22} & \cdots & A^{2l} & \cdots & A^{2i_4} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A^{j1} & A^{j2} & \cdots & A^{jl} & \cdots & A^{ji_4} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ A^{i_2 1} & A^{i_2 2} & \cdots & A^{i_2 l} & \cdots & A^{i_2 i_4} \end{bmatrix} \tag{18}$$

where $A^{jl} \in \mathbb{R}^{i_1 i_3}$, then the vector equation

$$\mathbf{z} = A\mathbf{x} \tag{19}$$

is equivalent to the matrix equation

$$Z = \sum_{j=1}^{i_2}\sum_{l=1}^{i_4} A^{jl}XE^{lj} \tag{20}$$

where E^{lj} is an $i_4 \times i_2$ matrix with 1 at position (lj) and 0 elsewhere.

The proof is omitted here for the sake of brevity. It can be found in [17, pp. 254–257].

The matrix K_{k+1} in (16) can be viewed as an $mn \times pq$ array of nq submatrices of dimension $m \times p$. Let K^{jl} denote the $m \times p$ submatrix of K_{k+1} at position (jl). Then, applying Proposition 1 to the second term on the right-hand side (RHS) of (17) with $A = K_{k+1}$ and $\mathbf{x} = \mathrm{vec}\,\tilde{Y}_{k+1}$ yields

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{n}\sum_{l=1}^{q} K_{k+1}^{jl}\tilde{Y}_{k+1}E^{lj}. \tag{21}$$

Equation (21) is the measurement update equation for the matrix state estimate at time t_{k+1}. From KF theory [10, p. 245], omitting the time subscript for clarity, the covariance matrix of the a posteriori estimation error at time t_{k+1} is

$$P_{k+1/k+1} = (I_{mn} - K\mathcal{H})P_{k+1/k}(I_{mn} - K\mathcal{H})^{T} + KRK^{T}. \tag{22}$$

C. Summary of the Matrix Kalman Filter

Initialization:

$$\hat{X}_{0/0} = \bar{X}_0 \tag{23}$$

$$P_{0/0} = \Pi_0. \tag{24}$$

Time Update equations:

$$\hat{X}_{k+1/k} = \sum_{r=1}^{\mu} \Theta_k^r \hat{X}_{k/k}\Psi_k^r \tag{25}$$

$$\Phi_k = \sum_{r=1}^{\mu}[(\Psi_k^r)^T \otimes \Theta_k^r] \tag{26}$$

$$P_{k+1/k} = \Phi_k P_{k/k}\Phi_k^T + Q_k. \tag{27}$$

Measurement Update equations:

$$\tilde{Y}_{k+1} = Y_{k+1} - \sum_{s=1}^{\nu} H_{k+1}^{s}\hat{X}_{k+1/k}G_{k+1}^{s} \tag{28}$$

$$\mathcal{H}_{k+1} = \sum_{s=1}^{\nu}[(G_{k+1}^{s})^{T} \otimes H_{k+1}^{s}] \tag{29}$$

$$S_{k+1} = \mathcal{H}_{k+1}P_{k+1/k}\mathcal{H}_{k+1}^{T} + R_{k+1} \tag{30}$$

$$K_{k+1} = P_{k+1/k}\mathcal{H}_{k+1}^{T}S_{k+1}^{-1} \tag{31}$$

$$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{n}\sum_{l=1}^{q} K_{k+1}^{jl}\tilde{Y}_{k+1}E^{lj} \tag{32}$$

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

$$K_{k+1} = \left.\underbrace{\begin{bmatrix} K_{k+1}^{11} & \cdots & K_{k+1}^{1l} & \cdots \\ \vdots & \ddots & \vdots & \ddots \\ K_{k+1}^{j1} & \cdots & K_{k+1}^{jl} & \cdots \\ \vdots & \ddots & \vdots & \ddots \end{bmatrix}}_{q \text{ matrices}}\right\} n \text{ matrices} \tag{33}$$

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

$$P_{k+1/k+1} = (I_{mn} - K\mathcal{H})P_{k+1/k}(I_{mn} - K\mathcal{H})^T + KRK^T. \tag{34}$$

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

$$\mathrm{var}\{\tilde{X}[i,j]\} = P[(j-1)m+i,(j-1)m+i] \tag{35a}$$

$$\mathrm{cov}\{\tilde{X}[i,j],\tilde{X}[k,l]\} = P[(j-1)m+i,(l-1)m+k] \tag{35b}$$

where $i,k = 1\ldots m$, and $j,l = 1\ldots n$. Equations (35) are proved using the definitions of the vec-operator and of the covariance matrix of a random matrix. The variable $\tilde{X}$ denotes either the a posteriori or the a priori estimation error as applicable, and P is the associated covariance matrix.

D. Discussion

The proposed filter 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 $n = q = 1$ the proposed algorithm yields exactly the KF. When $n \neq 1$ or $q \neq 1$ the MKF provides us with equations where the state matrix estimate is propagated and updated as a matrix using the coefficients of the matrix plant. Therefore, the estimation algorithm preserves the physical insight into the original plant.

The MKF includes as special cases two other matrix estimators presented in the literature, the matrix BLUE [13], and the simplified multivariate least-squares (SMLS) algorithm [15, p. 96]. Consider the special case of a time-invariant system with a time-invariant state-matrix, where the measurement model is described by $Y = HX + V$, H is full-rank, the rows of V are independently identically distributed (IID), and there is no a priori estimate, then developing an MKF using these assumptions leads to the matrix BLUE of [13]. The proof is detailed in [17, pp. 257–259]. The SMLS algorithm can be derived as a special case of the MKF algorithm as follows. Assume that the system is time invariant, the measurement is a row-matrix, which is expressed as $\mathbf{y}_{k+1}^T = \mathbf{h}_{k+1}^T X + \mathbf{v}_{k+1}^T$, the covariance matrix of the measurement noise $\mathbf{v}_{k+1}$ is the identity matrix, and the columns of the initial state matrix estimate are IID. Then, developing an MKF based on the above assumptions yields the SMLS algorithm. A detailed proof is given in [17, pp. 260–262].

IV. COMPARATIVE EXAMPLES

In this section, we present several examples taken from engineering problems, which illustrate analytically and numerically the advantage of the MKF over the vectorized filter. Note that, for the sake of brevity, the necessary developments leading to the matrix equations are only outlined.

A. Three-Dimensional Tracking Filter

This example is taken from a guidance problem, as presented in [18], with an intercepting missile heading towards an accelerating target, where the authors addressed the problem of estimating the relative position vector, the relative velocity vector, and the inertial target acceleration expressed along the interceptor sensor coordinate frame. Denoting the three-dimensional vectors by $\mathbf{p}_k$, $\mathbf{v}_k$, and $\mathbf{a}_k$, respectively, the authors defined a 3×3 state matrix X_k as follows:

$$X_k \stackrel{\Delta}{=} [\mathbf{p}_k \quad \mathbf{v}_k \quad \mathbf{a}_k] \tag{36}$$

and, using (36), developed the following discrete-time process matrix equation:

$$X_{k+1} = C(k+1,k)X_k F^T + W_k. \tag{37}$$

In (37), $C(k+1,k)$ is a 3×3 rotation matrix which is due to the rotation of the interceptor sensor coordinate frame, F is a 3×3 transformation matrix that is only a function of the discretization step size, that is, $t_{k+1} - t_k$, and W_k is a matrix of process noises. If the matrix equation (37) is vectorized, this will lead to the following 9×9 vector equation:

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k \tag{38}$$

where Φ_k is the 9×9 matrix

$$\Phi_k = F \otimes C(k+1,k) \tag{39}$$

and $\mathbf{x}_k$ and $\mathbf{w}_k$ are the vec-transforms of X_k and W_k, respectively. Alternatively, the system dynamics can be represented by three explicit 3×1 vector equations as follows:

$$\mathbf{p}_{k+1} = C(k+1,k)(f_{11}\mathbf{p}_k + f_{21}\mathbf{v}_k + f_{31}\mathbf{a}_k) + \mathbf{w}_k^p \tag{40a}$$

$$\mathbf{v}_{k+1} = C(k+1,k)(f_{12}\mathbf{p}_k + f_{22}\mathbf{v}_k + f_{33}\mathbf{a}_k) + \mathbf{w}_k^v \tag{40b}$$

$$\mathbf{a}_{k+1} = C(k+1,k)(f_{13}\mathbf{p}_k + f_{23}\mathbf{v}_k + f_{33}\mathbf{a}_k) + \mathbf{w}_k^a. \tag{40c}$$

Comparing the various expressions for the process equations (37) to (40), we see that (37) presents the advantage of synthetically describing the system dynamics through a single matrix equation. Also, it preserves the physical insight into the problem: the state equations for $\mathbf{p}_k$, $\mathbf{v}_k$, and $\mathbf{a}_k$, are written in the coordinate frame that is associated with the seeker sensor. From the physics of the problem, the seeker is rotating between two consecutive times t_k and t_{k+1}, and to account for this rotation, the three columns of the state matrix need to be projected onto the new seeker sensor's frame at t_{k+1}. Mathematically, this is conveniently done by left-multiplying the state matrix by the seeker's frame incremental rotation matrix $C(k+1,k)$, as in (37). Not only is the physical insight masked in (38) and (39) because of the use of the Kronecker product, but the vectorization also induces more computations in the associated filter. The latter fact will be emphasized and illustrated in the next section. Although (40) retain the physical insight, they have the drawback of losing the simple structure inherent to (37).

B. Quaternion Estimation

In aerospace systems a crucial problem is that of determining the spatial orientation of one Cartesian coordinate frame, say $\mathcal{B}$, with respect to another one, say $\mathcal{R}$; this problem is known as the attitude determination problem. One way of representing attitude is by means of the four Euler parameters, which are the components of a unit-norm vector called the quaternion of rotation, $\mathbf{q} \in \mathbb{R}^4$. It is well known that the quaternion dynamics can be modeled as the following process equation (see e.g. [3, pp. 511–512]):

$$\mathbf{q}_{k+1} = \Phi_k \mathbf{q}_k + \mathbf{w}_k \tag{41}$$

where Φ_k is a 4×4 matrix that is computed as

$$\Phi_k = \exp\left\{\frac{1}{2}\int_{t_k}^{t_{k+1}} \Omega_t dt\right\} \tag{42}$$

in which

$$\Omega_t = \begin{bmatrix} -[\boldsymbol{\omega}_t \times] & \boldsymbol{\omega}_t \\ -\boldsymbol{\omega}_t^T & 0 \end{bmatrix} \tag{43}$$

and $\boldsymbol{\omega}_t \in \mathbb{R}^3$ in (43) is the angular velocity vector of $\mathcal{B}$ with respect to $\mathcal{R}$ resolved in $\mathcal{B}$ and $[\boldsymbol{\omega}_t \times]$ is the skew-symmetric cross-product matrix. Usually $\boldsymbol{\omega}_t$ is the noisy output of a triad of gyros and $\mathbf{w}_k$ is a process noise vector that models the uncertainty due to the gyro noises. Also note the clear interpretation of the first term on the RHS of (41); it accounts for a rotation in $\mathbb{R}^4$ of the quaternion via the orthogonal transition matrix Φ_k. One very popular approach to the optimal estimation of the quaternion consists in extracting the normalized eigenvector that belongs to the largest positive eigenvalue of a special four-dimensional matrix K (see [3, pp. 426–428]). In a previous work [19], it was proposed to optimally estimate the elements of the matrix K itself, before feeding them to an eigenvector solver, and using the results of [20], the process equation for that matrix was written as a 4×4 matrix equation, as follows:

$$K_{k+1} = \Phi_k K_k \Phi_k^T + W_k \tag{44}$$

where the 4×4 matrix Φ_k is identical to that of (43), and W_k is a process noise matrix. Note that one can easily interpret the similarity transformation in the first term on the RHS as the underlying "rotation" of the system of eigenvectors, with one of them being the sought quaternion. If (44) is vectorized, this will lead to the following 16×1 vector equation:

$$\mathrm{vec}\, K_{k+1} = (\Phi_k \otimes \Phi_k)\mathrm{vec}\, K_k + \mathrm{vec}\, W_k. \tag{45}$$

In (45) the structure is lost, and the first term on the RHS is not straightforwardly interpreted. In addition, the filter computational burden will increase (see the next section).

C. Direction Cosine Matrix Estimation

Another example, also from the realm of attitude determination, illustrates the advantage of the matrix notation. Here, the DCM, denoted by D_k, is used as the attitude representation of a rotating frame $\mathcal{B}$ with respect to a reference frame $\mathcal{R}$. In this case the process equation for D_k is represented as the following 3×3 matrix equation:

$$D_{k+1} = \Phi_k D_k + W_k \tag{46}$$

where Φ_k is the 3×3 matrix

$$\Phi_k = \exp\left\{\int_{t_k}^{t_{k+1}} (-[\boldsymbol{\omega}_t \times])dt\right\} \tag{47}$$

and $\boldsymbol{\omega}_t$ is the angular velocity vector mentioned in the previous example. The first term on the RHS in (46) has a clear physical interpretation as it expresses the rotation of $\mathcal{B}$ with respect to $\mathcal{R}$. This physical insight is lost in the vectorized version of (46), which is written as follows:

$$\mathrm{vec}\, D_{k+1} = [I_9 \otimes \Phi_k]\mathrm{vec}\, D_k + \mathrm{vec}\, W_k. \tag{48}$$

As stated earlier, the use of the 9×9 transition matrix in (48) increases the filter computational burden. Besides the foregoing advantages, keeping the model equation in a matrix form allows to incorporate into the MKF the nonlinear orthogonality constraint on the DCM estimate in a straightforward manner. The approach consists in modeling the orthogonality constraint as a pseudomeasurement equation. Starting from the conventional orthogonality constraint equation, that is,

$$D_k^T D_k = I_3 \tag{49}$$

where $I_3 \in \mathbb{R}^3$ is the identity matrix, one can develop the equivalent equation

$$D_k = (\tfrac{3}{2} I_3 - \tfrac{1}{2} D_k^T D_k) D_k. \tag{50}$$

Note that using (49) in (50) yields the trivial identity $D_k = D_k$. The choice of (50) is not unique. In this case, it is inspired by an orthogonalization iterative formula introduced in [21]. Denoting the current estimate of D_k by $\hat{D}_k$, and using it in (50), yields

$$\hat{D}_k = (\tfrac{3}{2} I_3 - \tfrac{1}{2} \hat{D}_k^T \hat{D}_k) D_k + V_k^{ort} \tag{51}$$

where V_k^{ort} is a 3×3 modeling error matrix due to the errors in $\hat{D}_k$. The sequence V_k^{ort} is modeled, for convenience, as a zero-mean white noise sequence with covariance matrix $\eta_k^{ort} I_9$. Upon defining the following matrices:

$$Z_k \stackrel{\Delta}{=} \hat{D}_k \tag{52a}$$

$$\mathcal{H}_k \stackrel{\Delta}{=} (\tfrac{3}{2} I_3 - \tfrac{1}{2} \hat{D}_k^T \hat{D}_k) \tag{52b}$$

one can rewrite (51) as follows

$$Z_k = \mathcal{H}_k D_k + V_k^{ort}. \tag{53}$$

Equation (53) is a pseudomeasurement model equation where the pseudomeasurement is Z_k, the observation matrix is $\mathcal{H}_k$ and the noise is V_k^{ort}. This equation can readily be incorporated into the MKF framework. The associated measurement-update stage is such that it produces a new estimate by enforcing the orthogonality property on the current one. This is done via the choice of the filter parameter η_k^{ort}, which is a design parameter and requires tuning: a high value of η_k^{ort} will induce soft orthogonalization, and a low value of η_k^{ort} will induce a hard orthogonalization. It was shown in a previous work [22] that this type of constrained Kalman filtering produces smooth transients and avoids undesirable overshoots of the estimation errors. Since the orthogonality property is naturally written as a matrix equation, this greatly facilitates the analysis of the filter.

The decomposition of the matrix equation (49) into scalar equations yields the following six scalar constraints:

$$d_{11}^2 + d_{12}^2 + d_{13}^2 = 1 \tag{54a}$$

$$d_{21}^2 + d_{22}^2 + d_{23}^2 = 1 \tag{54b}$$

$$d_{31}^2 + d_{32}^2 + d_{33}^2 = 1 \tag{54c}$$

$$d_{11}d_{21} + d_{12}d_{22} + d_{13}d_{23} = 0 \tag{54d}$$

$$d_{11}d_{31} + d_{12}d_{32} + d_{13}d_{33} = 0 \tag{54e}$$

$$d_{21}d_{31} + d_{22}d_{32} + d_{23}d_{33} = 0 \tag{54f}$$

where d_{ij}, for $i, j = 1,2,3$, denotes the components of D_k. This results in a multiplication of equations, a loss of structure and, thus, of insight into the resulting filter. The higher the dimension n of the state matrix, the higher is the number of equations: $n(n+1)/2$. In addition, we notice that vectorizing equation (53) is possible, but would yield filter equations where both vectorized and matrix versions of the DCM estimates are involved. As a result, the analysis of the vectorized filter would not be as straightforward as that of the MKF.

TABLE I
Counts of FLOPs for Basic Matrix Operations [10, p. 258]

Operation	Multiplications	Additions
$A_{MN} \pm B_{MN}$	—	MN
$A_{MN}\ B_{NL}$	MNL	$ML(N-1)$

V. COMPUTATIONAL PERFORMANCE

A comparison of the number of floating point operations (FLOPs) per cycle in the MKF and in the vectorized KF (VKF) is presented here. One cycle consists of one time-propagation stage and one measurement-update stage. The computation requirement in FLOPs is based on Table I. For comparative purposes, we only consider the computations that are not common to both filters, and that consist of matrix algebraic operations. This excludes, for instance, matrix manipulations such as column extraction or concatenation. The specific equations of interest are summarized in Table II. For convenience, we recall in Table III the sizes of the various variables involved in the MKF and in the vectorized filter. Using Tables I–III to count the number of FLOPs per cycle in each filter, we obtain Table IV. Note that the general formulas for the MKF are only valid for $n > 1$ and $q > 1$. For the cases where $n = 1$ or $q = 1$, the matrix state-space model equations would not involve matrices like Ψ_k^r or G_{k+1}^s. Indeed, being simply scalars, these matrices would be incorporated in the other model variables; that is, Θ_k^r and H_{k+1}^s respectively, yielding the conventional vector model equations.

REMARK The count for the measurement-update stage in the MKF is based on a coding that exploits

TABLE II
Description of Equations that are Not Common to MKF and to Vectorized Filter

Stage	MKF Equations	Vectorized Filter Equations
Time-Propagation	$\hat{X}_{k+1/k} = \sum_{r=1}^{\mu} \Theta_k^r \hat{X}_{k/k} \Psi_k^r$	$\hat{\mathbf{x}}_{k+1/k} = \Phi_k \hat{\mathbf{x}}_{k/k}$
Residuals	$\tilde{Y}_{k+1} = Y_{k+1} - \sum_{s=1}^{\nu} H_{k+1}^s \hat{X}_{k+1/k} G_{k+1}^s$	$\tilde{\mathbf{y}}_{k+1} = \mathbf{y}_{k+1} - \mathcal{H}_{k+1} \hat{\mathbf{x}}_{k+1/k}$
Measurement-Update	$\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \sum_{j=1}^{n} \sum_{l=1}^{q} K_{k+1}^{jl} \tilde{Y}_{k+1} E^{lj}$	$\hat{\mathbf{x}}_{k+1/k+1} = \hat{\mathbf{x}}_{k+1/k} + K_{k+1} \tilde{\mathbf{y}}_{k+1}$

TABLE III
Sizes of Variables Involved in MKF and in Vectorized Filter

MKF Variables		Vectorized Filter Variables	
$\hat{X}_{k/k}$, $\hat{X}_{k+1/k}$	$m \times n$	$\hat{\mathbf{x}}_{k/k}$, $\hat{\mathbf{x}}_{k+1/k}$	$mn \times 1$
Θ_k^r	$m \times m$	Φ_k	$mn \times mn$
Ψ_k^r	$n \times n$		
Y_{k+1}, $\tilde{Y}_{k+1}$	$p \times q$	$\mathbf{y}_{k+1}$, $\tilde{\mathbf{y}}_{k+1}$	$pq \times 1$
H_{k+1}^s	$p \times m$	$\mathcal{H}_{k+1}$	$pq \times mn$
G_{k+1}^s	$n \times q$		
K_{k+1}^{jl}	$m \times p$	K_{k+1}	$mn \times pq$

TABLE IV
Number of FLOPs Per Cycle in MKF and in Vectorized Filter

Stage	MKF Equations	Vectorized Filter Equations
Time-Propagation	$2\mu m^2 n + 2\mu m n^2 - \mu mn + \mu m - m$	$2m^2n^2 - mn$
Residuals	$2\nu mpq + 2\nu mnq - \nu mq$	$2mnpq$
Measurement-Update	$2mnpq - mnq + mn - mq + m$	$2mnpq - mn$

Note: Count is limited to noncommon equations.

the structure of the update formula, as given in the third row of Table II. The steps are the following:

1) Extract the lth column from $\tilde{Y}_{k+1}$, for $l = 1,2,\ldots,q$, and denote it by $\tilde{Y}_{Cl}$.

2) Compute the vectors $\mathbf{u}(l,j) = K_{k+1}^{jl}\tilde{Y}_{Cl}$, for $l = 1,2,\ldots,q$, and $j = 1,2,\ldots,n$. Each vector $\mathbf{u}(l,j)$ is the partial correction to the jth column of $\hat{X}_{k+1/k}$ contributed by the lth column of $\tilde{Y}_{k+1}$.

3) Denote by $\Delta\hat{X}$ the total correction term in the estimate update formula. For each column j of $\Delta\hat{X}$, denoted by $\Delta\hat{X}_{Cj}$, add all the partial corrections over all the columns of $\tilde{Y}_{k+1}$; that is, do: $\Delta\hat{X}_{Cj} = \sum_{l=1}^{q}\mathbf{u}(l,j)$, for $j = 1,2,\ldots,n$.

4) Build the matrix $\Delta\hat{X}$ and add it to the old estimate; that is, do: $\hat{X}_{k+1/k+1} = \hat{X}_{k+1/k} + \Delta\hat{X}$.

Due to the presence of the parameters μ and ν in the matrix model only, a comparison of the expressions for the MKF and the vectorized filter is, in general, difficult. Nevertheless, as illustrated through the various examples brought in the preceding section, there is a motivation to consider the particular cases where μ and ν are small, at least with respect to the dimensions of the model variables. In that case, Table IV shows that there is a computational advantage to the MKF over the vectorized filter in the time-propagation stage and in the Residuals computation. Both filters have, however, similar complexity for the estimate update formula. To facilitate a numerical comparison, we introduce the relative computational efficiency ratio of the MKF with respect to the vectorized filter, denoted by ρ, and defined as follows:

$$\rho \triangleq \frac{N_{\mathrm{VKF}} - N_{\mathrm{MKF}}}{N_{\mathrm{VKF}}} \cdot 100 \tag{55}$$

where N_{MKF} and N_{VKF} are obtained by summing the operations needed for computing the expressions in the three rows of Table IV, for the matrix filter and for the vectorized filter, respectively. Various numerical results are summarized in Table V for $\mu = \nu = 1$. Case A is relevant to an estimator of the DCM involving the orthogonality pseudomeasurement model. In that case, using the MKF yields a value of 30% for the efficiency ratio ρ. In the example involving the estimation of the 4×4 K-matrix using 4×4 matrix observations of that matrix state, the ratio increases to 40%. Case C is relevant to an estimator of the DCM processing 3×1 vector measurements only. In that case, the relative advantage of using the MKF increases to 44%. A matrix estimator of the DCM equally using a vector measurement model and an orthogonality pseudomeasurement model (see [22]) has a relative efficiency ratio of 37%.

TABLE V
Relative Computational Efficiency of MKF for Particular Sizes of State Matrix and of Measured Matrix

Case	m	n	p	q	ρ (%)
A	3	3	3	3	30%
B	4	4	4	4	40%
C	3	3	3	1	44%
D	100	100	100	100	67%
E	100	100	1	1	98%

Case D addresses a large scale hypothetical estimation problem where both the state and the observation matrices are 100×100. It appears that the relative efficiency reaches an upper bound of 67%. If the observations are sequentially processed as scalars (case E in Table V), then the relative efficiency of the MKF is even greater and reaches 98%. The above results illustrate the relative numerical efficiency of the MKF with respect to the vectorized version in some cases that are motivated by engineering problems.

VI. NUMERICAL EXAMPLE

The numerical example that is proposed in this section belongs to the field of attitude determination (AD) from vector measurements. As mentioned earlier, AD algorithms from vector measurements usually consist of two stages (see e.g. [3, pp. 426–428]); in the first stage the measurement data is collected in a convenient matrix format, and in the second stage the attitude is extracted by some numerical method. We focus here on the first stage by designing an MKF that filters noise out of the matrix of measurements. This MKF is thus not an attitude estimator per se, but rather a "prefilter" of the measurements. Consider two Cartesian coordinate frames, $\mathcal{B}$ and $\mathcal{R}$, where $\mathcal{B}$ is attached to the spacecraft body, and $\mathcal{R}$ is a given inertial reference frame. Let the representations in $\mathcal{B}$ and $\mathcal{R}$ of any physical vector, like the Earth magnetic field, or the line of sight unit vector from the spacecraft to a celestial object, be denoted by $\mathbf{b}^o$ and $\mathbf{r}$, respectively; then, $\mathbf{b}^o$ and $\mathbf{r}$ are related by

$$\mathbf{b}^o = D\mathbf{r} \tag{56}$$

where D is the DCM matrix [3, p. 410]. The reference vector $\mathbf{r}$ is usually accurately known from tables or almanac, while $\mathbf{b}$ is measured on-board the spacecraft with some error. Denoting by $\mathbf{v}$ the measurement error, and assuming that N simultaneous measurements are performed at time t_k, yields

$$[\mathbf{b}_{k,1} \cdots \mathbf{b}_{k,N}] = [\mathbf{b}^o_{k,1} \cdots \mathbf{b}^o_{k,N}] + [\mathbf{v}_{k,1} \cdots \mathbf{v}_{k,N}] \tag{57}$$

which is equivalently written as

$$Y_k = X_k + V_k. \tag{58}$$

The matrices X_k, Y_k, and V_k, all in $\mathbb{R}^{3\times N}$, are obviously defined by (57). Equation (57) represents the matrix measurement equation of a plant, which is represented by the state matrix X_k. The goal is to design an MKF that performs a regression on the matrix measurements Y_k in order to optimally estimate X_k. The matrix state-space model for X_k is completed by a process equation as follows. The kinematics law of a rigid body in terms of the attitude matrix is described by the well-known difference equation [3, p. 512]

$$D_{k+1} = \Theta^o_k D_k \tag{59}$$

$$\Theta^o_k = \exp\{-[\boldsymbol{\omega}^o_k \times]\,\Delta t\} \tag{60}$$

where $\boldsymbol{\omega}^o_k$ is the true angular velocity vector of $\mathcal{B}$ with respect to $\mathcal{R}$, resolved in $\mathcal{B}$, $[\boldsymbol{\omega}^o_k \times]$ denotes the cross-product matrix of $\boldsymbol{\omega}^o_k$; that is, for $\mathbf{u} \in \mathbb{R}^3$, $[\boldsymbol{\omega}^o_k \times]\mathbf{u} \triangleq \boldsymbol{\omega}^o_k \times \mathbf{u}$, and Δt is the time increment, $\Delta t \triangleq t_{k+1} - t_k$. We assume here that, for each vector measurement, the rate of change of the associated reference vector $\mathbf{r}$ is negligible. This is the case when directions to stars are measured because stars are assumed fixed in an inertial reference frame. Furthermore, it is assumed that the same stars are observed at each epoch time. This happens for a spacecraft with an inertial-stabilized attitude, such that the same portion of the celestial sphere can be observed over time. The angular velocity vector is measured by a triad of body-mounted gyros. Using (56), (59), (60), and the definition of the state matrix X_k yields, after some algebraic manipulations, the following process equation

$$X_{k+1} = \Theta_k X_k + W_k \tag{61}$$

where Θ_k is obtained by substituting the measured angular velocity $\boldsymbol{\omega}_k$ for $\boldsymbol{\omega}^o_k$ in (60), and W_k is the process noise. Equation (61) is the state process equation of the matrix plant under consideration. Equations (61) and (58) constitute the matrix state-space model to which one can apply the general MKF.

The tested scenario is for a spinning spacecraft that undergoes nutation. The spin velocity is 0.464 r/min, the nutation rate is 1 r/hr, and the nutation angle is 22.5°. As a special case, the gyro noise is assumed to be a zero-mean white sequence with covariance matrix $Q^\epsilon_k = (0.01\ \text{deg/hr})^2 I_3$. An analytic expression for the 6×6 covariance matrix of W_k, Q_k, is computed as follows,

$$Q_k = [(\hat{X}_{k/k})^T \otimes I_3]\mathcal{L}Q^\epsilon_k \mathcal{L}^T[(\hat{X}_{k/k})^T \otimes I_3]^T \Delta t^2 \tag{62}$$

$$\mathcal{L}^T = [[\mathbf{e}_1 \times] \quad [\mathbf{e}_2 \times] \quad [\mathbf{e}_3 \times]] \tag{63}$$

where $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the standard basis in $\mathbb{R}^3$, and $\hat{X}_{k/k}$ denotes the a posteriori estimate of X_k at t_k. Equations (62) and (63) provide an initial value for Q_k, which is further refined by filter tuning. The same two directions are simultaneously observed at each sampling time. The reference-components vectors of the two observed directions are chosen as $[\mathbf{r}_1 \ \ \mathbf{r}_2] = [\mathbf{e}_1 \ \ \mathbf{e}_2]$. The vector observations noises are assumed to be zero-mean, white sequences with covariance matrices $R_{k,1} = 4 \cdot 10^{-4} I_3[\text{rad}^2]$ and $R_{k,2} = 4 \cdot 10^{-5} I_3[\text{rad}^2]$. They are also assumed to be uncorrelated, so that R_k, the covariance matrix of V_k, is a 6×6 block-diagonal matrix expressed as $R_k = \text{diag}(R_{k,1}, R_{k,2})$. The initial value of the state X_0 is a Gaussian matrix random variable with mean $\bar{X}_0 = [\mathbf{e}_1 \ \ \mathbf{e}_2]$ and covariance matrix $\Pi_0 = 0.2 I_6$, where I_6 denotes the 6×6 identity matrix. The vector

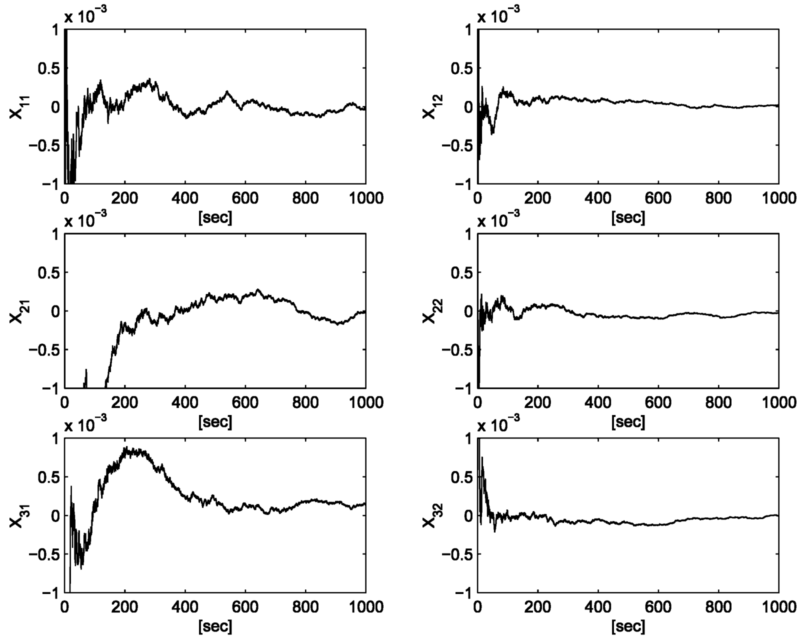


Fig. 1. Sample run. Time histories of estimation errors.

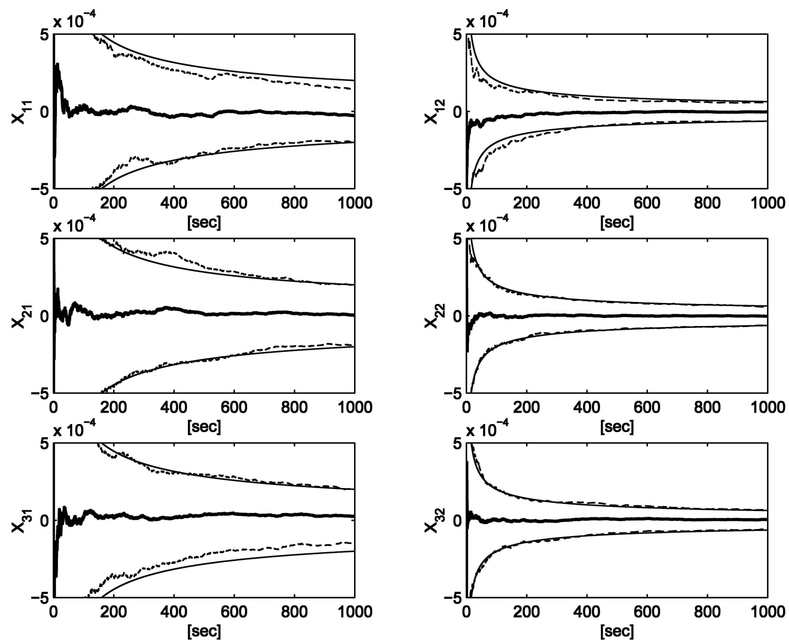


Fig. 2. Monte-Carlo simulation results. Full thick line: Monte-Carlo mean, broken thin lines: Monte-Carlo $\pm 1\ \sigma$, full thin lines: filter $\pm 1\ \sigma$.

measurement noises, the gyro noise, and the initial state are assumed to be uncorrelated with one another. The MKF algorithm is initialized with $\hat{X}_0 = O_6$, and $P_{0/0} = 5I_6$ (O_6 and I_6 are the zero matrix and the identity matrix, respectively, in $\mathbb{R}^6$).

Figs. 1 and 2 summarize the simulation results. Fig. 1 presents a sample run of the time histories of the estimation errors. The left-hand column of the plots corresponds to the first observed direction, while the plots in the right-hand column correspond to the second direction. It can be seen that the estimation errors are convergent. Because of the higher accuracy associated with the second vector observation, the estimator performance is better in the right-hand column, e.g., the transient phases are shorter and the steady-state values are smaller. These conclusions are confirmed when looking at the 100 runs Monte-Carlo simulation results (Fig. 2). The 1 σ-envelopes are clearly narrower in the right-hand column than in the left-hand column. Moreover, Fig. 2 shows that there is consistency between the actual variance and the filter-computed variance. Note that this happens in spite of the nonlinearities introduced in the process equation and in the expression for Q_k.

VII. CONCLUSION

It has been shown that stochastic systems described by linear matrix difference equations and observed through linear matrix measurement equations can be estimated by a matrix version of the ordinary KF. The proposed estimation algorithm uses a compact matrix notation to produce the matrix estimate and the estimation error covariance matrix in terms of the original plant coefficients. The MKF is a natural and straightforward extension of the ordinary KF, and includes, as special cases, other matrix filters previously introduced. Comparative examples motivated by engineering problems illustrated the notational advantage of the matrix filter over the vectorized filter. If the parameters μ and ν are small, the matrix filter presents a computational advantage, too. As a numerical example, an MKF was designed for solving the first stage of an AD problem.

REFERENCES

[1] Kalman, R. E.
A new approach to linear filtering and prediction problems.
Transactions on ASME, Journal of Basic Engineering, ser. D, **82** (Mar. 1960), 35–45.

[2] Kalman, R. E., and Bucy, R. S.
New results in linear filtering and prediction theory.
Transactions on ASME, Journal of Basic Engineering, ser. D, **83** (Mar. 1961), 95–108.

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

[4] 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.

[5] 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.

[6] Athans, M.
The matrix maximum principle.
Information and Control, **11**, 5–6 (1967), 592–606.

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

[8] Golub, G. H., and Van Loan, C. F.
An analysis of the total least-squares problem.
SIAM Journal of Numerical Analysis, **17** (1980), 883–893.

[9] Van Huffel, S., and Vandewalle, J.
The total least squares problem: Computational aspects and analysis.
Society for Industrial and Applied Mathematics, Philadelphia, PA, 1991.

[10] Mendel, J. M.
Lessons in Estimation Theory for Signal Processing, Communications, and Control.
Englewood Cliffs, NJ: Prentice-Hall, 1995.

[11] Neudecker, H.
Some theorems on matrix differentiation with special references to kronecker products.
Journal of American Statistical Associations, **64** (1969), 953–963.

[12] Tracy, D. S., and Dwyer, P. S.
Multivariate maxima and minima with matrix derivatives.
Journal of American Statistical Associations, **64** (1969), 1576–1594.

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

[14] MacRae, E. C.
Matrix derivatives with an application to an adaptive linear decision problem.
The Annals of Statistics, **2**, 2 (1974), 337–346.

[15] Goodwin, G. C., and Sin, K. S.
Adaptive Filtering Prediction and Control.
Englewood Cliffs, NJ: Prentice-Hall, 1984.

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

[17] Choukroun, D.
Novel methods for attitude estimation from vector observations.
Ph.D. dissertation, Technion, Israel Institute of Technology, 2003.

[18] Weiss, H., and Hexner, G.
Simple structure for a high-performance three-dimensional tracking filter.
Journal of Guidance, Control, and Dynamics, **27**, 3 (May–June 2004), 491–493.

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

[20] 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.

[21] 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.

[22] Choukroun, D., Weiss, H., Bar-Itzhack, I. Y., and Oshman, Y.
Direction cosine matrix estimation from vector observations using a matrix Kalman filter.
In *Proceedings of the AIAA Guidance, Navigation and Control Conference*, Austin, TX, Aug. 2003.

Daniel Choukroun (M'00) received the B.Sc. (summa cum laude), M.Sc., and Ph.D. degrees in aerospace engineering from the Technion–Israel Institute of Technology, Haifa, Israel, in 1997, 2000, and 2003, respectively. He also holds the title Engineer Under Instruction from the Ecole Nationale de l'Aviation Civile, Toulouse, France, received 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, where he received the Miriam and Aaron Gutwirth Special Excellency Award for achievement in research. Since 2003, he has been a postdoctoral fellow and a lecturer at the University of California, Los Angeles, in the Department of Mechanical and Aerospace Engineering. His research interests are in optimal estimation and control theory with applications to aerospace systems.

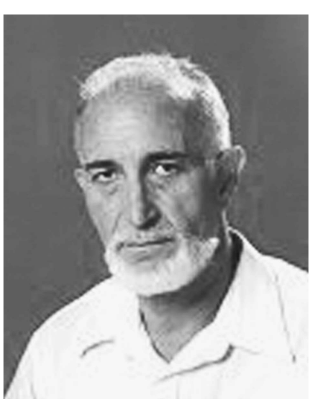
Haim Weiss (M'02) received his B.Sc. and M.Sc. degrees in electrical engineering, in 1966 and 1971, respectively, both from Technion–Israel Institute of Technology. In 1979 he received the Ph.D. degree from the University of Newcastle, Australia, also in electrical engineering.

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 a position of Adjunct Senior Teaching Fellow.

His research interests include estimation, guidance, spacecraft attitude determination and control, and inertial navigation.

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

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



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

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

From 1975 to 1981 he was with the Israeli Air Force, where he worked in the areas of structural dynamics and flutter analysis and flight testing. In 1987 he was a research associate at the Department of Mechanical and Aerospace Engineering of the State University of New York at Buffalo, where he was, in 1988, a visiting professor. Since 1989 he has been with the Department of Aerospace Engineering at the Technion–Israel Institute of Technology, where he is a professor. During the 1996/1997 and 1997/1998 academic years he spent a sabbatical with the Guidance, Navigation and Control Center of NASA's Goddard Space Flight Center, where he worked in research and development of spacecraft attitude estimation algorithms. He has consulted to RADA Electronic Industries Ltd., RAFAEL Ltd., and the Israeli Ministry of Defense. His research has been supported by the Israeli Aircraft Industries, the Israeli Ministry of Defense, the U.S. Air Force Office of Scientific Research (AFOSR), RAFAEL Ltd., and the Israeli Science Foundation (ISF), as well as by various Technion research grants.

Dr. Oshman is the President of the Israeli Association for Automatic Control, a national member organization of the International Federation of Automatic Control (IFAC). He is a member of the AIAA Guidance, Navigation and Control Technical Committee. He headed the Philadelphia Flight Control Laboratory at the Technion (1993–1996). His research interests are in advanced optimal estimation, information fusion and control methods and their application in Aerospace guidance, navigation and control systems. Of particular interest are interdisciplinary aerospace systems, including structural estimation and control, and health monitoring/fault detection and isolation (FDI) systems. He published over 120 journal and conference papers and book chapters 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 has been awarded the Best Paper Award of the 2004 AIAA Guidance, Navigation and Control Conference. He has received the Technion's Raymond and Miriam Klein Research Prize for his research on enhanced air-to-air missile tracking using target orientation observations (2002), and the Technion's Meir Hanin Research Prize for his work on spacecraft angular velocity estimation (2004). The latter work has been put to use in the Israeli AMOS-2 communication satellite. He has been on the program committees of over a dozen international conferences.

Dr. Oshman is an editor for Guidance and Control Systems for the *IEEE Transactions on Aerospace and Electronic Systems*, and a member of the editorial board of the *AIAA Journal of Guidance, Control and Dynamics*. He is an Associate Fellow of the AIAA.