JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 23, No. 2, March–April 2000

New, Guaranteed Positive Time Update for the Two-Step Optimal Estimator

N. Jeremy Kasdin*
Princeton University, Princeton, New Jersey 08544

A new time update for the two-step optimal filter is presented. This time update eliminates the occurrence of low eigenvalues in the first-step covariance for certain systems. This event left the first-step covariance singular (or negative definite) and often resulted in an unstable implementation. The new time update is guaranteed positive definite and significantly improves the performance of the two-step estimator on all systems. The two-step filter is an alternative to the standard recursive estimators that are applied to nonlinear measurement problems, such as the extended and iterated extended Kalman filters. It improves the estimate error by splitting the cost function minimization into two steps (a linear first step and a nonlinear second step) by defining a set of first-step states that are nonlinear combinations of the desired states. A linear approximation is made in the time update of the first-step states rather than in the measurement update as in conventional methods.

I. Introduction

THE two-step filter[1,2] is a new approach for nonlinear recursive estimation that substantially improves the estimate error relative to the extended Kalman filter (EKF) or the iterated extended Kalman filter (IEKF). It accomplishes this by dividing the estimation problem into steps, a linear first step and a nonlinear second step. The result is a filter that comes much closer to minimizing the desired cost and dramatically reduces the biases and mean-square error of the EKF.

One potential pitfall of the two-step estimator has been the approximate time update.[1,2] Because of the subtraction in its implementation, certain dynamic systems may produce nonpositive covariance.[3] This paper presents an alternative solution to that problem by introducing a new version of the time update that is guaranteed positive definite. Before introducing these modifications, however, it is worthwhile to present a quick review of the two-step filter algorithm.

Recall that the dynamic estimation problem can be formulated as the optimal solution that minimizes the following quadratic cost function (least-squares estimation):

$$J = \frac{1}{2}(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1)^T M_{x_1}^{-1}(\boldsymbol{x}_1 - \bar{\boldsymbol{x}}_1) + \frac{1}{2}\sum_{k=1}^{N-1} \boldsymbol{w}_k^T Q_k^{-1}\boldsymbol{w}_k + \frac{1}{2}\sum_{k=1}^{N}[z_k - F(\boldsymbol{x}_k, t_k)]^T R_k^{-1}[z_k - F(\boldsymbol{x}_k, t_k)] \tag{1}$$

subject to the (discrete) dynamic constraint on the desired $n \times 1$ state vector

$$\boldsymbol{x}_{k+1} = \Phi_k \boldsymbol{x}_k + \Gamma_k \boldsymbol{w}_k \tag{2}$$

with the $\ell \times 1$ measurement vector z_k given by

$$z_k = F(\boldsymbol{x}_k, t_k) + \boldsymbol{v}_k \tag{3}$$

As usual, $\boldsymbol{w}$ represents a vector of Gaussian independent and identically distributed (i.i.d.) process noise with covariance matrix Q_d, and v represents a vector of Gaussian i.i.d. measurement noise with covariance matrix R. The subscript k denotes the value at time step t_k. The vector $\boldsymbol{x}_1$ represents the a priori state estimate with error covariance M_{x_1}.

For linear measurement problems, that is, problems where the measurement function F_k is linear in the states ($z_k = H_k \boldsymbol{x}_k + \nu_k$), the well-known recursive Kalman filter can be derived as the minimum solution using, for example, the sweep method (in this case the estimate is also the maximum likelihood and minimum variance solution).[4] Unfortunately, there is no such general solution to the stated nonlinear problem.

The typical alternative approach is to derive the recursive filter directly, rather than as a derived solution, from the global minimum of the cost function in Eq. (1). To do this, the fact that the linear Kalman filter can be derived in a number of different but equivalent ways is exploited to derive the nonlinear filter in a similar fashion. One common approach for the linear problem is to minimize a static quadratic cost function and then to manipulate the resulting solution to derive a recursive form. In other words, the linear Kalman filter can be derived by instead minimizing the following linear cost function for a batch solution of the static state:

$$J = \frac{1}{2}\sum_{k=1}^{N}(z_k - H_k \boldsymbol{x})^T R^{-1}(z_k - H_k \boldsymbol{x}) \tag{4}$$

The resulting solution provides an optimal estimate of the states $\boldsymbol{x}$ in terms of H and z_k, represented by $\hat{\boldsymbol{x}}$. This batch solution can then be manipulated to find a recursive equation for the estimate $\hat{\boldsymbol{x}}$, that is, an equation for an update to the estimate, given an additional measurement, that does not require reformulating the batch solution. The resulting formula is the same as the Kalman filter measurement update. The time update is then added using the equations for stochastic process propagation.[5] The time update is independent of the measurement and thereby provides a new a priori estimate for the recursive measurement update.

This observation (i.e., that the static problem with time update is equivalent to the batch dynamic problem) is exploited in the derivation of the two-step estimator.[1,2] The nonlinear measurement problem is recast as a static batch problem and solved for the best recursive estimate of the states. This static nonlinear cost function is given by

$$J = \frac{1}{2}\sum_{k=1}^{N}[z_k - F(\boldsymbol{x}, t_k)]^T R^{-1}[z_k - F(\boldsymbol{x}, t_k)] \tag{5}$$

The two-step filter formulation rewrites this cost function as a linear measurement in an $m \times 1$ intermediate (first-step) state vector $\boldsymbol{y}$:

$$J_y = \frac{1}{2}\sum_{k=1}^{N}(z_k - H\boldsymbol{y}_k)^T R^{-1}(z_k - H\boldsymbol{y}_k) \tag{6}$$

where $\boldsymbol{y}_k = \boldsymbol{f}_k(\boldsymbol{x}_k)$.

Received 28 October 1998; accepted for publication 24 May 1999.

*Assistant Professor, Department of Mechanical and Aerospace Engineering.

The resulting linear cost function can be solved exactly for $\hat{y}_k$ and also reformulated as a recursive update. The desired state estimate $\hat{x}_k$ is found in a second-step minimization with cost function

$$J_x = \tfrac{1}{2}[\hat{y}_k - f_k(x_k)]^T P_y^{-1}[\hat{y}_k - f_k(x_k)] \tag{7}$$

where P_y is the first-step state estimate error covariance. This optimization is usually performed using a Gauss–Newton algorithm.[1,2] It is straightforward to show that this two-step approach exactly minimizes the desired static cost function in Eq. (5).[1] Once formulated as a recursive update, the linear time update on x_k is used in the same fashion as in the standard, linear Kalman filter.

In the previously published work,[1–3] an approximate time update was used to find the updated first-step state estimate $\bar{y}_{k+1}$ given the exact update of the second-step states $\bar{x}_{k+1}$. This approximation is based on linearizing $f(x)$ about the previous best estimate of $\bar{x}_{k+1}$ and $\hat{x}_k$ and finding the expected value and covariance. This is the only approximation in the two-step filter algorithm and is performed to first order in Refs. 1 and 2. Unfortunately, the resulting expression contains the difference of two matrix quantities. It can be shown[3] that for certain dynamic systems and measurements this expression for the first-step covariance will drop rank.

References 1 and 2 outline a third derivation of the EKF and two-step filter and present further arguments to validate the improvement provided by the two-step filter, particularly for static systems. The remainder of this paper is devoted to reviewing the potential problem in the previous first-order implementation (Sec. II) and to presenting a new form of the time update that eliminates it (Sec. III). The example in Sec. IV verifies the filter's performance.

For reference, the original two-step filter formulation[1] is summarized by the following set of equations. Given measurements

$$z_k = F(x_k, t_k) + v_k = H_k y_k + v_k \qquad k = 1, \ldots, N$$

dynamics

$$x_{k+1} = \Phi_k x_k + \Gamma_k w_k \qquad k = 1, \ldots, N-1$$

nonlinearity

$$y_k = f_k(x_k)$$

and initial conditions

$$\bar{x}_1, M_{x_1}$$

where

$$E[v_k] = 0 \qquad E\left[v_k v_k^T\right] = R_k$$

$$E[w_k] = 0 \qquad E\left[w_k w_k^T\right] = Q_k$$

For the first-step optimization, the measurement update is

$$\hat{y}_k = \bar{y}_k + P_{y_k} H_k^T R_k^{-1}(z_k - H_k \bar{y}_k) \qquad k = 1, \ldots, N$$

$$P_{y_k} = \left(M_{y_k}^{-1} + H_k^T R_k^{-1} H_k\right)^{-1}$$

and the time update is

$$\bar{y}_{k+1} = \hat{y}_k + f_{k+1}(\bar{x}_{k+1}) - f_k(\hat{x}_k) \qquad k = 1, \ldots, N-1$$

$$M_{y_{k+1}} = P_{y_k} + \left.\frac{\partial f_{k+1}}{\partial x}\right|_{\bar{x}_{k+1}} M_{x_{k+1}} \left.\frac{\partial f_{k+1}}{\partial x}\right|_{\bar{x}_{k+1}}^T - \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_k} P_{x_k} \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_k}^T$$

For the second-step optimization, the measurement update (Gauss–Newton iteration on i) is

$$\hat{x}_{k,i+1} = \hat{x}_{k,i} - H_{G_{k,i}}^{-1} q_{k,i}^T \qquad k = 1, \ldots, N$$

$$H_{G_{k,i}} = \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_{k,i}}^T P_{y_k}^{-1} \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_{k,i}}$$

$$q_{k,i} = -[\hat{y}_k - f_k(\hat{x}_{k,i})]^T P_{y_k}^{-1} \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_{k,i}}$$

$$P_{x_{k,i}} = H_{G_{k,i}}^{-1}$$

and the time update is

$$\bar{x}_{k+1} = \Phi_k \hat{x}_k \qquad k = 1, \ldots, N-1$$

$$M_{x_{k+1}} = \Phi_k P_{x_k} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T$$

One additional caveat should be mentioned. Because of the least-squares minimization necessary to determine the second-step states [Eq. (7)], the number of first-step states must equal or exceed the number of second-step states ($m \geq n$). Violation of this requirement would result in a nonunique minimum for the second-step optimization. For situations where other means of satisfying this requirement are not practical, the suggested approach is to simply augment the choice of first-step states with the second-step state vector x (see Refs. 1 and 2). This sensitivity of the filter to the relationship between y and x is a reflection of the problem of observability. Problems for which the number of desired states x exceeds the number of possible choices for first-step states y are ones that have so few measurements at each step that the dynamic model is heavily relied on for creating observability. In nonlinear problems this can result in very poor performance when using the EKF and IEKF (single site radar ranging[1,2] is an example of such a problem). Although the two-step filter can be made to perform quite well on these types of problems, with biases substantially lower than the EKF and IEKF, the designer is always better off adding additional measurements, when possible, to improve observability at each measurement update.

II. Original Two-Step Time Update and Ill Conditioning

The original first-step time update was derived using the identities[1,2]

$$y_k = f_k(x_k, t_k) \tag{8}$$

and

$$y_{k+1} = y_k + f_{k+1}(x_{k+1}) - f_k(x_k) \tag{9}$$

where x is the desired $n \times 1$ state vector. The time update for the first step states y is found by taking the expectation of Eq. (9). Unfortunately, the nonlinear measurement $f(x)$ makes this impossible in practice. The expectation is made tractable by expanding the last two terms to first order in the second-step estimation error[1,2]:

$$f_k(x_k) \approx f_k(\hat{x}_k) + \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_k} (x_k - \hat{x}_k) \tag{10}$$

and likewise for $f_{k+1}(x_{k+1})$ about $\bar{x}_{k+1}$. By taking the expected value of this expanded version and assuming the distribution of estimation error is unbiased, the first-step time update is as presented in Sec. I:

$$\bar{y}_{k+1} \approx \hat{y}_k + f_{k+1}(\bar{x}_{k+1}) - f_k(\hat{x}_k) \tag{11}$$

The covariance is likewise found by subtracting Eq. (11) from Eq. (9), expanding to first order, squaring, and taking the expected value. After much algebra, the result is again as presented in Sec. I:

$$M_{y_{k+1}} \approx P_{y_k} + \left.\frac{\partial f_{k+1}}{\partial x}\right|_{\bar{x}_{k+1}} M_{x_{k+1}} \left.\frac{\partial f_{k+1}}{\partial x}\right|_{\bar{x}_{k+1}}^T - \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_k} P_{x_k} \left.\frac{\partial f_k}{\partial x}\right|_{\hat{x}_k}^T \tag{12}$$

These equations have the desirable property that they are updates for the first-step states and covariance. That is, they are small corrections on the previous estimates as in the traditional Kalman filter. However, it is not difficult to see that the covariance update in Eq. (12) has potential problems. A true covariance is always positive definite and symmetric. Because of the subtraction in Eq. (12), though, the covariance of the first-step states can potentially drop rank or go negative. This could be catastrophic for the filter.

Unfortunately, such behavior has been seen in simulations, particularly those involving the second-order filter.[2] Garrison et al. have shown that this event is predictable and depends on the system, measurement, and trajectory.[3] This can be seen by first formulating the matrix C:

$$C \equiv P_y - \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\hat{x}} P_x \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\hat{x}}^T \tag{13}$$

and then performing a numerical rank test on the following matrix,

$$\operatorname{rank}\left(\left[C, \frac{\partial \boldsymbol{f}}{\partial x_1}, \frac{\partial \boldsymbol{f}}{\partial x_2}, \ldots, \frac{\partial \boldsymbol{f}}{\partial x_m}\right], \varepsilon\right) < n \tag{14}$$

where the partial derivatives of $\boldsymbol{f}$ are computed along a reference trajectory and ε is a specified tolerance. The points where this matrix drops rank are those at which the first-step covariance will become ill conditioned.

The importance of Eqs. (13) and (14) is not in their ability to predict where the time update may fail, but rather in their ability to explain the nature of such a failure.[3] Although it may be possible to prevent failure of the algorithm using these equations by monitoring Eq. (14), this has not been shown to work in all cases, and it greatly complicates the filter. The only proposed solution to date[3] has been a small augmentation to the time update in Eq. (12) to keep the small eigenvalues positive,

$$\begin{aligned} M_{y_{k+1}} \approx P_{y_k} &+ \left.\frac{\partial \boldsymbol{f}_{k+1}}{\partial \boldsymbol{x}}\right|_{\bar{x}_{k+1}} M_{x_{k+1}} \left.\frac{\partial \boldsymbol{f}_{k+1}}{\partial \boldsymbol{x}}\right|_{\bar{x}_{k+1}}^T \\ &- \left.\frac{\partial \boldsymbol{f}_k}{\partial \boldsymbol{x}}\right|_{\hat{x}_k} P_{x_k} \left.\frac{\partial \boldsymbol{f}_k}{\partial \boldsymbol{x}}\right|_{\hat{x}_k}^T + \varepsilon I \end{aligned} \tag{15}$$

Here, ε is a small positive scalar chosen large enough to prevent P_y from becoming ill conditioned. Though this seems to work for the test problems tried, it is ad hoc and has not been proven to be a global solution. In the next section an altogether new time update approach is presented that guarantees a positive definite, symmetric covariance.

III. New Time Update

Before presenting the new algorithm, it is necessary to expand the class of problems to which the two-step filter is applied. Previously, the dynamics were limited to linear, discrete equations. This was done merely to simplify presentation of the algorithm and maintain focus on the changes to the measurement update. As the two-step filter did not change the time update on the second step states $\boldsymbol{x}$, there was no reason to exclude nonlinear, continuous dynamics. It is simply necessary to integrate numerically $\boldsymbol{x}$ and its covariance using any standard approach.

For the new approach, however, it is necessary to return to the more general continuous formulation. The problem is now stated as a search for the best estimate of the state vector $\boldsymbol{x}$ (in a least-squares sense) at time t_k given the nonlinear measurements

$$z_k = \boldsymbol{F}(\boldsymbol{x}_k, t_k) + \boldsymbol{v}_k \tag{16}$$

where $\boldsymbol{x}$ is subject to the dynamic constraint

$$\dot{\boldsymbol{x}} = A(\boldsymbol{x}, t) + B(\boldsymbol{x}, t)\boldsymbol{w}(t) \tag{17}$$

and $\boldsymbol{w}(t)$ is a white noise stochastic process with autospectral density matrix Q. In other words, $\boldsymbol{x}(t)$ is the solution to the Ito stochastic differential equation given by Eq. (17).

The measurement update is found in the same manner as before, by forming a first-step state vector as a nonlinear combination of the desired second-step states $\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$ and using the equations in Sec. I with the discrete measurements from Eq. (16).[1,2]

For the time update we must find an expression for propagating $\boldsymbol{y}$ given the stochastic differential equation for $\boldsymbol{x}$ in Eq. (17). This is done by using the standard expression for the Ito chain rule. That is, the time derivative of a nonlinear function of $\boldsymbol{x}$ can be found using the modified chain rule of the Ito calculus and Eq. (17)[6,7] Doing this on $\boldsymbol{f}(\boldsymbol{x})$ results in the following stochastic differential equation for each component of $\boldsymbol{y}$:

$$\begin{aligned} \dot{y}_i = \frac{\partial f_i}{\partial t} + \frac{\partial f_i}{\partial \boldsymbol{x}} A(\boldsymbol{x}, t) + \frac{1}{2}\operatorname{tr}\left\{\frac{\partial^2 f_i}{\partial \boldsymbol{x}^2} B(\boldsymbol{x}, t) Q B^T(\boldsymbol{x}, t)\right\} \\ + \frac{\partial f_i}{\partial \boldsymbol{x}} B(\boldsymbol{x}, t)\boldsymbol{w}(t) \end{aligned} \tag{18}$$

where tr{ } refers to the trace of the enclosed matrix expression. This equation is now used along with Eq. (17) to numerically propagate $\boldsymbol{x}$ and $\boldsymbol{y}$.

Recall that in the two-step filter formulation it is a requirement that there be at least as many first-step states as second-step states. For cases where such a factorization of $F(\boldsymbol{x}, t)$ is difficult to find, it was proposed that the measurement simply be augmented with the state to form a larger first-step state vector. Now, this augmentation is required because the states $\boldsymbol{x}$ and $\boldsymbol{y}$ and their covariance must be simultaneously numerically integrated. Thus, the first-step states are always going to be given by an expression that looks like

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \hat{\boldsymbol{F}}(\boldsymbol{x}, t) \\ \boldsymbol{x}(t) \end{bmatrix} \tag{19}$$

where $\hat{\boldsymbol{F}}(\boldsymbol{x}, t)$ may be just $\boldsymbol{F}(\boldsymbol{x}, t)$, as proposed earlier, or it may be some other convenient combination.

What remains needed is an algorithm for performing this nonlinear propagation. Although this is an area of current research, a conventional, Runge–Kutta based technique for the numerical integration has so far been adequate. In this approach, the nonlinearities are expanded to first order in the estimate error and assumed central. The expected value is taken to find the deterministic differential equations for the estimates,

$$\begin{aligned} \dot{\bar{\boldsymbol{y}}} &= \begin{bmatrix} \dfrac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{F}}(\bar{\boldsymbol{x}}, t) \\ \dot{\bar{\boldsymbol{x}}} \end{bmatrix} \\ &= \begin{bmatrix} \left.\dfrac{\partial f_1}{\partial t}\right|_{\hat{x}} + \left.\dfrac{\partial f_1}{\partial \boldsymbol{x}}\right|_{\hat{x}} A(\hat{\boldsymbol{x}}, t) + \dfrac{1}{2}\operatorname{tr}\left\{\left.\dfrac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\right|_{\hat{x}_k} B(\hat{\boldsymbol{x}}, t) Q B^T(\hat{\boldsymbol{x}}, t)\right\} \\ \vdots \\ \left.\dfrac{\partial f_\ell}{\partial t}\right|_{\hat{x}} + \left.\dfrac{\partial f_\ell}{\partial \boldsymbol{x}}\right|_{\hat{x}} A(\hat{\boldsymbol{x}}, t) + \dfrac{1}{2}\operatorname{tr}\left\{\left.\dfrac{\partial^2 f_\ell}{\partial \boldsymbol{x}^2}\right|_{\hat{x}} B(\hat{\boldsymbol{x}}, t) Q B^T(\hat{\boldsymbol{x}}, t)\right\} \\ A(\hat{\boldsymbol{x}}, t) \end{bmatrix} \\ &\equiv \tilde{A}(\hat{\boldsymbol{x}}, t) \end{aligned} \tag{20}$$

For the time update, this equation is propagated using the classical, third- or fourth-order Runge–Kutta formula to find $\bar{\boldsymbol{y}}_{k+1}$.

The covariance is also found by linearizing the augmented dynamics $\tilde{A}(\boldsymbol{x}, t)$ and then squaring and taking the expected value to find the following linear differential equation for the covariance:

$$\dot{P}_y = \left.\frac{\partial \tilde{A}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}}\right|_{\hat{x}} P_y \left.\frac{\partial \tilde{A}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}}\right|_{\hat{x}}^T + \tilde{B}(\hat{\boldsymbol{x}}, t) Q \tilde{B}^T(\hat{\boldsymbol{x}}, t) \tag{21}$$

where

$$\tilde{B}(\boldsymbol{x}, t) = \begin{bmatrix} \dfrac{\partial f_1}{\partial \boldsymbol{x}} B(\boldsymbol{x}, t) \\ \vdots \\ \dfrac{\partial f_\ell}{\partial \boldsymbol{x}} B(\boldsymbol{x}, t) \\ B(\boldsymbol{x}, t) \end{bmatrix} \tag{22}$$

This equation for the covariance is most easily propagated using the discrete update version,

$$P_y(k+1) = \Phi(t_{k+1}, t_k) P_y(k) \Phi^T(t_{k+1}, t_k) + Q_d(k) \tag{23}$$

where the state transition matrix is the solution of the linear differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, t_0) = \left.\frac{\partial \tilde{A}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}}\right|_{\hat{x}} \Phi(t, t_0) \tag{24}$$

and the discrete process noise covariance matrix is the solution of the integral

$$Q_d(k) = \Phi_k\left\{\int_{t_{k-1}}^{t_k} \Phi^{-1}(\tau)\tilde{B}(\hat{\boldsymbol{x}}, \tau)Q\tilde{B}^T(\hat{\boldsymbol{x}}, \tau)\Phi^{-T}(\tau)\,\mathrm{d}\tau\right\}\Phi_k^T \tag{25}$$

Equations (24) and (25) are most easily solved using the algorithm due to Van Loan[8] (see also Ref. 9). Equation (23) is the new time update for P_y. It is guaranteed to be positive definite and symmetric, thus eliminating the robustness problems of the original two-step filter.

This new version of the two-step estimator may now be summed up as follows.

Given measurments

$$z_k = F(\boldsymbol{x}_k, t_k) + \boldsymbol{v}_k = H_k\boldsymbol{y}_k + \boldsymbol{v}_k \qquad k = 1, \ldots, N$$

dynamics

$$\dot{\boldsymbol{x}} = A(\boldsymbol{x}, t) + B(\boldsymbol{x}, t)\boldsymbol{w}(t)$$

nonlinearity

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix}\hat{\boldsymbol{F}}(\boldsymbol{x}, t)\\ \boldsymbol{x}(t)\end{bmatrix}$$

and initial conditions

$$\bar{\boldsymbol{x}}_1, M_{x_1}$$

where

$$E[\boldsymbol{v}_k] = 0 \qquad E\left[\boldsymbol{v}_k\boldsymbol{v}_k^T\right] = R_k$$

$$E[\boldsymbol{w}(t)] = 0 \qquad E[\boldsymbol{w}(t)\boldsymbol{w}(t+\tau)^T] = Q\delta(\tau)$$

For the first-step optimization, the measurment update is

$$\hat{\boldsymbol{y}}_k = \bar{\boldsymbol{y}}_k + P_{y_k}H_k^T R_k^{-1}(z_k - H_k\bar{\boldsymbol{y}}_k) \qquad k = 1, \ldots, N$$

$$P_{y_k} = \left(M_{y_k}^{-1} + H_k^T R_k^{-1} H_k\right)^{-1}$$

and the time update is

$$\dot{\bar{\boldsymbol{y}}} = \begin{bmatrix}\frac{\mathrm{d}}{\mathrm{d}t}\hat{\boldsymbol{F}}(\bar{\boldsymbol{x}}, t)\\ \dot{\bar{\boldsymbol{x}}}\end{bmatrix}$$

$$= \begin{bmatrix}\left.\frac{\partial f_1}{\partial t}\right|_{\hat{x}} + \left.\frac{\partial f_1}{\partial \boldsymbol{x}}\right|_{\hat{x}} A(\hat{\boldsymbol{x}}, t) + \frac{1}{2}\mathrm{tr}\left\{\left.\frac{\partial^2 f_1}{\partial \boldsymbol{x}^2}\right|_{\hat{x}_k} B(\hat{\boldsymbol{x}}, t)QB^T(\hat{\boldsymbol{x}}, t)\right\}\\ \vdots\\ \left.\frac{\partial f_\ell}{\partial t}\right|_{\hat{x}} + \left.\frac{\partial f_\ell}{\partial \boldsymbol{x}}\right|_{\hat{x}} A(\hat{\boldsymbol{x}}, t) + \frac{1}{2}\mathrm{tr}\left\{\left.\frac{\partial^2 f_\ell}{\partial \boldsymbol{x}^2}\right|_{\hat{x}} B(\hat{\boldsymbol{x}}, t)QB^T(\hat{\boldsymbol{x}}, t)\right\}\\ A(\hat{\boldsymbol{x}}, t)\end{bmatrix}$$

$$\equiv \tilde{A}(\hat{\boldsymbol{x}}, t)$$

$$\dot{P}_y = \left.\frac{\partial \tilde{A}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}}\right|_{\hat{x}} P_y \left.\frac{\partial \tilde{A}(\boldsymbol{x}, t)}{\partial \boldsymbol{x}}\right|_{\hat{x}}^T + \tilde{B}(\hat{\boldsymbol{x}}, t)Q\tilde{B}^T(\hat{\boldsymbol{x}}, t)$$

where

$$\tilde{B}(\boldsymbol{x}, t) = \begin{bmatrix}\frac{\partial f_1}{\partial \boldsymbol{x}}B(\boldsymbol{x}, t)\\ \vdots\\ \frac{\partial f_\ell}{\partial \boldsymbol{x}}B(\boldsymbol{x}, t)\\ B(\boldsymbol{x}, t)\end{bmatrix}$$

For the second-step optimization, the measurement update (Gauss–Newton iteration on i) is

$$\hat{\boldsymbol{x}}_{k,i+1} = \hat{\boldsymbol{x}}_{k,i} - H_{G_{k,i}}^{-1}\boldsymbol{q}_{k,i}^T \qquad k = 1, \ldots, N$$

$$H_{G_{k,i}} = \left.\frac{\partial \boldsymbol{f}_k}{\partial \boldsymbol{x}}\right|_{\hat{x}_{k,i}}^T P_{y_k}^{-1}\left.\frac{\partial \boldsymbol{f}_k}{\partial \boldsymbol{x}}\right|_{\hat{x}_{k,i}}$$

$$\boldsymbol{q}_{k,i} = -[\hat{\boldsymbol{y}}_k - \boldsymbol{f}_k(\hat{\boldsymbol{x}}_{k,i})]^T P_{y_k}^{-1}\left.\frac{\partial \boldsymbol{f}_k}{\partial \boldsymbol{x}}\right|_{\hat{x}_{k,i}} \qquad P_{x_{k,i}} = H_{G_{k,i}}^{-1}$$

IV. Example

The simple two-state example of Ref. 3 is used to demonstrate the effectiveness of the new technique. This example was chosen because it readily demonstrates the problem of ill conditioning in the original two-step estimator, yet is straightforward and simple to implement. It was designed to be representative (in two dimensions) of a more complex problem of relative navigation in elliptical orbits.[10] This orbital mechanics problem was also seen to exhibit the ill conditioning problem with a catastrophic failure of the filter.

The example is a purely kinematic motion of a particle on a spiral trajectory defined by a constant angular velocity ω_0 and a constant radial velocity v_0 (see Fig. 1). The desired (second-step) state vector consists of the two-dimensional position of the particle (x_1, x_2). The motion of the particle in a Cartesian coordinate system is described by two first-order, nonlinear differential equations,

$$\begin{aligned}\frac{\mathrm{d}x_1}{\mathrm{d}t} &= \frac{x_1 v_0}{\sqrt{x_1^2 + x_2^2}} - x_2\omega_0 + w_1\\ \frac{\mathrm{d}x_2}{\mathrm{d}t} &= \frac{x_2 v_0}{\sqrt{x_1^2 + x_2^2}} + x_1\omega_0 + w_2\end{aligned} \tag{26}$$

Here, w_1 and w_2 are process noise terms with diagonal covariance matrix Q.

The objective of the filter is to find estimates of the states x_1 and x_2 given only range measurements from a fixed point located at (1, 0). The measurement equation is thus given by

$$z = \sqrt{(x_1 - 1)^2 + x_2^2} \tag{27}$$

The new two-step filter is formed using the augmented first-step state vector

$$\boldsymbol{y} = \begin{bmatrix}\sqrt{(x_1 - 1)^2 + x_2^2}\\ x_1\\ x_2\end{bmatrix} \tag{28}$$

with measurement matrix $H = [1\ \ 0\ \ 0]$.

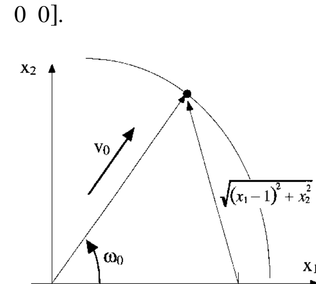

Fig. 1 Kinematic particle on a spiral trajectory.

Table 1 Parameter values for the example problem

Parameters	Values
Time step, Δt	0.0002
Filter initial state, $\hat{\boldsymbol{x}}_0$	(2.0, 0.0)
True initial state, $\boldsymbol{x}(0)$	(2.583, 0.313)
Measurement covariance, R	10^{-4}
Filter initial second-step state covariance, P_{x_0}	diag{0.25, 0.25}
Filter initial first-step state covariance, P_{y_0}	$\begin{bmatrix} 0.226 & 0.209 & 0.000 \\ 0.209 & 0.247 & 0.000 \\ 0.000 & 0.000 & 0.248 \end{bmatrix}$
Second-step continuous process noise, Q	diag$\{10^{-12}, 10^{-12}\}$

Table 1 lists the values used in the simulation.[3] The particle starts out at the position (2.583, 0.313) and propagates for 6 s with $\omega_0 = 1$ and $v_0 = 1$. Although a process noise term is used in the filter model, no process noise is added to the simulation of the truth model. Note that the very small time step is used to display better the sudden drop in eigenvalue of P_y in the original first step filter. Plots of the reference trajectory and of the simulated noisy measurements are shown in Fig. 2.

The original two-step estimator was used to obtain estimates of x_1 and x_2 as in Ref. 3. The result is shown in Fig. 3. As expected, Fig. 3 shows that the first-step covariance drops rank twice during the trajectory. The first such occurrence proves catastrophic for the estimator.

Figure 4 shows the results of using the new time update in the two-step filter on this example problem. For the example, the new dynamic equations of motion for the first-step states were found using the formula in Eq. (20),

$$\dot{\boldsymbol{y}} = \begin{bmatrix} \dot{r} \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \dfrac{\sqrt{x_1^2 + x_2^2}}{r} v_0 - \dfrac{x_1 v_0}{r\sqrt{x_1^2 + x_2^2}} + \dfrac{x_2 \omega_0}{r} + \dfrac{q}{2r} \\ \dfrac{x_1 v_0}{\sqrt{x_1^2 + x_2^2}} - x_2 \omega_0 \\ \dfrac{x_2 v_0}{\sqrt{x_1^2 + x_2^2}} + x_1 \omega_0 \end{bmatrix}$$

$$+ \begin{bmatrix} \dfrac{x_1 - 1}{r} & \dfrac{x_2}{r} \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{29}$$

where $r = \sqrt{[(x_1 - 1)^2 + x_2^2]}$ and q is the magnitude of each diagonal element of Q. The new two-step filter was implemented as described here using a fourth-order Runge–Kutta update on the dynamics from Eq. (29) and the linearized covariance update given in Eqs. (21–25).

After only a brief comparison of Figs. 3 and 4, the dramatic improvement afforded by the new time update is apparent. The first-step covariance is always positive and full rank and well-behaved. The estimate error rapidly becomes small and behaves as predicted by the covariance bounds. As a further validation of the new two-step estimator, a series of simulations were performed comparing performance with the EKF using identical random seeds. The results of one such run are shown in Fig. 5. Again, the improvement provided by the two-step estimator is significant. Note that the average result of a Monte Carlo simulation is not shown. Rather, Fig. 5 actually displays the best EKF result. Figure 6 shows a comparison of the bias in the new two-step estimator and the EKF after 100 averages. It was found that in 100 runs a small percentage of the EKF cases resulted in a wildly divergent error that substantially skewed the EKF averages. To date, there has been no evidence of poor performance in the two-step estimator.

Also note from Fig. 5 that the predicted covariance of the EKF is actually smaller than the two-step filter. However, because of the large biases in the EKF, the two-step covariance is a much better

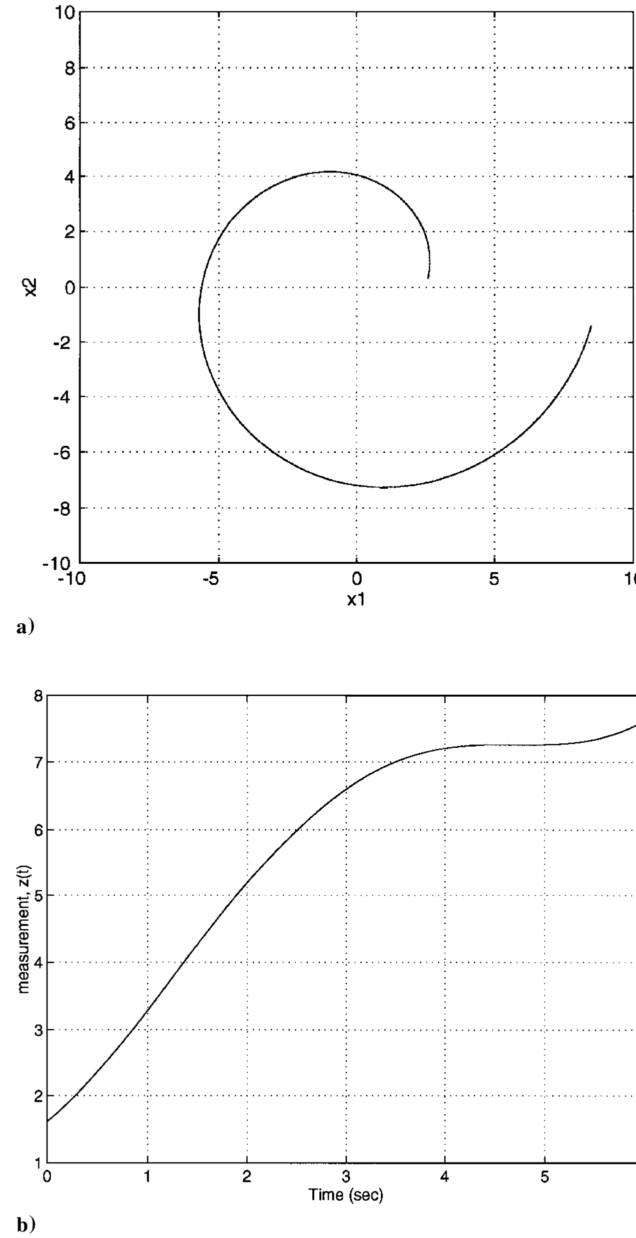

Fig. 2 Example problem trajectory and measurements.

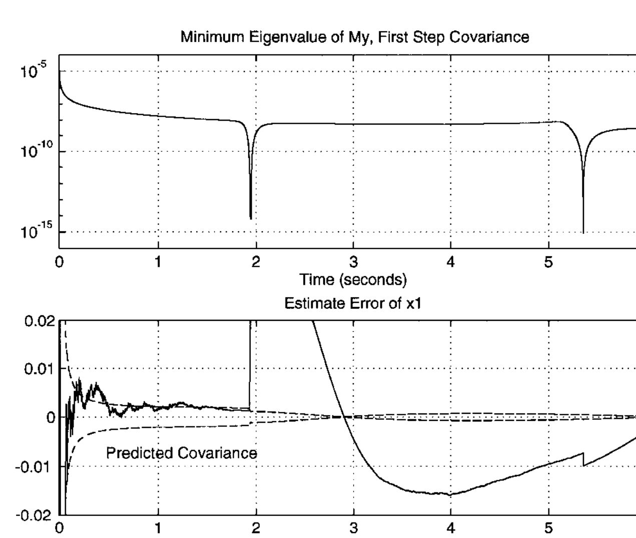

Fig. 3 Original two-step estimator results.

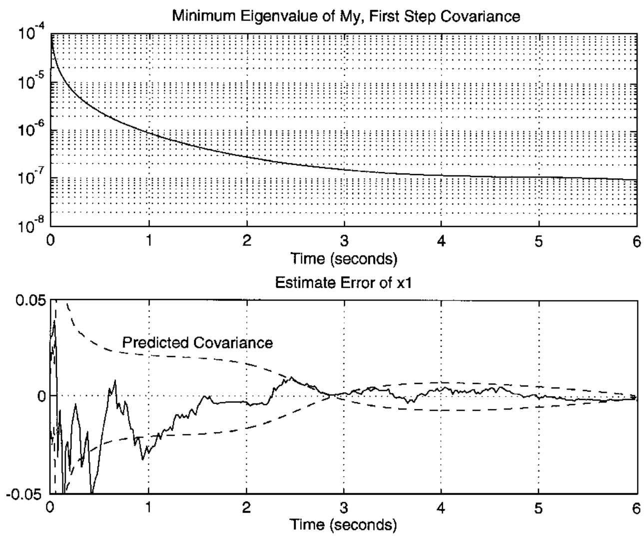

Fig. 4 New two-step estimator results.

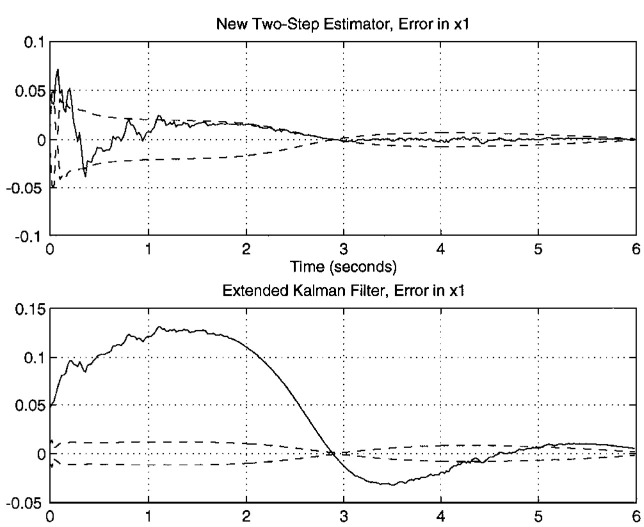

Fig. 5 Comparison of new two-step filter with EKF.

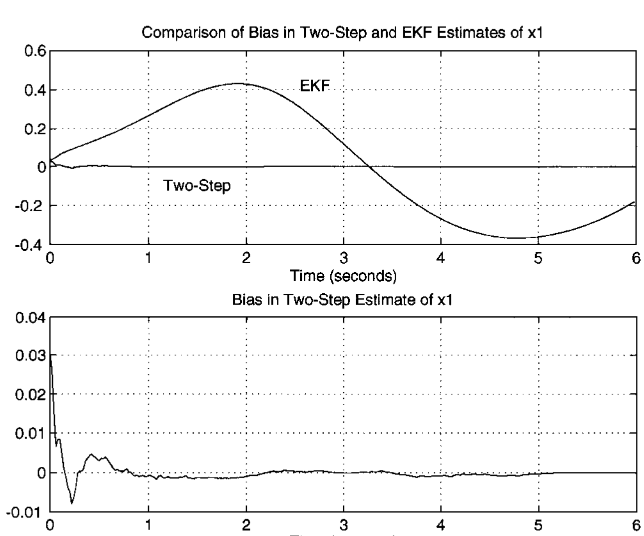

Fig. 6 Comparison of bias in EKF and two-step estimator (100 averages).

predictor of actual performance. It is this underestimation of the error in the EKF that is one of the prime culprits of its large biases.

V. Practical Considerations

Finally, some comments are in order regarding implementation of the two-step filter. After a first review, the filter appears substantially more complicated than the EKF. In fact, very little extra work is required to formulate this new version of the two-step estimator. The various Jacobians of the measurement and dynamics need to be found in either case. Given the dramatic improvement in performance and robustness, the small amount of additional analysis required to formulate the filter is well worth the effort.

It is true that the two-step filter will take longer to execute than the EKF. There are two reasons for this, the added Gauss–Newton search to find the optimal second-step states and the increased dimension of the first-step states. No specific performance studies have been done to date as that would require a real-time implementation optimized for speed in each case. The straightforward (though admittedly suboptimal) MATLAB® studies done for this paper showed a 2–2.5 times speed increase for the two-step estimator. These are very conservative estimates as they include various housekeeping and data storage steps not present in a real-time implementation.

In the unlikely event that the two-step implementation proves to be too slow for a given application, there are steps that can be taken to speed performance at the expense of slight degradations in accuracy. In particular, the Gauss–Newton step can be eliminated as both the measurement and time update are formulated in terms of the first-step states. Inaccuracies arise because the linearizations are now no longer found with the optimal values of the states. Simulations have shown very little degradation with still significant improvement over the EKF.

Finally, the same numerical and word length concerns that arise in the standard Kalman filter and the EKF arise here. Because the format of this filter is the same as the Kalman filter, the same techniques used there (e.g., the square-root implementation) can be applied here. Reference 2 has a thorough discussion of the square-root implementation for the original two-step filter. Reference 2 also discusses various techniques for initializing the covariance of the first-step states that applies equally well to this new version of the filter.

VI. Conclusions

The two-step estimator is a new technique for performing recursive estimation on problems with nonlinear measurements. It substantially improves filter performance (as measured by bias and mean-square error) over the traditional estimation techniques (the EKF and IEKF). The improvement, in essence, came from moving the approximations out of the filter measurement update and into the time update by splitting the problem into two steps (a linear estimation and a nonlinear fit).

This paper presented a new approach to performing the first-step time update in the filter. This new approach eliminates the potential problems associated with ill-conditioned first-step covariance matrices that sometimes appeared in the original formulation. By using the Ito stochastic chain rule, it became possible to derive a differential description for the time variation of the first-step states. This equation is used directly, with traditional techniques, to perform the first-step time update. This new time update has the enormous advantage of being guaranteed positive at all times.

Simulation results verify the effectiveness of the new time update. It outperforms both the EKF and the original two-step estimator. With this new tool, the two-step estimator should prove to be an extremely useful technique for performing estimation of systems with nonlinear measurements.

References

[1] Haupt, G. T., Kasdin, N. J., Keiser, G. M., and Parkinson, B. W., "An Optimal Recursive Iterative Algorithm for Discrete Nonlinear Least-Squares Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 3, 1996, pp. 643–649.

[2] Kasdin, N. J., and Haupt, G. T., "Second-Order Correction and Numerical Considerations for the Two-Step Optimal Estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 2, 1997, pp. 362–369.

[3]Garrison, J. L., Axelrad, P., and Kasdin, N. J., "Ill-Conditioned Covariance Matrices in the First-Order Two-Step Estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 5, 1998, pp. 754–760.

[4]Bryson, A. E., and Ho, Y. C., *Applied Optimal Control*, Hemisphere, Washington, DC, 1975, pp. 151.

[5]Stengel, R. F., *Stochstic Optimal Control*, Wiley, New York, 1986, pp. 315–340.

[6]Maybeck, P. S., *Stochastic Models, Estimation, and Control, Volume 2*, Academic, New York, 1982, p. 186.

[7]Kloeden, P. E., and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer–Verlag, Berlin, 1992, pp. 90–98.

[8]Van Loan, C. F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control*, Vol. AC-23, June 1978, pp. 395–404.

[9]Kasdin, N. J., "Discrete Simulation of Colored Noise and Stochastic Processes and $1/f^{\alpha}$ Power Law Noise Generation," *Proceedings of the IEEE*, Vol. 83, No. 5, 1995, pp. 802–827.

[10]Garrison, J. L., "Recursive Nonlinear Estimation for Relative Navigation in Elliptical Orbits," Ph.D. Dissertation, Dept. of Aerospace Engineering Sciences, Univ. of Colorado, Boulder, CO, Dec. 1997.