Gain-Free Square Root Information Filtering Using the Spectral Decomposition

Yaakov Oshman*
State University of New York at Buffalo, Buffalo, New York

A new square root state estimation algorithm is introduced, that operates in the information mode in both the time and the measurement update stages. The algorithm, called the V-Lambda filter, is based on the spectral decomposition of the covariance matrix into a $V\Lambda V^T$ form, where V is the matrix whose columns are the eigenvectors of the covariance matrix, and Λ is the diagonal matrix of its eigenvalues. The algorithm updates a normalized state estimate along with the information matrix square root factors, thus doing away with the gain computation. Both stages of the filter constitute equation-free algorithms and thus ideally suit parallel processing implementations. Singular value decomposition is used as a sole computational tool in both the eigenvectors/eigenvalues and the normalized state estimate updates, rendering a complete estimation scheme with exceptional numerical stability and precision. The distinct square root nature of the new algorithm is demonstrated numerically via a typical example, which compares the performance of the V-Lambda filter to that of the corresponding conventional Kalman algorithm. Belonging to the class of square root estimation algorithms, the new filter has all the virtues of a true square root routine. However, the new formulation also provides its user with invaluable insight into the heart of the estimation process, which is a unique characteristic of the V-Lambda filters.

I. Introduction

IT is now widely recognized that the filtering algorithms developed by Kalman[1] and Kalman and Bucy[2] may suffer from numerical instability when implemented in practice. Soon after the introduction of these algorithms, it was shown that, especially when implemented on relatively short-word-length computers, they may lead to the computation of negative definite covariance matrices.[3,4] Except for being a theoretical impossibility, they also may cause filter divergence. It is important to note here that this numerical phenomenon may occur in an unanticipated fashion, and that the filter may have started diverging even before the negative covariance eigenvalues have been observed.[5]

Many ad hoc solutions have been offered in the literature to this problem.[6] However, recognizing the fact that the problem is caused by the use of a numerically unstable algorithm, it is now commonly agreed in the estimation community that the best solution is to use the so-called square root estimation algorithms, which are inherently numerically stable. In fact, today these algorithms are considered to be the only reliable means of applying linear estimation theory to practical estimation problems. By the nature of the square root approach to linear estimation, this class of algorithms contains many different methods. The common feature to all of these algorithms is that they all use some decomposition (termed: the square root decomposition) of the estimation error covariance matrix into its square root factors, thereby replacing the covariance matrix by its factors in each of the computation stages so that the covariance itself is never explicitly computed. When needed, the covariance can be easily reconstructed via its square root factors. Obtained this way, the covariance is assured to be symmetric positive semidefinite. Moreover, it has been shown that in some ill-conditioned cases these algorithms can deliver up to twice the accuracy of the conventional Kalman algorithm (their accuracy when implemented on a certain word length computer is comparable to the accuracy of the regular algorithm implemented with double that word length).[5] Among these methods are those based on the QR factorization[7] and Bierman's U-D method[8,9] that uses a UDU^T decomposition, where U is a unit upper triangular matrix and D is diagonal.

Recently, two new square root filtering algorithms were introduced.[10] These algorithms, called V-Lambda filters, use the spectral decomposition of the error covariance matrix into a $P \equiv V\Lambda V^T$ form, where V is the matrix of the eigenvectors of P, and Λ is the diagonal matrix of its eigenvalues. The measurement update scheme, common to both algorithms, operates in the information mode and furnishes the a posteriori V and $\Lambda^{-1/2}$ factors via a singular value decomposition. Two time update algorithms were proposed (a continuous one and a discrete one), which both operate in covariance mode, rendering the resulting filters hybrid. As opposed to other square root routines, in which the choice of the type of square root decomposition to be used is based solely on computational efficiency considerations, the V-Lambda algorithms provide their user with an invaluable insight into the estimation process. Being based on the spectral decomposition of the covariance they enable the user to monitor closely the eigenfactors (eigenvalues/eigenvectors), which are continuously available, so that singularities may be revealed as they develop. Moreover, using these algorithms makes it very easy to identify those state subsets that become nearly dependent (a problem that cannot be easily resolved using other square root methods) (Ref. 5, p. 100; Ref. 7, p. 72). The reader also is referred to Ref. 11 for an enlightening explanation of the role of the covariance eigenfactors in the analysis of the observability of certain linear combinations of state variables.

In comparison with other information square root algorithms that update a normalized estimate of the state and thus avoid the necessity of gain computation,[5] the V-Lambda algorithms presented in Ref. 10 update the state estimate directly; this, of course, makes it necessary to compute the gain matrix, and, while two alternative gain computation algorithms were outlined, it would still be desirable (from the computational point of view) to develop a V-Lambda information algorithm which is "gain-free," in a similar fashion to other square root information routines.

In the present paper, we introduce a new formulation of the discrete time V-Lambda measurement update algorithm. The

Received Nov. 16, 1987; revision received April 21, 1988.

*Visiting Assistant Professor, Department of Mechanical and Aerospace Engineering; currently, Lecturer, Faculty of Aeronautical Engineering, Technion—Israel Inst. of Technology, Haifa, Israel. Member AIAA.

new formulation uses the singular value decomposition technique to provide the updated normalized estimate of the state along with the updated square root factors of the information matrix. This formulation does away with the gain computation and, at the same time, renders the state estimate update algorithm numerically robust (being computed via the singular value decomposition). A new time update V-Lambda algorithm is introduced too, which operates in the information mode as well. The new time update algorithm is also based on the singular value decomposition technique and is combined with the measurement update algorithm to form a complete V-Lambda information filter.

Both of the new algorithms are derived using a dynamic programming approach, as opposed to the direct approach based on the Kalman filter equations, used in Ref. 10. Being based solely on the numerically robust singular value decomposition, the new V-Lambda filter offers its user excellent numerical reliability and accuracy qualities, together with the provision of the covariance eigenfactors, which, as observed before, adds physical insight to square root filtering. The recent vast development in the area of parallel computation of the singular value decomposition using multiprocessor arrays[12–14] also makes the new filtering algorithm computationally attractive.

In the next section we describe the method of dynamic programming as applied to optimal linear filtering. The results presented in this section are used in Secs. III and IV, where both stages of the V-Lambda filter, namely, the measurement update stage and the time propagation stage, are derived. These algorithms are tailored in Sec. V to form a complete square root state estimator. In Sec. VI we present a numerical example, which serves to demonstrate the superior numerical characteristics of the V-Lambda algorithm via a comparison with the conventional Kalman filtering algorithm. Concluding remarks are offered in the final section.

II. Dynamic Programming Approach to Linear Estimation

In this section we present a dynamic programming approach to the problem of linear optimal estimation. This approach, and the development that follows in this section, are based on the dynamic programming solution given in Ref. 15 to the general (nonlinear) filtering problem. The purpose of this section is twofold: 1) to acquaint the reader briefly with the method, and 2) to prepare the theoretical basis for developing the square root filtering algorithm in later sections.

In the ensuing, we will use the following notation to describe the discrete-time stochastic process whose state is to be estimated:

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

where $x_k \in R^n$ is the state vector, $\{w_k\} \in R^p$ a Gaussian white sequence with zero mean and positive definite covariance Q_k, and $F_k \in R^{n,n}$ the dynamics (transition) matrix.

It is assumed that the initial state vector x_0 is random and has a Gaussian distribution with mean μ and positive definite covariance P_0.

The measurement equation is

$$y_k = H_k x_k + v_k \tag{2}$$

where $y_k \in R^m$ is the measurement vector, $\{v_k\} \in R^m$ the Gaussian white measurement noise sequence with zero mean and positive definite covariance R_k, and $H_k \in R^{m,n}$ the measurement matrix.

It is further assumed that the measurement noise, the process noise, and the initial condition random vector are not correlated.

Before proceeding with the development, it should be noted that the Gaussian distribution assumption is made in order to facilitate the use of the dynamic programming approach. Indeed, as will be proved in the sequel, the filtering algorithm that will be derived is the optimal *linear* filter for non-Gaussian systems as well.

Our goal here is to find the optimal estimate $\hat{X}$ of the sequence $X^n = \{x_0, x_1, \ldots, x_n\}$, given the measurements $Y^n = \{y_0, y_1, \ldots, y_n\}$. This estimate is defined as that $\hat{X}$ which brings to minimum the following cost function:

$$\underset{X|Y}{E} \; \{\|X^n - \hat{X}\|_W^2\} \tag{3}$$

where E denotes the expectation operator (conditioned here on the measurements), $\|\cdot\|$ is the Euclidean vector norm, and W is an arbitrary positive definite weighting matrix. As is well known,[16] the optimal estimator (in the sense that was defined earlier), is the conditional expectation

$$\hat{X} = E\{X^n | Y^n\}$$

In the Gaussian case,[6] this conditional expectation coincides with the mode of the conditional probability density function (pdf) $f_{X|Y}(x|y)$. Since we are dealing with a linear dynamical system whose inputs (and initial condition) are Gaussian distributed, we will use the following procedure to find the optimal estimator:

1) We will write an explicit expression for the pdf in our case.

2) Using the dynamic programming approach, we will find the mode of this pdf (at which the function attains maximum), which will be the optimal estimator sought for.

Proceeding with the first stage, we use Bayes rule to write

$$f_{X|Y}(x_0,\ldots,x_n|y_0,\ldots,y_n) = \frac{f_{Y|X}(y_0,\ldots,y_n|x_0,\ldots,x_n) f_X(x_0,\ldots,x_n)}{f_Y(y_0,\ldots,y_n)}$$

Using the measurement equation (2), the Markovian nature of the process (1), and the fact that the measurement noise is a white sequence, we can express the conditional pdf as follows:

$$f_{X|Y}(x_0,\ldots,x_n|y_0,\ldots,y_n) = \frac{\prod_{k=0}^{n} f_{v_k}(y_k - H_k x_k) f_0(x_0) \prod_{k=1}^{n} f_k(x_k|x_{k-1})}{f_Y(y_0,\ldots,y_n)} \tag{4}$$

where f_{v_k} is the (Gaussian) pdf of v_k, $f_0(x_0)$ the (Gaussian) pdf of x_0, $f_k(x_k|x_{k-1})$ the conditional pdf of x_k conditioned on x_{k-1}, and $f_Y(y_0,\ldots,y_n)$ the joint pdf of the measurements Y^n.

From the system equation (1) it follows that $f_k(x_k|x_{k-1})$ is Gaussian, with mean $F_{k-1}x_{k-1}$ and covariance $G_{k-1}Q_{k-1}G_{k-1}^T$. Without loss of generality, we can assume this covariance to be nonsingular (since the system's state equations can always be rearranged so that this condition is satisfied). Therefore, we can rewrite Eq. (4) in the following explicit form:

$$\begin{aligned} f_{X|Y}(x_0,\ldots,x_n|y_0,\ldots,y_n) &= C(F_k,G_k,Q_k,R_k) \\ &\times \exp\biggl(-\frac{1}{2}\sum_{k=0}^{n}\|y_k - H_k x_k\|_{R_k^{-1}}^2 - \frac{1}{2}\|x_0 - \mu\|_{P_0^{-1}}^2 \\ &\quad -\frac{1}{2}\sum_{k=0}^{n-1}\|x_{k+1} - F_k x_k\|_{[GQG^T]_k^{-1}}^2\biggr) \end{aligned} \tag{5}$$

where $C(F_k,G_k,Q_k,R_k)$ is a constant that is determined by the system's parameters.

Having written the conditional pdf in an explicit form, we now turn to the second stage of the development, namely, that of finding the mode of that function. Since, as can be seen from Eq. (5), the pdf is an exponential function, it follows that an equivalent problem to that of finding its mode is the following one.

Minimize the cost function:

$$I_n = \frac{1}{2}\|x_0 - \mu\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{k=0}^{n}\|y_k - H_k x_k\|^2_{R_k^{-1}} + \frac{1}{2}\sum_{k=0}^{n-1}\|x_{k+1} - F_k x_k\|^2_{[GQG^T]_k^{-1}} \tag{6}$$

with respect to the sequence $\{x_0,...,x_n\}$.

Note here that, in light of the system equation (1), the last term in Eq. (6) also can be expressed as

$$\frac{1}{2}\sum_{k=0}^{n-1}\|G_k w_k\|^2_{[GQG^T]_k^{-1}}$$

Therefore, the problem of minimizing I_n in Eq. (6) is equivalent to that of minimizing the following function:

$$J_n = \frac{1}{2}\|x_0 - \mu\|^2_{P_0^{-1}} + \frac{1}{2}\sum_{k=0}^{n}\|y_k - H_k x_k\|^2_{R_k^{-1}} + \frac{1}{2}\sum_{k=0}^{n-1}\|w_k\|_{Q_k^{-1}} \tag{7a}$$

with respect to (w.r.t.) the sequences $\{x_0,...,x_n\}$ and $\{w_0,...,w_{n-1}\}$, subject to the equality constraints:

$$x_{k+1} = F_k x_k + G_k w_k \qquad k = 0,1,...,n-1. \tag{7b}$$

This is a $(n+1)$-steps constrained minimization problem whose solution yields the required sequence of state estimates $\{\hat{x}_0,...,\hat{x}_n\}$, as well as the sequence of smoothed estimates of the process noise: $\{\hat{w}_0,...,\hat{w}_{n-1}\}$. The rest of this section is devoted to the solution of this minimization problem, using the dynamic programming method.

In principle, the minimization problem (7) can be solved using some method of parameter optimization. This will yield the required results (as described above), however, this method of solution corresponds to a batch processing of the measurements. In general, a recursive solution (which will furnish, at each time, t_n, only the last optimal estimate, $\hat{x}_n$, based on the measurements history Y^n) is more desirable than the corresponding batch algorithm because of obvious implementation considerations. Therefore, we will reformulate the minimization problem in a dynamic programming setting, which will then allow us to outline a recursive solution (filter).

In the sequel, we shall assume that the dynamics matrix is invertible (this is a common assumption in information filtering algorithms; the dynamics matrix is guaranteed to be invertible when the discrete mathematical model of the system is derived via a discretization of an originally continuous model). We may, therefore, rewrite the system equation in the following way:

$$x_k = F_k^{-1}[x_{k+1} - G_k w_k] \tag{8}$$

Now define the following sequence of functions:

$$S_0(x_0) := \|x_0 - \mu\|^2_{P_0^{-1}} + \|y_0 - H_0 x_0\|^2_{R_0^{-1}} \tag{9a}$$

and for $n = 1,2,...$:

$$S_n(x_n) := \min_{w_0,...,w_n}\left(\|x_0 - \mu\|^2_{P_0^{-1}} + \sum_{k=0}^{n}\|y_k - H_k x_k\|^2_{R_k^{-1}} + \sum_{k=0}^{n-1}\|w_k\|^2_{Q_k^{-1}}\right) \tag{9b}$$

subject to the equality constraints, Eq. (7b).

A few remarks are in order here. First, note that because of the constraints of Eq. (7b) $S_n(x_n)$ is indeed a function of x_n only [and not of the entire sequence X^n that appears in the second term of Eq. (9b)]. Using these constraints we can express any x_i (for $i \leq n$) in terms of x_n only. Note also that S_n is not a function of the process noise sequence $\{w_k\}_{k=0}^{n-1}$, since it is defined as the result of a minimization w.r.t. this sequence. Finally, comparing the definition of $S_n(x_n)$ in Eqs. (9) and the minimization problem defined in Eqs. (7), we observe that a further minimization of $S_n(x_n)$ w.r.t. x_n will result in the x_n that minimizes J_n, which is the required optimal filtered estimate $\hat{x}_{n/n}$.

Since we are interested in a recursive algorithm, let us rewrite Eq. (9b) in the following way:

$$\begin{aligned} S_n(x_n) &= \min_{w_{n-1}}\left(\min_{w_0,...,w_{n-2}}\left\{\|x_0 - \mu\|^2_{P_0^{-1}} + \sum_{k=0}^{n-1}\|y_k - H_k x_k\|^2_{R_k^{-1}} + \sum_{k=0}^{n-2}\|w_k\|^2_{Q_k^{-1}}\right\} + \|y_n - H_n x_n\|^2_{R_n^{-1}} + \|w_{n-1}\|^2_{Q_{n-1}^{-1}}\right) \\ &= \min_{w_{n-1}}\left\{S_{n-1}(F_{n-1}^{-1}[x_n - G_{n-1}w_{n-1}]) + \|y_n - H_n x_n\|^2_{R_n^{-1}} + \|w_{n-1}\|^2_{Q_{n-1}^{-1}}\right\} \end{aligned} \tag{10}$$

which is a recursive equation in the functions S_i. It can be shown[17] that a recursive solution of the estimation problem can be obtained by assuming solutions of the form

$$S_n(x_n) = \|x_n - \hat{x}_{n/n}\|^2_{P_{n/n}^{-1}} + r_n^2 \tag{11}$$

[which means that $S_n(x_n)$ attains its minimum at the a posteriori estimate $\hat{x}_{n/n}$] and proceeding with a step-by-step minimization (r_n^2 is the minimization residual). The following remarks are made:

1) Minimization of $S_0(x_0)$ w.r.t. x_0 provides the updated (a posteriori) estimate $\hat{x}_{0/0}$. On examination of Eq. (9a), it is seen that this estimate is obtained by augmenting the a priori information on x_0 with the additional measurement information in such a way that each information is weighted according to its certainty level (represented here by the covariance inverse).

2) The outlined procedure furnishes, at each time step, the a posteriori estimate. Although this may be satisfactory in many cases, it is easy to see that, through a reformulation of the problem, the minimization process can be reorganized into two consecutive stages such that, in addition to the a posteriori estimate, the a priori estimate also will be obtained, thus conforming to the more common two-stage representation of the Kalman filter.

To meet this end, Eq. (10) will be rewritten as

$$S_n(x_n) = \min_{w_{n-1}}\left(S_{n-1}(F_{n-1}^{-1}[x_n - G_{n-1}w_{n-1}]) + \|w_{n-1}\|^2_{Q_{n-1}^{-1}}\right) + \|y_n - H_n x_n\|^2_{R_n^{-1}}$$

which also can be written as

$$S_n(x_n) = T_n(x_n) + \|y_n - H_n x_n\|^2_{R_n^{-1}} \tag{12}$$

where $T_n(x_n)$ is defined as:

$$T_n(x_n) := \min_{w_{n-1}}\left(S_{n-1}(F_{n-1}^{-1}[x_n - G_{n-1}w_{n-1}]) + \|w_{n-1}\|^2_{Q_{n-1}^{-1}}\right) \tag{13}$$

Examining Eq. (12), we see that S_n is the sum of two terms, the second of which expresses the information contributed solely by the last acquired measurement (y_n). Remembering that x_n for which S_n attains its minimum is the a posteriori estimate $\hat{x}_{n/n}$ (which is based on the whole measurement history Y^n), we conclude that the value of x_n that minimizes

$T_n(x_n)$ is the a priori (time-propagated) estimate, which is based on the measurements $Y^{n-1} = \{y_0,...,y_{n-1}\}$. It can be shown[15] that there also exists a recursive equation for the functions T_n, whose solution, augmented by a minimization at each step, furnishes the a priori estimate at each time step. This solution can be written as:

$$T_n(x_n) = \|x_n - \hat{x}_{n/n-1}\|^2_{P^{-1}_{n/n-1}} + \rho_n^2 \tag{14}$$

which is analogous to the form of the solution for S_n [Eq. (11)] (ρ_n^2 is the minimization residual). Using Eq. (14) in Eq. (12), we obtain

$$S_n(x_n) = \|x_n - \hat{x}_{n/n-1}\|^2_{P^{-1}_{n/n-1}} + \|y_n - H_n x_n\|^2_{R_n^{-1}} + \rho_n^2 \tag{15}$$

which (noting that $\hat{x}_{n/n}$ is the result of the minimization of the last expression) means that the a priori estimate $\hat{x}_{n/n-1}$ is treated by the estimation process as a "measurement," in addition to y_n (this is true for any information filter[5]). In a similar fashion, if in Eq. (13) $S_{n-1}(x_{n-1})$ is expressed in terms of $\hat{x}_{n-1/n-1}$ using Eq. (11), the following expression is obtained for $T_n(x_n)$:

$$T_n(x_n) = \min_{w_{n-1}} (\|F^{-1}_{n-1}[x_n - G_{n-1}w_{n-1}] - \hat{x}_{n-1/n-1}\|^2_{P^{-1}_{n-1/n-1}} + \|w_{n-1}\|^2_{Q^{-1}_{n-1}}) \tag{16}$$

Summarizing the presentation of the dynamic programming approach to optimal linear estimation, the two-stage estimation process is composed of the following recursive algorithm:

1) *Measurement update*: having obtained the a priori (time propagated) estimate, the a posteriori filtered estimate $\hat{x}_{n/n}$ is obtained by minimizing $S_n(x_n)$ [Eq. (15)] w.r.t. x_n.

2) *Time update*: having obtained the a posteriori estimate, the a priori estimate $\hat{x}_{n/n-1}$ is obtained by minimizing $T_n(x_n)$ [Eq. (16)] w.r.t. x_n.

The method presented in this section will be used next to develop both stages of the V-Lambda information filter. We start with the measurement update algorithm.

III. Gain-free V-Lambda Measurement Update Algorithm

The measurement update problem is as follows: Given the square root spectral factors $V_{k/k-1}$ and $\Lambda^{-1/2}_{k/k-1}$ of the a priori information matrix $P^{-1}_{k/k-1}$, where $P_{k/k-1}$ is the a priori estimation error covariance, $V_{k/k-1}$ is the eigenvectors matrix, $\Lambda_{k/k-1}$ is the diagonal eigenvalues matrix and $P_{k/k-1} = V_{k/k-1}\Lambda_{k/k-1}V^T_{k/k-1}$, and given the a priori normalized state estimate $\hat{d}_{k/k-1}$ [defined below in Eq. (18)], compute the a posteriori square root factors $V_{k/k}$ and $\Lambda^{-1/2}_{k/k}$, and the updated normalized estimate $\hat{d}_{k/k}$, defined as

$$\hat{d}_{k/k} := \Lambda^{-1/2}_{k/k} V^T_{k/k}\hat{x}_{k/k} \tag{17}$$

The solution to the measurement update problem is summarized in the next theorem.

Theorem 3.1: V-Lambda Measurement Update. Given: the time propagated factors $V_{k/k-1}$ and $\Lambda^{-1/2}_{k/k-1}$, the nonsingular measurement noise covariance R_k, and the a priori normalized estimate $\hat{d}_{k/k-1}$, where

$$\hat{d}_{k/k-1} := \Lambda^{-1/2}_{k/k-1} V^T_{k/k-1}\hat{x}_{k/k-1} \tag{18}$$

define an augmented matrix $A_k \in R^{n+m,n}$ as

$$A_k := \begin{pmatrix} \Lambda^{-1/2}_{k/k-1} V^T_{k/k-1} \\ R_k^{-1/2} H_k \end{pmatrix} \tag{19}$$

and perform a singular value decomposition of it to obtain

$$A_k = Y_k \begin{pmatrix} \Sigma_k \\ 0 \end{pmatrix} U_k^T \tag{20}$$

Then, the measurement updated spectral factors are related to the singular value decomposition factors of A_k as follows

$$V_{k/k} = U_k \tag{21a}$$

$$\Lambda^{-1/2}_{k/k} = \Sigma_k \tag{21b}$$

Moreover, define b_k as

$$b_k := \begin{pmatrix} \hat{d}_{k/k-1} \\ R_k^{-1/2} y_k \end{pmatrix} \tag{22}$$

and premultiply it by Y_k^T; then, partitioning the resulting vector in accordance with the partition of b_k, the updated normalized estimate is found as follows:

$$Y_k^T b_k = \begin{pmatrix} \hat{d}_{k/k} \\ f_2 \end{pmatrix} \tag{23}$$

where the meaning of the m-vector f_2 will become clear in the ensuing [in Eq. (31)]. (Note: $Y_k \in R^{n+m,n+m}$ and $U_k \in R^{n,n}$ are the orthogonal matrices of the eigenvectors of $A_k A_k^T$ and $A_k^T A_k$, respectively, and $\Sigma_k \in R^{n,n}$ is a diagonal matrix whose nonzero elements are the eigenvalues of $A_k^T A_k$.)

Proof: As shown in Sec. II (Eq. 15), the optimal estimate of x_k, based on the measurements $\{y_0, y_1, ..., y_k\}$ can be obtained by minimizing the cost function

$$S_k(x_k) = \|x_k - \hat{x}_{k/k-1}\|^2_{P^{-1}_{k/k-1}} + \|y_k + H_k x_k\|^2_{R_k^{-1}} \tag{24}$$

with respect to x_k. Using the square root factors of $P^{-1}_{k/k-1}$ and R_k^{-1}, Eq. (24) can be rewritten as

$$S_k(x_k) = \|\Lambda^{-1/2}_{k/k-1} V^T_{k/k-1}(x_k - \hat{x}_{k/k-1})\|^2 + \|R_k^{-1/2}(H_k x_k - y_k)\|^2 = \left\| \begin{pmatrix} \Lambda^{-1/2}_{k/k-1} V^T_{k/k-1} \\ R_k^{-1/2} H_k \end{pmatrix} x_k - \begin{pmatrix} \Lambda^{-1/2}_{k/k-1} V^T_{k/k-1}\hat{x}_{k/k-1} \\ R_k^{-1/2} y_k \end{pmatrix} \right\|^2 \tag{25}$$

Now, using the augmented matrix A_k defined in Eq. (19), the normalized state estimate $\hat{d}_{k/k-1}$ defined in Eq. (18), and the vector b_k defined in Eq. (22), Eq. (25) can be written as

$$S_k(x_k) = \|A_k x_k - b_k\|^2 \tag{26}$$

Minimization of S_k w.r.t. x_k is now a standard least squares (LS) problem. We use the singular value decomposition to solve it as follows.

Perform the singular value decomposition of A_k as in Eq. (20). Then Eq. (26) becomes

$$S_k(x_k) = \|Y_k \begin{pmatrix} \Sigma_k \\ 0 \end{pmatrix} U_k^T x_k - b_k\|^2$$

which, after premultiplication by the orthogonal (and hence norm-preserving) matrix Y_k^T, becomes

$$S_k(x_k) = \|\begin{pmatrix} \Sigma_k U_k^T x_k \\ 0 \end{pmatrix} - Y_k^T b_k\|^2 \tag{27}$$

Partitioning the second term in Eq. (27) as follows

$$Y_k^T b_k = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \qquad f_1 \in R^n \tag{28}$$

we obtain

$$S_k(x_k) = \|\Sigma_k U_k^T x_k - f_1\|^2 + \|f_2\|^2 \tag{29}$$

Now, clearly, the minimum of $S_k(\boldsymbol{x}_k)$ with respect to $\boldsymbol{x}_k$ is reached when

$$\Sigma_k U_k^T \boldsymbol{x}_k = \boldsymbol{f}_1 \tag{30}$$

from which we obtain the updated state estimate

$$\hat{\boldsymbol{x}}_{k/k} = U_k \Sigma_k^{-1} \boldsymbol{f}_1$$

where from Eqs. (22), (23), and (28) it is clear that $\boldsymbol{f}_1$ consists of the first n elements of the column vector

$$Y_k^T \begin{pmatrix} \hat{\boldsymbol{d}}_{k/k-1} \\ R_k^{-1/2} \boldsymbol{y}_k \end{pmatrix}$$

The minimum value of the cost function (the estimation residual) is

$$S_k(\hat{\boldsymbol{x}}_{k/k}) = \|\boldsymbol{f}_2\|^2 \tag{31}$$

We have yet to prove the expressions for the updated square root information factors (i.e., the factors of $P_{k/k}^{-1}$), and to accomplish that we use the following well-known result.[7] For the solution $\boldsymbol{x}_{LS}$ that minimizes the LS criterion $J = \|A\boldsymbol{x} - \boldsymbol{b}\|^2$, the error covariance matrix is given by

$$P := E\{[\boldsymbol{x} - \boldsymbol{x}_{LS}][\boldsymbol{x} - \boldsymbol{x}_{LS}]^T\} = (A^T A)^{-1} \tag{32}$$

Employing Eq. (32) in our case we obtain

$$P_{k/k}^{-1} = U_k \Sigma_k^2 U_k^T \tag{33}$$

but

$$P_{k/k}^{-1} \equiv V_{k/k} \Lambda_{k/k}^{-1} V_{k/k}^T \tag{34}$$

then, because of the uniqueness of the spectral decomposition, necessarily

$$\Lambda_{k/k}^{-1/2} = \Sigma_k \tag{35a}$$

$$V_{k/k} = U_k \tag{35b}$$

Using Eqs. (35) in Eq. (30) and noting the definition Eq. (17), we observe that the updated normalized state estimate is

$$\hat{\boldsymbol{d}}_{k/k} = \boldsymbol{f}_1 \tag{36}$$

which ends the proof.

Discussion

As we have outlined above, the updated normalized estimate is read from the vector $Y_k^T \boldsymbol{b}_k$. This, however, does not mean that the matrix Y_k [whose dimension is $(n+m) \times (n+m)$] is actually formed and stored. According to the standard Golub-Reinsch algorithm[7,18] Y_k is not computed explicitly but rather is applied as it evolves during the consecutive orthogonal transformation process, to the vector $\boldsymbol{b}_k$ (indeed, when one wishes to compute Y_k one has to define the entries to that algorithm in a particular way).

Another remark that pertains to the computation efficiency is the following. When there are many measurements to be processed (i.e., when m is large relative to n), one may apply an orthogonal transformation to the $(n+m) \times (n+m)$ matrix $[A_k\ \boldsymbol{b}_k]$ to obtain an $(n+1) \times (n+1)$ upper triangular matrix $[C_k\ \boldsymbol{h}_k]$, as a preliminary step before the singular value decomposition.[19] Thus, if $[A_k\ \boldsymbol{b}_k]$ and $[C_k\ \boldsymbol{h}_k]$ are related by

$$Q[A_k | \boldsymbol{b}_k] = \begin{array}{c} \begin{array}{cc} n \text{ cols.} & 1 \text{ col.} \end{array} \\ \left(\begin{array}{c|c} C_k & \boldsymbol{h}_k \\ \hline 0 & 0 \end{array}\right) \end{array} \begin{array}{l} n+1 \text{ rows} \\ m-1 \text{ rows} \end{array}$$

where $Q \in R^{n+m,n+m}$ is orthogonal, then it can be verified easily that

$$\|A_k \boldsymbol{x}_k - \boldsymbol{b}_k\|^2 = \|C_k \boldsymbol{x}_k - \boldsymbol{h}_k\|^2 \qquad \text{for all} \quad \boldsymbol{x}_k$$

and

$$A_k^T A_k = C_k^T C_k$$

Therefore, the same results are obtained if singular value decomposition is applied to C_k and $\boldsymbol{h}_k$ instead of A_k and $\boldsymbol{b}_k$. At the same time, the saving in computer storage may be substantial because the transformation can be arranged so that rows or groups of rows of $[A_k | \boldsymbol{b}_k]$ are processed sequentially in forming $[C_k | \boldsymbol{h}_k]$; thus the minimal storage required for this processing would be $(n+1)(n+2)/2$ locations for the upper triangular matrix $[C_k | \boldsymbol{h}_k]$ plus $n+1$ locations for one row of $[A_k | \boldsymbol{b}_k]$, as compared to the $(n+m)(n+1)$ locations needed for all the elements of $[A_k | \boldsymbol{b}_k]$.

In the next section we present a V-Lambda time update algorithm that, when combined with the current measurement update algorithm, forms a complete square root estimator.

IV. V-Lambda Time Update: An Information Algorithm

Given the a posteriori square root information factors $V_{k/k}$ and $\Lambda_{k/k}^{-1/2}$, where $P_{k/k}^{-1} \equiv V_{k/k} \Lambda_{k/k}^{-1} V_{k/k}^T$, the next theorem shows how to propagate these factors in time to get the a priori factors $V_{k+1/k}$ and $\Lambda_{k+1/k}^{-1/2}$.

Theorem 4.1: V-Lambda Time Update. Given: the measurement updated factors $V_{k/k}$ and $\Lambda_{k/k}^{-1/2}$, the nonsingular transition matrix F_k, the input gain matrix $G_k \in R^{n,p}$ and the nonsingular process noise covariance $Q_k \in R^{p,p}$, the time-propagated spectral factors are computed according to the following algorithm.

Define the augmented array $B_k \in R^{p+n,p+n}$:

$$B_k := \begin{pmatrix} Q_k^{-1/2} & 0 \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} G_k & \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} \end{pmatrix} \tag{37}$$

and perform a partial triangularization of it; that is, find an orthogonal transformation τ such that[20]

$$\tau B_k = \begin{pmatrix} M_k & L_k \\ 0 & N_k \end{pmatrix} \tag{38}$$

where $M_k \in R^{p,p}$ is upper triangular. Now, perform a singular value decomposition of N_k in Eq. (38) to obtain

$$N_k = W_k S_k Z_k^T \tag{39}$$

where W_k, Z_k are the orthogonal matrices of the left and right singular vectors of N_k, respectively, and S_k is the diagonal singular values matrix; then, the a priori eigenfactors at t_{k+1} are given by

$$\Lambda_{k+1/k}^{-1/2} = S_k \tag{40a}$$

$$V_{k+1/k} = Z_k \tag{40b}$$

and the time propagation of the state estimate is performed according to

$$\hat{\boldsymbol{x}}_{k+1/k} = F_k \hat{\boldsymbol{x}}_{k/k} \tag{41}$$

(which is in accordance with the regular Kalman filter algorithm).

Proof: Following the result derived in Sec. II, we have to minimize the following cost function in order to find the time propagation algorithm:

$$T_{k+1}(x_{k+1}) = \min_{w_k} \left(\|F_k^{-1}[x_{k+1} - G_k w_k] - \hat{x}_{k/k}\|^2_{P^{-1}_{k/k}} + \|w_k\|^2_{Q_k^{-1}} \right) \tag{42}$$

with respect to x_{k+1}. We shall perform this minimization in the following two stages:

1) Minimization of the right-hand side of Eq. (42) w.r.t. w_k.

2) Minimization of the result of stage (1) w.r.t. x_{k+1}.

To perform the first stage of the minimization, write Eq. (42) in the following (algebraically equivalent) form:

$$T_{k+1}(x_{k+1}) = \min_{w_k} \left\| \begin{pmatrix} Q_k^{-1/2} \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} G_k \end{pmatrix} w_k - \begin{pmatrix} 0 \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}[x_{k+1} - F_k \hat{x}_{k/k}] \end{pmatrix} \right\|^2 \tag{43}$$

where use has been made of the spectral factors of the information matrix that appears in Eq. (42). Now, perform a triangularization of the augmented matrix that premultiplies w_k in Eq. (43), i.e., find an orthogonal transformation τ such that

$$\tau \begin{pmatrix} Q_k^{-1/2} \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} G_k \end{pmatrix} = \begin{pmatrix} M_k \\ 0 \end{pmatrix} \tag{44}$$

where M_k is upper triangular. Since the orthogonal transformation τ is norm-preserving, use of Eq. (44) enables us to rewrite Eq. (43) as

$$T_{k+1}(x_{k+1}) = \min_{w_k} \left\| \begin{pmatrix} M_k w_k \\ 0 \end{pmatrix} - \begin{pmatrix} f_k^{(1)} \\ f_k^{(2)} \end{pmatrix} \right\|^2 = \min_{w_k} \|M_k w_k - f_k^{(1)}\|^2 + \|f_k^{(2)}\|^2 \tag{45}$$

where the vectors $f_k^{(1)}, f_k^{(2)}$ are defined by

$$\begin{pmatrix} f_k^{(1)} \\ f_k^{(2)} \end{pmatrix} := \tau \begin{pmatrix} 0 \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}[x_{k+1} - F_k \hat{x}_{k/k}] \end{pmatrix} \tag{46}$$

From Eq. (45) it is easy to see that the minimum value of $T_{k+1}(x_{k+1})$ w.r.t. w_k occurs at that w_k for which the first term in Eq. (45) is zero, and then

$$T_{k+1}(x_{k+1}) = \|f_k^{(2)}\|^2 \tag{47}$$

In Eq. (47) the dependence of T_{k+1} on x_{k+1} is implicit. For the purpose of the next stage of the minimization, we have to express this dependence in an explicit form. To meet this end, we rewrite Eq. (46) using a partitioned form of the matrix τ:

$$\begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix} \begin{pmatrix} 0 \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}[x_{k+1} - F_k \hat{x}_{k/k}] \end{pmatrix} = \begin{pmatrix} \tau_{12} \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}[x_{k+1} - F_k \hat{x}_{k/k}] \\ \tau_{22} \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}[x_{k+1} - F_k \hat{x}_{k/k}] \end{pmatrix} = \begin{pmatrix} f_k^{(1)} \\ f_k^{(2)} \end{pmatrix} \tag{48}$$

Substituting $f_k^{(2)}$ from Eq. (48) into Eq. (47) and using Eq. (38), we obtain

$$T_{k+1}(x_{k+1}) = \|N_k[x_{k+1} - F_k \hat{x}_{k/k}]\|^2 = \|x_{k+1} - F_k \hat{x}_{k/k}\|^2_{N_k^T N_k} \tag{49}$$

which, employing the singular value decomposition of N_k in Eq. (39), can be rewritten as

$$T_{k+1}(x_{k+1}) = \|x_{k+1} - F_k \hat{x}_{k/k}\|^2_{Z_k S_k^2 Z_k^T} \tag{50}$$

From the last result it is easily seen that the minimizing value of x_{k+1} is

$$\hat{x}_{k+1/k} = F_k \hat{x}_{k/k} \tag{51}$$

Furthermore, using the LS result stated in the preceding section [Eq. (32)], we readily obtain

$$P^{-1}_{k+1/k} = Z_k S_k^2 Z_k^T \tag{52}$$

from which, after a comparison with the spectral decomposition of $P^{-1}_{k+1/k}$, Eqs. (40) result, ending the proof.

V. The V-Lambda Information Filter

Having developed the algorithms for both the measurement update and the time update stages, the square root V-Lambda algorithm is complete. For convenience, it is summarized in Table 1. We note, however, that in the presentation of the dynamic programming method, which was used to develop the filtering algorithm, it was assumed that the various stochastic processes driving the system and the measurement, as well as the random initial condition, are all Gaussian distributed. This implies that the resulting filtering algorithm is optimal under the Gaussian distribution condition and, at the same time, gives rise to the question: What can be said about the filter's optimality when this condition is not met? The rest of this section is concerned with this question. It will be shown that when the system is not Gaussian, the V-Lambda algorithm is still the optimal linear filter (i.e., the optimal filter among the restricted class of linear filters).

The method by which we show this follows. As is well known, the (conventional) Kalman filter is the optimal linear filter, even in cases where the system and/or measurement are driven by non-Gaussian noises.[16] Hence, in order to show that the V-Lambda filter has the same property (although derived under the assumption of Gaussian noise), it suffices to show that the V-Lambda filter is algebraically equivalent to the Kalman filter. We prove, therefore, the following theorem.

Theorem 5.1: The V-Lambda filter (Table 1) is algebraically equivalent to the (conventional) Kalman filter.

Proof: We break the proof into two parts, each one of which will examine one of the two algorithms comprising the corresponding two stages of the V-Lambda filter.

The measurement update stage. Using the measurement update equations (19), (20), and (21) (Theorem 3.1), we can write

$$\begin{pmatrix} \Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T \\ R_k^{-1/2} H_k \end{pmatrix} = Y_k \begin{pmatrix} \Lambda_{k/k}^{-1/2} \\ 0 \end{pmatrix} V_{k/k}^T$$

Premultiplying each side of this equation by its transpose, we obtain for the left-hand side:

$$(V_{k/k-1} \Lambda_{k/k-1}^{-1/2} \quad H_k^T R_k^{-1/2}) \begin{pmatrix} \Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T \\ R_k^{-1/2} H_k \end{pmatrix} = V_{k/k-1} \Lambda_{k/k-1}^{-1} V_{k/k-1}^T + H_k^T R_k^{-1} H_k \equiv P^{-1}_{k/k-1} + H_k^T R_k^{-1} H_k \tag{53}$$

and, similarly, for the right-hand side:

$$V_{k/k} (\Lambda_{k/k}^{-1/2} \quad 0) Y_k^T Y_k \begin{pmatrix} \Lambda_{k/k}^{-1/2} \\ 0 \end{pmatrix} V_{k/k}^T = V_{k/k} \Lambda_{k/k}^{-1} V_{k/k}^T \equiv P^{-1}_{k/k} \tag{54}$$

From a comparison of Eqs. (53) and (54) we obtain

$$P^{-1}_{k/k} = P^{-1}_{k/k-1} + H_k^T R_k^{-1} H_k \tag{55}$$

which is the Kalman filter measurement update formulated in information mode.[16] Thus we have proved that the algorithm

Table 1 V-Lambda filtering algorithm

System model:	$x_{k+1} = F_k x_k + G_k w_k, \quad x \in R^n, w \in R^p, \quad E\{w_k\} = 0, \quad E\{w_j w_k^T\} = Q_k \delta_{jk}$
Measurement model:	$y_k = H_k x_k + v_k, \quad y \in R^m, \quad E\{v_k\} = 0, \quad E\{v_j v_k^T\} = R_k \delta_{jk}$
Initial conditions:	$E\{x_0\} = \mu \quad E\{[x_0 - \mu][x_0 - \mu]^T\} = P_0$
State estimate:	$\hat{x}_{k+1/k} = F_k \hat{x}_{k/k}$

Time update

Spectral factors $B_k := \begin{pmatrix} Q_k^{-1/2} & 0 \\ \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} G_k & \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} \end{pmatrix}$

$\tau B_k = \begin{pmatrix} M_k & L_k \\ 0 & N_k \end{pmatrix}$ τ orthogonal[20] $M_k \in R^{p,p}$ upper triangular

$N_k \xrightarrow[\text{decomposition}]{\text{singular value}} W_k S_k Z_k^T$; Read: $V_{k+1/k} = Z_k \quad \Lambda_{k+1/k}^{-1/2} = S_k$

Measurement update

$A_k := \begin{pmatrix} \Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T \\ R_k^{-1/2} H_k \end{pmatrix} \qquad b_k = \begin{pmatrix} \hat{d}_{k/k-1} \\ R_k^{-1/2} y_k \end{pmatrix}$

$A_k \xrightarrow[\text{decomposition}]{\text{singular value}} Y_k \begin{pmatrix} \Sigma_k \\ 0 \end{pmatrix} U_k^T \qquad \hat{d}_{s/t} := \Lambda_{s/t}^{-1/2} V_{s/t}^T \hat{x}_{s/t}$

Read: $V_{k/k} = U_k, \quad \Lambda_{k/k}^{-1/2} = \Sigma_k, \quad Y_k^T b_k = \begin{pmatrix} \hat{d}_{k/k} \\ f_2 \end{pmatrix} \quad f_2 \in R^m$

for measurement update of the information eigenfactors [Eqs. (19–21)] is algebraically equivalent to the corresponding Kalman filter algorithm. To complete the proof of this part, we still have to prove the equivalence of the state estimate algorithm. This is done next.

Using Eqs. (22) and (23) we can write

$$Y_k^T \begin{pmatrix} \hat{d}_{k/k-1} \\ R_k^{-1/2} y_k \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix}^T \begin{pmatrix} \hat{d}_{k/k-1} \\ R_k^{-1/2} y_k \end{pmatrix} = \begin{pmatrix} \hat{d}_{k/k} \\ f_2 \end{pmatrix}$$

from which we have

$$Y_{11}^T \hat{d}_{k/k-1} + Y_{21}^T R_k^{-1/2} y_k = \hat{d}_{k/k} \tag{56}$$

Also, using Eq. (53) we obtain

$$\begin{pmatrix} \Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T \\ R_k^{-1/2} H_k \end{pmatrix} = \begin{pmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{pmatrix} \begin{pmatrix} \Lambda_{k/k}^{-1/2} \\ 0 \end{pmatrix} V_{k/k}^T = \begin{pmatrix} Y_{11} \Lambda_{k/k}^{-1/2} V_{k/k}^T \\ Y_{21} \Lambda_{k/k}^{-1/2} V_{k/k}^T \end{pmatrix}$$

hence

$$\Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T = Y_{11} \Lambda_{k/k}^{-1/2} V_{k/k}^T \tag{57a}$$

$$R_k^{-1/2} H_k = Y_{21} \Lambda_{k/k}^{-1/2} V_{k/k}^T \tag{57b}$$

From Eqs. (57) we can express Y_{11}, Y_{21} in terms of the information eigenfactors and the measurement geometry and statistics:

$$Y_{11} = \Lambda_{k/k-1}^{-1/2} V_{k/k-1}^T V_{k/k} \Lambda_{k/k}^{1/2} \tag{58a}$$

$$Y_{21} = R_k^{-1/2} H_k V_{k/k} \Lambda_{k/k}^{1/2} \tag{58b}$$

Now substitute Eqs. (58a) and (58b) into Eq. (56):

$$\Lambda_{k/k}^{1/2} V_{k/k}^T V_{k/k-1} \Lambda_{k/k-1}^{-1/2} \hat{d}_{k/k-1} + \Lambda_{k/k}^{1/2} V_{k/k}^T H_k^T R_k^{-1/2} R_k^{-1/2} y_k = \hat{d}_{k/k}$$

from which, after some algebraic manipulation and using the definitions (17) and (18) we obtain

$$P_{k/k-1}^{-1} \hat{x}_{k/k-1} + H_k^T R_k^{-1} y_k = P_{k/k}^{-1} \hat{x}_{k/k}$$

which is the state measurement update equation in the information mode Kalman filter.[16] This completes the proof of the first part of the theorem, namely, that the V-Lambda measurement update is algebraically equivalent to the corresponding Kalman filter algorithm.

The time update stage. Expressing the transformation matrix τ in partitioned form [as it appears in Eq. (48)] and using Eq. (38), we have

$$N_k = \tau_{22} \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1}$$

from which, using the singular value decomposition factors of N_k in Eq. (39) and their relations to the propagated eigenfactors in Eq. (40), we obtain the following expression for τ_{22}:

$$\tau_{22} = W_k \Lambda_{k+1/k}^{-1/2} V_{k+1/k}^T F_k V_{k/k} \Lambda_{k/k}^{1/2} \tag{59}$$

Again, using the partitioned form of τ in Eq. (44), we have

$$\tau_{21} Q_k^{-1/2} + \tau_{22} \Lambda_{k/k}^{-1/2} V_{k/k}^T F_k^{-1} G_k = 0$$

from which, using Eq. (59), we obtain the following expression for τ_{21}:

$$\tau_{21} = -W_k \Lambda_{k+1/k}^{-1/2} V_{k+1/k}^T G_k Q_k^{1/2} \tag{60}$$

Since the transformation τ is orthogonal, its block elements must satisfy

$$\tau_{21} \tau_{21}^T + \tau_{22} \tau_{22}^T = I \tag{61}$$

where I is the identity matrix. Upon substitution of Eqs. (59) and (60) into Eq. (61), and after some manipulation, we obtain

$$G_k Q_k G_k^T + F_k P_{k/k} F_k^T = P_{k+1/k}$$

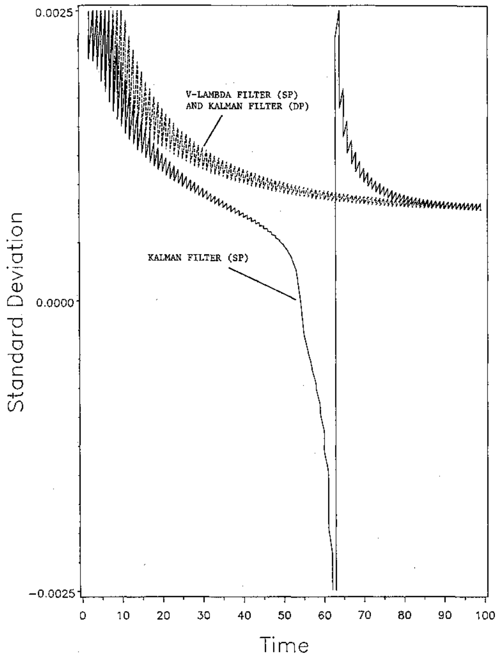

Fig. 1 Standard deviation of the estimation error of x_1.

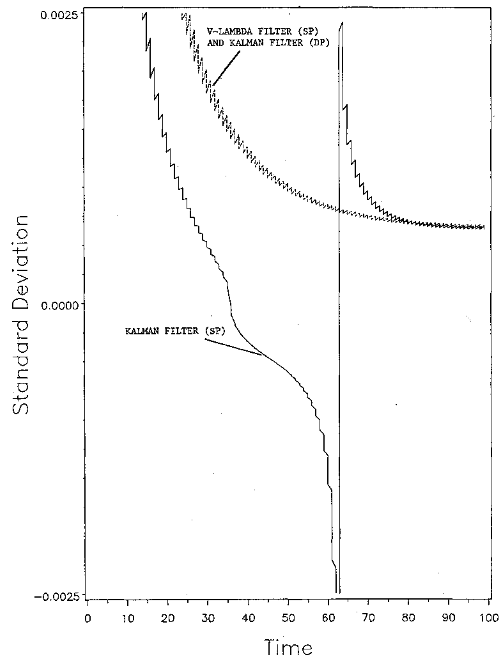

Fig. 2 Standard deviation of the estimation error of x_2.

which is the covariance time update in the Kalman filter. This proves that the time update algorithm of the V-Lambda filter is algebraically equivalent to the conventional Kalman algorithm [note that the state update algorithm (41) is identical to the corresponding Kalman filter algorithm]. Thus the proof of the theorem is completed.

Having proved Theorem 5.1, which shows that the new algorithm is the optimal linear filter also in the non-Gaussian case, the presentation of the new V-Lambda filter is completed. In the next section we demonstrate its numerical robustness with an example.

VI. Filtering Example

In this section we present the results of a simple filtering example, in order to demonstrate that the new V-Lambda algorithm works satisfactorily and to demonstrate the superior numerical stability and accuracy of the new algorithm when compared to the convenient Kalman filter algorithm.

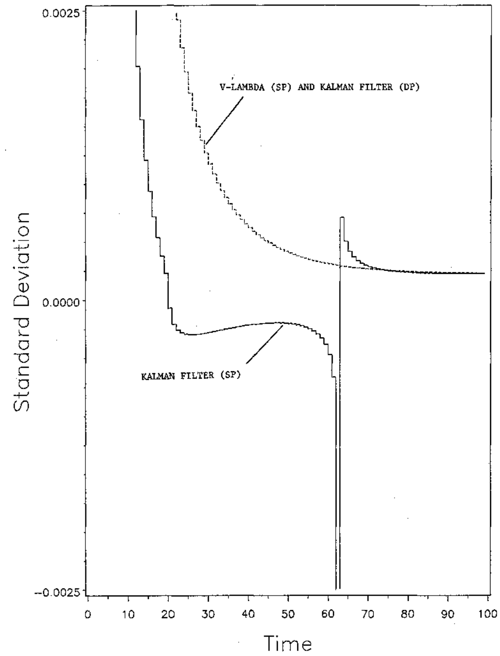

Fig. 3 Standard deviation of the estimation error of x_3.

Example 6.1

The dynamical system is described by the following discrete-time mathematical model:

$$x_{k+1} = F_k x_k + w_k \tag{62a}$$

$$y_k = H_k + v_k \tag{62b}$$

where

$$x \equiv [x_1, x_2, x_3]^T \quad F_k = \begin{pmatrix} 1 & 0.1 & 0.05 \\ 0 & 1 & 0.1 \\ 0 & 0 & 1 \end{pmatrix} \quad H_k = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0.01 & 0 \end{pmatrix}$$

$$E\{w_k\} = \mathbf{0} \qquad E\{w_j w_k^T\} = \text{diag}\,\{0.1E-7, 0.1E-7, 0.1E-8\}\delta_{jk}$$

$$E\{v_k\} = \mathbf{0} \qquad E\{v_j v_k^T\} = \text{diag}\,\{0.1E-4, 0.1E-4\}\delta_{jk}$$

$$P_0 \equiv E\{[x_0 - E\{x_0\}][x_0 - E\{x_0\}]^T\}$$

$$= \text{diag}\,\{0.25E+5, 0.25E+5, 0.25E+5\}$$

The initial state vector and its estimate are chosen as

$$x_0 = [0.01, 0.1, 1]^T, \qquad \hat{x}_0 = [1, 0.5, 0.005]$$

Both the V-Lambda filter and the conventional Kalman filter are used to obtain the estimate of the state vector. The V-Lambda filter is implemented in single precision (SP), while the conventional filter is used both in single and in double precision (DP). All runs were performed on a DEC VAX 8650 machine. In Figs. 1–3 we show the time histories of the standard deviations of the estimation error components as obtained by the three filters used. As can be seen from these figures, the V-Lambda filter and the DP conventional filter behave identically. The SP conventional filter loses numerical significance after about 20 s of estimation, which results in negative variances along the diagonal of the covariance matrix (the corresponding square roots are plotted as negative values in these figures). In Figs. 4–6 the time histories of the absolute values of the estimation error components are shown, as computed by the three filters. It is interesting to note that

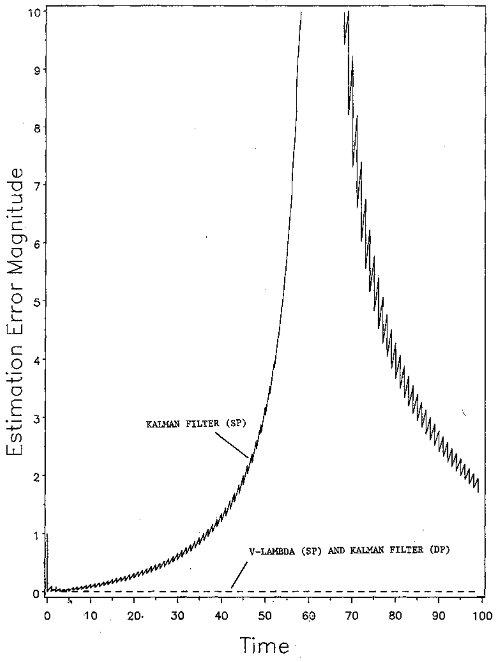

Fig. 4 Absolute value of the estimation error of x_1.

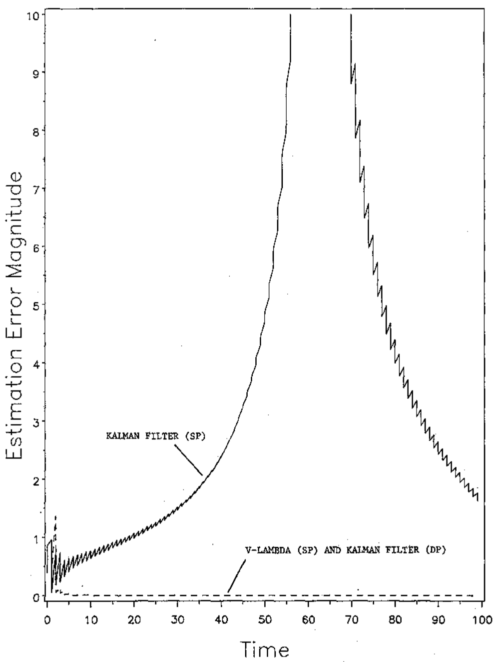

Fig. 5 Absolute value of the estimation error of x_2.

although the SP conventional filter diverges at a certain time point in the process, it regains stability afterwards; this phenomenon also has been observed and explained by Bellantoni and Dodge.[3]

VII. Conclusions

A new, information type, V-Lambda square root filtering algorithm is presented in this paper. The new algorithm is based on the singular value decomposition as the main computational tool, which renders it exceptionally numerically robust and accurate. The continuous availability of the covariance eigenfactors to the user is an additional merit of the proposed method.

Compared to other state-of-the-art square root filters, the new algorithm is computationally more costly (because of its reliance on the singular value decomposition). However, the advantages of the V-Lambda filter justify the additional com-

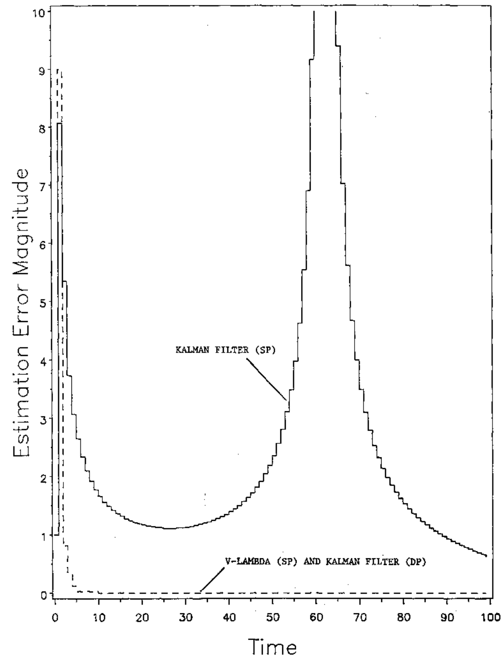

Fig. 6 Absolute value of the estimation error of x_3.

putation load. Moreover, because of the increased popularity of the singular value decomposition as a design tool in control problems and with the current vast development in the areas of parallel computation of the singular value decomposition, it is anticipated that the new algorithm eventually will become increasingly attractive as its computational requirements are reduced.

Acknowledgment

The author wishes to acknowledge a stimulating discussion with the late Dr. G. J. Bierman, an eminent leader in square root filtering, which motivated the search for the algorithm presented in this paper.

References

[1]Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 82D, March 1960, pp. 35–45.

[2]Kalman, R. E. and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," *Transactions of the ASME, Journal of Basic Engineering*, Vol. 83, 1961, pp. 95–108.

[3]Bellantoni, J. F. and Dodge, K. F., "A Square Root Formulation of the Kalman-Schmidt Filter," *AIAA Journal*, Vol. 5, July 1967, pp. 1309–1314.

[4]Potter, J. E. and Stern, R. G., "Statistical Filtering of Space Navigation Measurements," *Proceedings of the AIAA Guidance and Control Conference*, AIAA, New York, 1963.

[5]Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation*, Academic, New York, 1977.

[6]Maybeck, P. S., *Stochastic Models, Estimation and Control*, Vol. 1, Academic, New York, 1979.

[7]Lawson, C. L. and Hanson, R. J., *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[8]Bierman, G. J., "Measurement Updating Using the U-D Factorization," *Automatica*, Vol. 12, July 1976, pp. 375–382.

[9]Thornton, C. L. and Bierman, G. J., "Gram-Schmidt Algorithms for Covariance Propagation," *Proceedings of the IEEE Conference on Decision and Control*, Inst. of Electrical and Electronics Engineers, New York, 1975, pp. 489–498.

[10]Oshman, Y. and Bar-Itzhack, I. Y., "Square Root Filtering Via Covariance and Information Eigenfactors," *Automatica*, Vol. 22, 1986, pp. 599–604.

[11]Ham, F. M. and Brown, R. G., "Observability, Eigenvalues, and Kalman Filtering," *IEEE Transactions on Aerospace and Electrical Systems*, Vol. AES-19, 1983, pp. 269–273.

[12]Brent, R. P. and Luk, F. T., "The Solution of Singular Value and

Symmetric Eigenvalue Problems on Multiprocessor Arrays," *SIAM Journal of Scientific Statistical Computing*, Vol. 6, Jan. 1985, pp. 69–84.

[13]Luk, F. T., "Architectures for Computing Eigenvalues and Singular Value Decompositions," *Proceedings of SPIE, Vol. 614: Highly Parallel Signal Processing Architectures*, Bellingham, WA, 1986.

[14]Schimmel, D. E. and Luk, F. T., "A New Systolic Array for the Singular Value Decomposition," Cornell Univ., Ithaca, NY, TR EE-CEG-85-7, 1985.

[15]Cox, H., "On the Estimation of State Variables and Parameters for Noisy Dynamic Systems," *IEEE Transactions on Automatic Control*, Vol. AC-9, Jan. 1964, pp. 5–12.

[16]Anderson, B. D. O. and Moore, J. B., *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979, p. 26.

[17]Oshman, Y., *Eigenfactors Methods in Optimal Linear Estimation*, D.Sc. Dissertation, Dept. of Aeronautical Engineering, Technion—Israel Institute of Technology, Haifa, (in Hebrew), 1986.

[18]Golub, G. H. and Reinsch, C., "Singular Value Decomposition and Least Squares Solutions," *Numerische Mathematik*, Vol. 14, 1970, pp. 403–420.

[19]Hanson, R. J. and Lawson, C. L., "Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems," *Mathematics of Computation*, Vol. 23, Oct. 1969, pp. 787–812.

[20]Golub, G. H. and Van Loan, C. F., *Matrix Computations*, Johns Hopkins Univ. Press, Baltimore, MD, 1983, p. 146.