

(Preprint) AAS 12-xxx

RECURSIVE IMPLEMENTATIONS OF THE CONSIDER FILTER

Renato Zanetti* and Chris D'Souza†

One method to account for parameters errors in the Kalman filter is to consider their effect in the so-called Schmidt-Kalman filter. This work addresses issues that arise when implementing a consider Kalman filter as a real-time, recursive algorithm. A favorite implementation of the Kalman filter as an onboard navigation subsystem is the UDU formulation. A new way to implement a UDU consider filter is proposed. The non-optimality of the recursive consider filter is also analyzed, and a modified algorithm is proposed to overcome this limitation.

INTRODUCTION

Consider analysis was introduced by S. F. Schmidt of NASA Ames in the mid 1960s as a means to account for errors in both the dynamic and measurement models due to uncertain parameters [1]. The consider Kalman filter, also called the Schmidt-Kalman filter resulted from this body of work. The consider approach is especially useful when parameters have low observability or when the extra computational power to estimate them is not deemed necessary [2].

Schmidt's approach is based on minimum variance estimation, Jazwinski [3] details the derivation of the optimal consider Kalman filter. In Section 8.2 of his book, Bierman [4] disputes the optimality of the Schmidt-Kalman filter, at least in its sequential implementation. Nevertheless the Schmidt-Kalman filter has received considerable attention in recent years. Tapley *et al.* [5] give an ample description of the consider Kalman filter, they provide a different methodology than Jazwinski and arrive to a different formulation. Woodbury *et al.* provide new insight into considering parameters in the measurement model [6]. Equivalent formulations to the consider filter were also studied [7, 8] and applied to Mars entry navigation [9] and orbit determination [10].

While the Schmidt-Kalman filter is very well known, not much attention has been given to actual implementations of it in a real-time recursive estimation algorithm. Onboard estimators commonly utilize the UDU formulation, which guarantees symmetry and positive definiteness of the covariance. To date there is no direct way of including a consider update into the UDU formulation, this paper provides a simple algorithm to obtain this capability. Another common practice is to process measurements one at the time (which is inevitable when utilizing the UDU algorithm). When processing measurements one at the time in the

*Senior Member of the Technical Staff, Vehicle Dynamics and Controls, The Charles Stark Draper Laboratory, 17629 El Camino Real, Suite 470, Houston, Texas, 77058. rzanetti@draper.com

†GN&C Autonomous Flight Systems Engineer, Aeroscience and Flight Mechanics Division, NASA Johnson Space Center EG6, 2101 NASA Parkway, Houston, Texas, 77058. chris.dsouza@nasa.gov

Schmidt-Kalman filter the order in which they are processed affects the final result. This work analyzes this phenomenon and proposes a globally optimal consider filter, which addresses the objections raised by Bierman.

THE KALMAN FILTER UPDATE

This work assumes linear measurements and dynamics, the extension to the nonlinear case can be readily obtained using standard extended Kalman filter techniques. Let \mathbf{y} be a set of measurements of a state vector \mathbf{x} corrupted by zero mean noise $\boldsymbol{\eta}$ with covariance \mathbf{R}

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}, \quad (1)$$

where \mathbf{H} is the measurement mapping (or sensitivity) matrix. Let $\hat{\mathbf{x}}^-$ be an unbiased estimate of \mathbf{x} with corresponding estimation error covariance given by \mathbf{P}^- . The *a priori* estimation error is given by

$$\mathbf{e}^- = \mathbf{x} - \hat{\mathbf{x}}^-. \quad (2)$$

The unbiased linear update based upon $\hat{\mathbf{x}}^-$ and \mathbf{y} produces the *a posteriori* estimate given by

$$\hat{\mathbf{x}}^+ = \hat{\mathbf{x}}^- + \mathbf{K}(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^-), \quad (3)$$

where \mathbf{K} is some determinist matrix of appropriate dimensions to be determined. The *a posteriori* estimation error expressed as

$$\mathbf{e}^+ = \mathbf{x} - \hat{\mathbf{x}}^+ = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{e}^- - \mathbf{K}\boldsymbol{\eta}. \quad (4)$$

Assuming that the measurement error $\boldsymbol{\eta}$ and the *a priori* estimation error, \mathbf{e}^- are uncorrelated and each are zero mean, we find that the *a posteriori* estimation error covariance is given by the so-called Joseph update:

$$\mathbf{P}^+ = \mathbb{E} \left\{ (\mathbf{x} - \hat{\mathbf{x}}^+) (\mathbf{x} - \hat{\mathbf{x}}^+)^T \right\} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^-(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T, \quad (5)$$

where \mathbf{I} is the identity matrix of appropriate dimension and \mathbf{R} is the measurement error covariance matrix.

Defining the covariance of the measurement residuals, \mathbf{W} , as

$$\mathbf{W} \triangleq \mathbf{H}\mathbf{P}^-\mathbf{H}^T + \mathbf{R}, \quad (6)$$

the updated (*a posteriori*) covariance is equivalently written as

$$\mathbf{P}^+ = \mathbf{P}^- - \mathbf{K}\mathbf{H}\mathbf{P}^- - \mathbf{P}^-\mathbf{H}^T\mathbf{K}^T + \mathbf{K}\mathbf{W}\mathbf{K}^T. \quad (7)$$

Notice that up to now no assumptions have been made as to the choice of \mathbf{K} and the Joseph update equation is valid for all \mathbf{K} , as is Eq. (7). The standard Kalman gain \mathbf{K}_{opt} minimizes the trace of the updated covariance matrix, \mathbf{P}^+

$$\mathbf{K}_{opt} = \mathbf{P}^-\mathbf{H}^T\mathbf{W}^{-1}. \quad (8)$$

Substituting this into Eq. (7) it follows that

$$\mathbf{P}_{opt}^+ = \mathbf{P}^- - \mathbf{K}_{opt} \mathbf{H} \mathbf{P}^- \quad (9)$$

$$= \mathbf{P}^- - \mathbf{K}_{opt} \mathbf{W} \mathbf{K}_{opt}^T \quad (10)$$

$$= \mathbf{P}^- - \mathbf{P}^- \mathbf{H}^T \mathbf{W}^{-1} \mathbf{H} \mathbf{P}^- \quad (11)$$

These equations are only valid for the optimal gain, \mathbf{K}_{opt} .

THE SCHMIDT-KALMAN FILTER UPDATE

Suppose now that we partition \mathbf{x} into n_s “estimated” states, \mathbf{s} , and n_p “consider” parameters, \mathbf{p} as

$$\mathbf{x} \triangleq \begin{bmatrix} \mathbf{s} \\ \mathbf{p} \end{bmatrix}, \quad (12)$$

preserving the fact that $n_x = n_s + n_p$, i.e. the total number of states is unchanged. Let us now partition the relevant matrices accordingly. Thus,

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{ss} & \mathbf{P}_{sp} \\ \mathbf{P}_{ps} & \mathbf{P}_{pp} \end{bmatrix} \quad (13)$$

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_s & \mathbf{H}_p \end{bmatrix} \quad (14)$$

$$\mathbf{K}_{opt} = \begin{bmatrix} \mathbf{K}_{s,opt} \\ \mathbf{K}_{p,opt} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{P}_{sp}^- \mathbf{H}_p^T \\ \mathbf{P}_{ps}^- \mathbf{H}_s^T + \mathbf{P}_{pp}^- \mathbf{H}_p^T \end{bmatrix} \mathbf{W}^{-1}. \quad (15)$$

As in the previous section

$$\begin{aligned} \mathbf{W} &= \mathbf{H} \mathbf{P}^- \mathbf{H}^T + \mathbf{R} \\ &= \mathbf{H}_s \mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{H}_s \mathbf{P}_{sp}^- \mathbf{H}_p^T + \mathbf{H}_p \mathbf{P}_{ps}^- \mathbf{H}_s^T + \mathbf{H}_p \mathbf{P}_{pp}^- \mathbf{H}_p^T + \mathbf{R} \end{aligned} \quad (16)$$

The updated portions of the covariance are

$$\mathbf{P}_{ss}^+ = \mathbf{P}_{ss}^- - \mathbf{K}_s \mathbf{H} \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_s^T + \mathbf{K}_s \mathbf{W} \mathbf{K}_s^T \quad (17)$$

$$\mathbf{P}_{sp}^+ = \mathbf{P}_{sp}^- - \mathbf{K}_s \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_p^T + \mathbf{K}_s \mathbf{W} \mathbf{K}_p^T \quad (18)$$

$$\mathbf{P}_{ps}^+ = \mathbf{P}_{ps}^- - \mathbf{K}_p \mathbf{H} \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_s^T + \mathbf{K}_p \mathbf{W} \mathbf{K}_s^T \quad (19)$$

$$\mathbf{P}_{pp}^+ = \mathbf{P}_{pp}^- - \mathbf{K}_p \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_p^T + \mathbf{K}_p \mathbf{W} \mathbf{K}_p^T \quad (20)$$

These equations are derived from Eq. (7) and are therefore valid for any choice \mathbf{K}_s and \mathbf{K}_p .

We now choose the \mathbf{K}_s and \mathbf{K}_p carefully such that the gain for the states is the optimal value $\mathbf{K}_{s,opt}$ from Eq. (15) and allowing \mathbf{K}_p to be (as yet) unspecified. Thus, substituting

for the value of $\mathbf{K}_{s,opt}$ for the three components in Eqs. (17)–(19), it becomes

$$\mathbf{P}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_{s,opt} \mathbf{W} \mathbf{K}_{s,opt}^T & \mathbf{P}_{sp}^- - \mathbf{K}_{s,opt} \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_{s,opt}^T & \mathbf{P}_{pp}^- - \mathbf{K}_p \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_p^T + \mathbf{K}_p \mathbf{W} \mathbf{K}_p^T \end{bmatrix} \quad (21)$$

This equation is valid for any value of \mathbf{K}_p . Notice that there is no \mathbf{K}_p in the cross-covariance between \mathbf{s} and \mathbf{p} . Therefore, what is remarkable about this equation is that once the optimal $\mathbf{K}_{s,opt}$ is chosen, the cross-covariance between \mathbf{s} and \mathbf{p} is independent of the choice of \mathbf{K}_p . We will take advantage of this property in due course.

We note that the updated (*a posteriori*) state is given by

$$\begin{bmatrix} \mathbf{s}^+ \\ \mathbf{p}^+ \end{bmatrix} = \begin{bmatrix} \mathbf{s}^- \\ \mathbf{p}^- \end{bmatrix} + \begin{bmatrix} \mathbf{K}_s \\ \mathbf{K}_p \end{bmatrix} (\mathbf{y} - \mathbf{H}_s \hat{\mathbf{s}}^- - \mathbf{H}_p \hat{\mathbf{p}}^-), \quad (22)$$

a consider filter is one in which the parameters, \mathbf{p} , are not updated. From Eq. (22) it is observed that this can be achieved by taking $\mathbf{K}_p = \mathbf{O}$. With this in mind, we start again from the most general covariance update given in Eq. (7) and we obtain that the *a posteriori* consider covariance matrix is

$$\mathbf{P}_{con}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_s \mathbf{H} \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix} - \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_s^T + \mathbf{K}_s \mathbf{W} \mathbf{K}_s^T & \mathbf{P}_{sp}^- - \mathbf{K}_s \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_s^T & \mathbf{P}_{pp}^- \end{bmatrix} \quad (23)$$

To find the optimal consider \mathbf{K}_s we take the derivative of the trace of the covariance matrix with respect to \mathbf{K}_s and set it to zero

$$\frac{\partial}{\partial \mathbf{K}_s} (\text{trace} [\mathbf{P}_{con}^+]) = - \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix}^T \mathbf{H}^T - \begin{bmatrix} \mathbf{P}_{ss}^- \\ \mathbf{P}_{ps}^- \end{bmatrix}^T \mathbf{H}^T + \mathbf{K}_{s,opt} \mathbf{W}^T + \mathbf{K}_{s,opt} \mathbf{W} = \mathbf{O}. \quad (24)$$

The result is the same as for the globally optimal Kalman filter, so that the optimal consider Kalman filter, also called the Schmidt-Kalman filter can be conveniently calculated from the optimal filter by zeroing out the rows corresponding to the consider states.

$$\mathbf{K}_{s,opt} = (\mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{P}_{sp}^- \mathbf{H}_p^T) \mathbf{W}^{-1} \quad (25)$$

So that the consider gain is

$$\mathbf{K}_{con} = \begin{bmatrix} \mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{P}_{sp}^- \mathbf{H}_p^T \\ \mathbf{O} \end{bmatrix} \mathbf{W}^{-1} = \begin{bmatrix} \mathbf{K}_{s,opt} \\ \mathbf{O} \end{bmatrix} \quad (26)$$

The state update is

$$\mathbf{s}^+ = \mathbf{s}^- + \mathbf{K}_s (\mathbf{y} - \mathbf{H}_s \hat{\mathbf{s}}^- - \mathbf{H}_p \hat{\mathbf{p}}^-) \quad (27)$$

$$\mathbf{p}^+ = \mathbf{p}^-. \quad (28)$$

It must be stated emphatically that the full updated covariance matrix (the *a posteriori* matrix) must be calculated by means of the Joseph update. We cannot use Eq. (11), for that equation was valid only for the optimal gain; we have chosen a sub-optimal gain or a non-globally optimal gain (*i.e.* the optimal gain corresponding to $\mathbf{K}_p = \mathbf{O}$). Substituting $\mathbf{K}_{s,opt}$ into Eq. (23) we obtain

$$\mathbf{P}_{con}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_{s,opt} \mathbf{W} \mathbf{K}_{s,opt}^T & \mathbf{P}_{sp}^- - \mathbf{K}_{s,opt} \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_{s,opt}^T & \mathbf{P}_{pp}^- \end{bmatrix} \quad (29)$$

Several comments are in order:

1. When using the Schmidt-Kalman filter, the *a priori* and *a posteriori* covariance of the parameters (\mathbf{P}_{pp}) are the same.
2. The *a posteriori* covariance matrix of the states and the cross-covariance between the states and the parameters are the same regardless of whether one uses the Schmidt-Kalman filter or the optimal Kalman filter.

It follows that the *a posteriori* consider covariance can be obtained from the *a posteriori* optimal covariance by replacing \mathbf{P}_{pp}^+ with \mathbf{P}_{pp}^- .

Finally, item (2) above *does not* mean that \mathbf{P}_{ss} from the optimal Kalman filter will be the same for all time as for the Schmidt-Kalman filter. Rather, it means that given the same *a priori* covariance the two algorithms will produce the same *a posteriori* covariance for \mathbf{P}_{ss} and \mathbf{P}_{sp} . However, since \mathbf{P}_{pp} is different between the two algorithms, after the very first update, the subsequent *a priori* covariances will be different henceforth, producing different results. We will expand to this point in due time.

THE UDU UPDATE

A common strategy to implement covariances in real-time flight software is to use the UDU formulation [11]. Thus, the covariance matrix is decomposed as

$$\mathbf{P} = \mathbf{U} \mathbf{D} \mathbf{U}^T \quad (30)$$

where \mathbf{U} is an upper triangular matrix with ones on the main diagonal, and \mathbf{D} is a diagonal matrix. In order to efficiently update \mathbf{U} and \mathbf{D} with measurements, the so-called ‘rank-one’ update proposed by Agee and Turner is sometimes used [12]. The rank-one (covariance) update requires that measurements are processed one at a time and has the form

$$\mathbf{P}^+ = \mathbf{P}^- + c \mathbf{a} \mathbf{a}^T \quad (31)$$

where c is a scalar and \mathbf{a} is a vector. If we let $\mathbf{P}^+ = \mathbf{U}^+ \mathbf{D}^+ (\mathbf{U}^+)^T$ and $\mathbf{P}^- = \mathbf{U}^- \mathbf{D}^- (\mathbf{U}^-)^T$, then the update must have the form

$$\mathbf{U}^+ \mathbf{D}^+ (\mathbf{U}^+)^T = \mathbf{U}^- \mathbf{D}^- (\mathbf{U}^-)^T + c \mathbf{a} \mathbf{a}^T \quad (32)$$

The algorithm for the rank-one update is backwards-recursive for \mathbf{U}^+ and \mathbf{D}^+ , starting with the (n, n) element (D_{nn}) and ending with the $(1, 1)$ element (D_{11}) and it is seen in Table 1.

For $j = n, n - 1, \dots, 3, 2$ set $c_n = c$ and recursively calculate:
$D_{jj}^+ = D_{jj}^- + c_j a_j^2$
$a_k := a_k - a_j U_{kj}^- \quad k = 1, 2, \dots, j - 1$
$U_{kj}^+ = U_{kj}^- + c_k a_j a_k / D_{jj}^+ \quad k = 1, 2, \dots, j - 1$
$c_{j-1} = c_j D_{jj}^+ / D_{jj}^-$
and finally compute
$D_{11}^+ = D_{11}^- + c_1 a_1^2$

Table 1. Backwards-Recursive Rank-One Update

For scalar measurements \mathbf{H} and \mathbf{K} become vectors and \mathbf{W} is a scalar. We recall that from Eq. (11) that the *optimal update* of the covariance matrix is

$$\mathbf{P}_{opt}^+ = \mathbf{P}^- - \mathbf{P}^- \mathbf{H}^T \mathbf{W}^{-1} \mathbf{H} \mathbf{P}^-$$

Thus,

$$\begin{aligned} \mathbf{U}^+ \mathbf{D}^+ (\mathbf{U}^+)^T &= \mathbf{U}^- \mathbf{D}^- (\mathbf{U}^-)^T - \mathbf{U}^- \mathbf{D}^- (\mathbf{U}^-)^T \mathbf{H}^T \mathbf{W}^{-1} \mathbf{H} \mathbf{U}^- \mathbf{D}^- (\mathbf{U}^-)^T \\ &= \mathbf{U}^- \left[\mathbf{D}^- - \frac{1}{W} \mathbf{D}^- (\mathbf{U}^-)^T \mathbf{H}^T \mathbf{H} \mathbf{U}^- \mathbf{D}^- \right] (\mathbf{U}^-)^T \end{aligned}$$

Defining

$$\mathbf{a} \triangleq \mathbf{D}^- (\mathbf{U}^-)^T \mathbf{H}^T \tag{33}$$

$$c \triangleq -\frac{1}{W} \tag{34}$$

$\mathbf{U}^+ \mathbf{D}^+ (\mathbf{U}^+)^T$ becomes

$$\mathbf{U}^+ \mathbf{D}^+ (\mathbf{U}^+)^T = \mathbf{U}^- [\mathbf{I} \mathbf{D}^- \mathbf{I} + c \mathbf{a} \mathbf{a}^T] (\mathbf{U}^-)^T,$$

where \mathbf{I} is the identity matrix (which is upper triangular with ones on the diagonal). We now proceed with the rank-one update of the term in the square bracket, defining an intermediate step to the solution as

$$\tilde{\mathbf{U}} \tilde{\mathbf{D}} \tilde{\mathbf{U}}^T \triangleq \mathbf{I} \mathbf{D}^- \mathbf{I} + c \mathbf{a} \mathbf{a}^T$$

The updated matrices are given by

$$\mathbf{U}^+ = \mathbf{U}^- \tilde{\mathbf{U}} \quad \text{and} \quad \mathbf{D}^+ = \tilde{\mathbf{D}} \tag{35}$$

Since c is negative, this measurement update involves the potential loss of precision due to subtraction of two positive numbers which are close to one another, a modification of the rank-one update is due to Carlson [13], resulting in a forward-recursive algorithm, one which is numerically stable. This modified algorithm is only valid for the optimal Kalman update, \mathbf{K}_{opt} . The updated \mathbf{U} and \mathbf{D} (\mathbf{U}^+ and \mathbf{D}^+ , respectively) and the optimal Kalman gain matrix, \mathbf{K}_{opt} , are produced by this algorithm, which is detailed in Table 2.

$$\begin{aligned}
& \text{Given } \mathbf{H}, \bar{\mathbf{U}}, \bar{\mathbf{D}}, \mathbf{R} = R \\
& \mathbf{f} = \bar{\mathbf{U}}^T \mathbf{H}^T \\
& \mathbf{v} = \bar{\mathbf{D}} \mathbf{f} \\
& \bar{\mathbf{K}}_1 = [v_1 \quad 0 \quad \cdots \quad \mathbf{0}]^T \\
& \alpha_1 = R + v_1 f_1 \\
& d_1 = d_{11} = \left(\frac{R}{\alpha_1}\right) \bar{d}_{11} \\
& \text{For } j = 2, 3, \dots, n, \text{ recursively calculate:} \\
& \alpha_j = \alpha_{j-1} + v_j f_j \\
& d_j = d_{jj} = \left(\frac{\alpha_{j-1}}{\alpha_j}\right) \bar{d}_{jj} \\
& \lambda_j = -(f_j / \alpha_{j-1}) \\
& \mathbf{U}_j = \bar{\mathbf{U}}_j + \lambda_j \bar{\mathbf{K}}_{j-1} \\
& \bar{\mathbf{K}}_j = \bar{\mathbf{K}}_{j-1} + v_j \bar{\mathbf{U}}_j \\
& \text{and finally compute} \\
& \mathbf{K}_{opt} = \bar{\mathbf{K}}_n / \alpha
\end{aligned}$$

Table 2. The Forward-Recursive *Modified* Rank-One Update

The updated state is computed as before, *i.e.*

$$\mathbf{x}^+ = \mathbf{x}^- + \mathbf{K}_{opt} (\mathbf{y} - \mathbf{H}\mathbf{x}^-) \quad (36)$$

THE SCHMIDT-KALMAN FILTER UDU UPDATE

Recall that the Schmidt-Kalman update to the covariance matrix due to the measurements was stated in Eq. (29) as

$$\mathbf{P}_{con}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_{s,opt} \mathbf{W} \mathbf{K}_{s,opt}^T & \mathbf{P}_{sp}^- - \mathbf{K}_{s,opt} \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_{s,opt}^T & \mathbf{P}_{pp}^- \end{bmatrix}$$

Unfortunately, the structure of the matrix in Eq. (37) does not allow for a rank-one update as we developed earlier. We stated earlier that to obtain the updated consider covariance we can begin with the updated optimal covariance and replace \mathbf{P}_{pp}^+ with \mathbf{P}_{pp}^- . Now, from Eqs. (10) and (11), we have

$$\mathbf{P}_{opt}^+ = \mathbf{P}^- - \mathbf{K}_{opt} \mathbf{W} \mathbf{K}_{opt}^T \quad (37)$$

From these relations, we see that the optimal *a posteriori* covariance is equal to the *a priori* covariance minus a delta covariance. We can, therefore, harness this idea to help us along. With the optimal value for \mathbf{K}_p , we obtain

$$\mathbf{P}_{opt}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_{s,opt} \mathbf{W} \mathbf{K}_{s,opt}^T & \mathbf{P}_{sp}^- - \mathbf{K}_{s,opt} \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_{s,opt}^T & \mathbf{P}_{pp}^- - \mathbf{K}_{p,opt} \mathbf{W} \mathbf{K}_{p,opt}^T \end{bmatrix} \quad (38)$$

This, of course can be written as

$$\mathbf{P}_{opt}^+ = \begin{bmatrix} \mathbf{P}_{ss}^- - \mathbf{K}_{s,opt} \mathbf{W} \mathbf{K}_{s,opt}^T & \mathbf{P}_{sp}^- - \mathbf{K}_{s,opt} \mathbf{H} \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix} \\ \mathbf{P}_{ps}^- - \begin{bmatrix} \mathbf{P}_{sp}^- \\ \mathbf{P}_{pp}^- \end{bmatrix}^T \mathbf{H}^T \mathbf{K}_{s,opt}^T & \mathbf{P}_{pp}^- \end{bmatrix} - \begin{bmatrix} \mathbf{O}_{n_s \times n_s} & \mathbf{O}_{n_s \times n_p} \\ \mathbf{O}_{n_p \times n_s} & \mathbf{K}_{p,opt} \mathbf{W} \mathbf{K}_{p,opt}^T \end{bmatrix} \quad (39)$$

Thus, we can rewrite Eq. (39) as

$$\mathbf{P}_{opt}^+ = \mathbf{P}_{con}^+ - \begin{bmatrix} \mathbf{O}_{n_s \times n_s} & \mathbf{O}_{n_s \times n_p} \\ \mathbf{O}_{n_p \times n_s} & \mathbf{K}_{p,opt} \mathbf{W} \mathbf{K}_{p,opt}^T \end{bmatrix} \quad (40)$$

Recalling from Eq.(15) that the optimal gain is partitioned as

$$\mathbf{K}_{opt} \triangleq \begin{bmatrix} \mathbf{K}_{s,opt} \\ \mathbf{K}_{p,opt} \end{bmatrix}.$$

Defining \mathbf{S} as

$$\mathbf{S} \triangleq \begin{bmatrix} \mathbf{O}_{n_s \times n_s} & \mathbf{O}_{n_s \times n_p} \\ \mathbf{O}_{n_p \times n_s} & \mathbf{I}_{n_p \times n_p} \end{bmatrix} \quad (41)$$

Eq. (40) becomes

$$\mathbf{P}_{opt}^+ = \mathbf{P}_{con}^+ - \mathbf{W} (\mathbf{S} \mathbf{K}_{opt}) (\mathbf{S} \mathbf{K}_{opt})^T \quad (42)$$

We note that \mathbf{S} is an $n_x \times n_x$ matrix (recalling that $n_x = n_s + n_p$) and \mathbf{K}_{opt} is an $n_x \times 1$ vector, because we are processing the measurements as scalars. Therefore $\mathbf{S} \mathbf{K}_{opt}$ is an $n_x \times 1$ vector. Therefore,

$$\mathbf{P}_{con}^+ = \mathbf{P}_{opt}^+ + \mathbf{W} (\mathbf{S} \mathbf{K}_{opt}) (\mathbf{S} \mathbf{K}_{opt})^T \quad (43)$$

This has the same form as the rank-one update as found in Eq. (31), *i.e.* $\mathbf{P}^+ = \mathbf{P}^- + c \mathbf{a} \mathbf{a}^T$. While earlier the scalar was negative for the optimal update creating potential numerical issues, now

$$\mathbf{a} = \mathbf{S} \mathbf{K}_{opt} \quad (44)$$

$$c = \mathbf{W}. \quad (45)$$

With this in mind, we can use the (un-modified) rank-one update which is a backward-recursive update. If, for example, all the consider parameters are in the top part of the state-space, we can effectively reduce the computations because the second update will not modify the columns of \mathbf{U} and \mathbf{D} corresponding to the estimated states.

Finally, the states (and the consider parameters) are updated as

$$\mathbf{s}^+ = \mathbf{s}^- + \mathbf{K}_s (\mathbf{y} - \mathbf{H}_s \hat{\mathbf{s}}^- - \mathbf{H}_p \hat{\mathbf{p}}^-) \quad (46)$$

$$\mathbf{p}^+ = \mathbf{p}^- \quad (47)$$

Of course, this is identical to what was obtained in Section 3.

Therefore, the procedure is as follows: first perform a complete rank-one measurement update with the optimal Kalman Gain (\mathbf{K}_{opt}) according to the *modified* rank-one update (as in Table 2) – on the full covariance matrix. Second, perform another rank-one update with $\mathbf{a} = \mathbf{S}\mathbf{K}_{opt}$ and $c = W$, according to the (un-modified) rank-one update (as in Table 1). If the consider parameters are always the same across flight we place them conveniently to reduce computations.

OPTIMAL CONSIDER FILTER

Through a logic construction, Bierman [11] states that the recursive Schmidt-Kalman filter is non-optimal because processing data as multiple batches provides a less accurate result than processing it all at once. We now demonstrate that statement and we propose a solution to overcome the nonoptimality of the Schmidt-Kalman filter.

Suppose we divide the measurement vector into two uncorrelated batches $\mathbf{y}^T = [\mathbf{y}_a^T \quad \mathbf{y}_b^T]$. It is well-known that processing the two batches together or sequentially using the Kalman filter equations produces the same result [14], i.e. defining

$$\mathbf{P}_{opt}^+ = (\mathbf{I} - \mathbf{K}_{opt}\mathbf{H}) \mathbf{P}^-, \quad \mathbf{K}_{opt} = \mathbf{P}^- \mathbf{H}^T (\mathbf{H}\mathbf{P}^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (48)$$

$$\mathbf{P}_{opt}^a = (\mathbf{I} - \mathbf{K}_{opt}^a \mathbf{H}_a) \mathbf{P}^-, \quad \mathbf{K}_{opt}^a = \mathbf{P}^- \mathbf{H}_a^T (\mathbf{H}_a \mathbf{P}^- \mathbf{H}_a^T + \mathbf{R}_a)^{-1} \quad (49)$$

$$\mathbf{P}_{opt}^b = (\mathbf{I} - \mathbf{K}_{opt}^b \mathbf{H}_b) \mathbf{P}_{opt}^a, \quad \mathbf{K}_{opt}^b = \mathbf{P}_{opt}^a \mathbf{H}_b^T (\mathbf{H}_b \mathbf{P}_{opt}^a \mathbf{H}_b^T + \mathbf{R}_b)^{-1} \quad (50)$$

then $\mathbf{P}_{opt}^+ = \mathbf{P}_{opt}^b$. Processing all the measurements at once using the consider gain

$$\mathbf{K}_{con} = \begin{bmatrix} \mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{P}_{sp}^- \mathbf{H}_p^T \\ \mathbf{O} \end{bmatrix} (\mathbf{H}_s \mathbf{P}_{ss}^- \mathbf{H}_s^T + \mathbf{H}_s \mathbf{P}_{sp}^- \mathbf{H}_p^T + \mathbf{H}_p \mathbf{P}_{ps}^- \mathbf{H}_s^T + \mathbf{H}_p \mathbf{P}_{pp}^- \mathbf{H}_p^T + \mathbf{R})^{-1}$$

we have that $\hat{\mathbf{s}}^+$, \mathbf{P}_{ss}^+ , \mathbf{P}_{ps}^+ have the same values for both the Schmidt-Kalman filter and the regular Kalman filter. If we process the first batch of measurements with the consider gain we have that $\mathbf{P}_{ss,opt}^a = \mathbf{P}_{ss,con}^a$, $\mathbf{P}_{ps,opt}^a = \mathbf{P}_{ps,con}^a$, but $\mathbf{P}_{pp,opt}^a \neq \mathbf{P}_{pp,con}^a = \mathbf{P}_{pp}^-$. Therefore

$$\begin{aligned} \mathbf{H}_s \mathbf{P}_{ss,con}^a \mathbf{H}_s^T + \mathbf{H}_s \mathbf{P}_{sp,con}^a \mathbf{H}_p^T + \mathbf{H}_p \mathbf{P}_{ps,con}^a \mathbf{H}_s^T + \mathbf{H}_p \mathbf{P}_{pp,con}^a \mathbf{H}_p^T + \mathbf{R}_b &\geq \\ \mathbf{H}_s \mathbf{P}_{ss,opt}^a \mathbf{H}_s^T + \mathbf{H}_s \mathbf{P}_{sp,opt}^a \mathbf{H}_p^T + \mathbf{H}_p \mathbf{P}_{ps,opt}^a \mathbf{H}_s^T + \mathbf{H}_p \mathbf{P}_{pp,opt}^a \mathbf{H}_p^T + \mathbf{R}_b & \end{aligned} \quad (51)$$

hence the rows of \mathbf{K}_{con}^b corresponding to the estimated states will be “smaller” than the corresponding rows of \mathbf{K}_{opt}^b , resulting in a smaller update.

Previously we derived the “optimal” consider filter by minimizing the trace of the posterior covariance. The optimality holds only if all measurements are processed at once and for a given prior covariance. Processing all measurements at once is more accurate than processing them in two batches, which in turn is more accurate than processing them in three batches, and so on. While this result might not seem of great consequences (usually only a few measurements are available at each given time), it has one larger implication: the fact that processing all measurements together or in batches produces equal results it is the key to optimality of the recursive Kalman filter. In the Kalman filter only the current measurements are incorporated, this measurement update is followed by propagation to the time of the next measurements. This procedure is equivalent to processing all the current and past measurements at once. In standard aerospace applications, measurements can be incorporated in hundreds of batches, one for each time the filter update phase is invoked. It is therefore essential to use a recursive estimator.

In previous sections it was noticed that we can incorporate the measurements using the regular Kalman filter update of the states, while leaving the parameters unchanged. The full covariance can be updated first using the entire Kalman gain and we then substitute the optimal \mathbf{P}_{pp}^+ with the prior \mathbf{P}_{pp}^- . Adopting this strategy we obtain exactly the same result as the Schmidt-Kalman filter. More importantly we could update the covariance and the states in many batches, and only after the last batch is processed we would substitute \mathbf{P}_{pp}^+ with \mathbf{P}_{pp}^- .

An optimal recursive Schmidt-Kalman filter that gives identical results to a batch estimator is obtained by implementing a regular Kalman filter in which the entire covariance is updated and propagated at all times, but only the estimated states are updated using the measurements, while the consider parameters remain unchanged across measurements. As discussed earlier the optimal consider filter and the Kalman filter share the same estimate of the states, their covariance, and their cross-covariance with the consider parameters. The covariance of the parameters of this optimal consider filter does not represent their actual uncertainty, instead it represents the uncertainty the parameters would have if they were optimally estimated. If it is desired to know what the uncertainty of the parameters actually is, another, smaller, covariance of only the parameters needs to be carried. This “true” covariance is never actually used by the filter nor changes in the measurement update phase.

NUMERICAL RESULTS

In this section we present a very simple numerical example to demonstrate the theory developed in this paper. A single state s and parameter p for the state vector $\mathbf{x}^T = [s \ p]$. The state is measured directly, while the parameter is a systematic measurement error

$$y_k = s_k + p_k + \eta_k = \begin{bmatrix} 1 & 1 \end{bmatrix} \mathbf{x}_k + \eta_k = \mathbf{H}\mathbf{x}_k + \eta_k \quad (52)$$

where η_k is a zero mean, white sequence with variance $R_k = 1$ and uncorrelated from any other error source. The true state is a random walk while the parameter is a first order

Markov process

$$\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \boldsymbol{\nu}_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{-\Delta t_k/\tau} \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} \nu_k \\ \mu_k \end{bmatrix} \quad (53)$$

where ν_k and μ_k are zero mean, uncorrelated white sequences with variances given by $Q_k = 1$ and $(1 - e^{-2\Delta t_k/\tau})P_{pp,ss}$, respectively. The steady-state value of the Markov process variance is chosen as $P_{pp,ss} = 1$ and the time constant τ is such that $e^{-200/\tau} = 0.5$. An initial unbiased estimate $\hat{\mathbf{x}}_0^T = [0 \ 0]$ of the true state vector is available and has associated estimation error covariance given by

$$\mathbf{P}_0 = \begin{bmatrix} 10 & 3 \\ 3 & 1 \end{bmatrix}. \quad (54)$$

A first measurement y_0 is assumed to be available at time $t_0 = 0$ and a second measurement y_1 becomes available at time $t_1 = 100$. A standard Kalman filter implementation performs a first update and produces

$$\mathbf{P}_{KF}^+(t_0) = \begin{bmatrix} 0.6111 & 0.1111 \\ 0.1111 & 0.1111 \end{bmatrix}. \quad (55)$$

The next step is a time update after which

$$\mathbf{P}_{KF}^-(t_1) = \begin{bmatrix} 1.6111 & 0.0786 \\ 0.0786 & 0.5556 \end{bmatrix}. \quad (56)$$

Finally a second update is performed that produces

$$\mathbf{P}_{KF}^+(t_1) = \begin{bmatrix} 0.7522 & -0.2438 \\ -0.2438 & 0.4346 \end{bmatrix}. \quad (57)$$

From the discussion above and since the initial uncertainty of the parameter coincides with its steady-state value we know that an optimal consider filter provides an estimate with error covariance

$$\mathbf{P}_{OCKF}^+(t_1) = \begin{bmatrix} 0.7522 & -0.2438 \\ -0.2438 & 1 \end{bmatrix}, \quad (58)$$

however applying the Schmidt-Kalman filter equations recursively we obtain after the first update

$$\mathbf{P}_{CKF}^+(t_0) = \begin{bmatrix} 0.6111 & 0.1111 \\ 0.1111 & 1 \end{bmatrix}, \quad (59)$$

after the propagation

$$\mathbf{P}_{CKF}^-(t_1) = \begin{bmatrix} 1.6111 & 0.0786 \\ 0.0786 & 1 \end{bmatrix}, \quad (60)$$

and finally after the second update is

$$\mathbf{P}_{CKF}^+(t_1) = \begin{bmatrix} 0.8535 & -0.4051 \\ -0.4051 & 1 \end{bmatrix}. \quad (61)$$

The optimal consider filter algorithm proposed in this paper carries the covariance of the Kalman filter with the understanding the the portion corresponding to the covariance of the consider parameters is fictitious. At any time it is desired to know the actual covariance it can be built from the Kalman filter covariance and an externally carried covariance of the parameters. For the simple example in this section the parameter’s covariance stays constant at its steady state value. To demonstrate the validity of this method a Monte Carlo simulation is performed. True initial states are obtained sampling from a Gaussian distribution with mean $\hat{\mathbf{x}}_0$ and covariance \mathbf{P}_0 . Each Monte Carlo run also samples different values for the process noise and measurement noise, with zero mean and covariance as specified above. The propagation phase is given by

$$\hat{\mathbf{x}}_{k+1} = \Phi_k \hat{\mathbf{x}}_k. \quad (62)$$

Figures 1 to 3 show the results of 100 Monte Carlo runs for each algorithm. The figures convey that all the algorithms perform correctly since their predicted estimation error is consistent with its actual value. Figure 4 shows a comparison of the three algorithms. All the covariances shown in the plots are 3σ values.

CONCLUSIONS

This paper analyzes recursive implementations of the consider filter. A recursive implementation is essential for onboard estimation systems. A common strategy for aerospace navigation is to carry the filter’s estimation error covariance utilizing the UDU formulation. This paper introduces a formulation of the UDU consider filter. The non-optimality of the recursive consider filter is discussed. An algorithm is proposed to overcome this limitation and produce optimal consider filter results. A numerical example is shown to demonstrate the validity of the proposed approach.

REFERENCES

- [1] S. F. Schmidt, “Application of State-Space Methods to Navigation Problems,” *Advances in Control Systems*, Vol. 3, 1966, pp. 293–340.
- [2] D. Woodbury and J. Junkins, “On the Consider Kalman Filter,” *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, August 2010.
- [3] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Vol. 64 of *Mathematics in Sciences and Engineering*. New York, New York 10003: Academic Press, 1970.
- [4] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Vol. 128 of *Mathematics in Sciences and Engineering*. Academic Press, 1978.
- [5] B. D. Tapley, B. E. Schutz, and G. H. Born, *Statistical Orbit Determination*. Elsevier Academic Press, 2004.
- [6] D. Woodbury, M. Majji, and J. Junkins, “Considering Measurement Model Parameter Errors in Static and Dynamic Systems,” *Proceedings of the George H. Born Symposium*, May 2010.
- [7] R. Zanetti, *Advanced Navigation Algorithms for Precision Landing*. PhD thesis, The University of Texas at Austin, Austin, Texas, December 2007.
- [8] R. Zanetti and R. H. Bishop, “Kalman Filters with Uncompensated Biases,” *Journal of Guidance, Control, and Dynamics*, Vol. 35, January–February 2012, pp. 327–330.
- [9] R. Zanetti and R. H. Bishop, “Entry Navigation Dead-Reckoning Error Analysis: Theoretical Foundations of the Discrete-Time Case,” *Proceedings of the AAS/AIAA Astrodynamics Specialist Conference held August 19–23, 2007, Mackinac Island, Michigan*, Vol. 129 of *Advances in the Astronautical Sciences*, 2007, pp. 979–994. AAS 07-311.

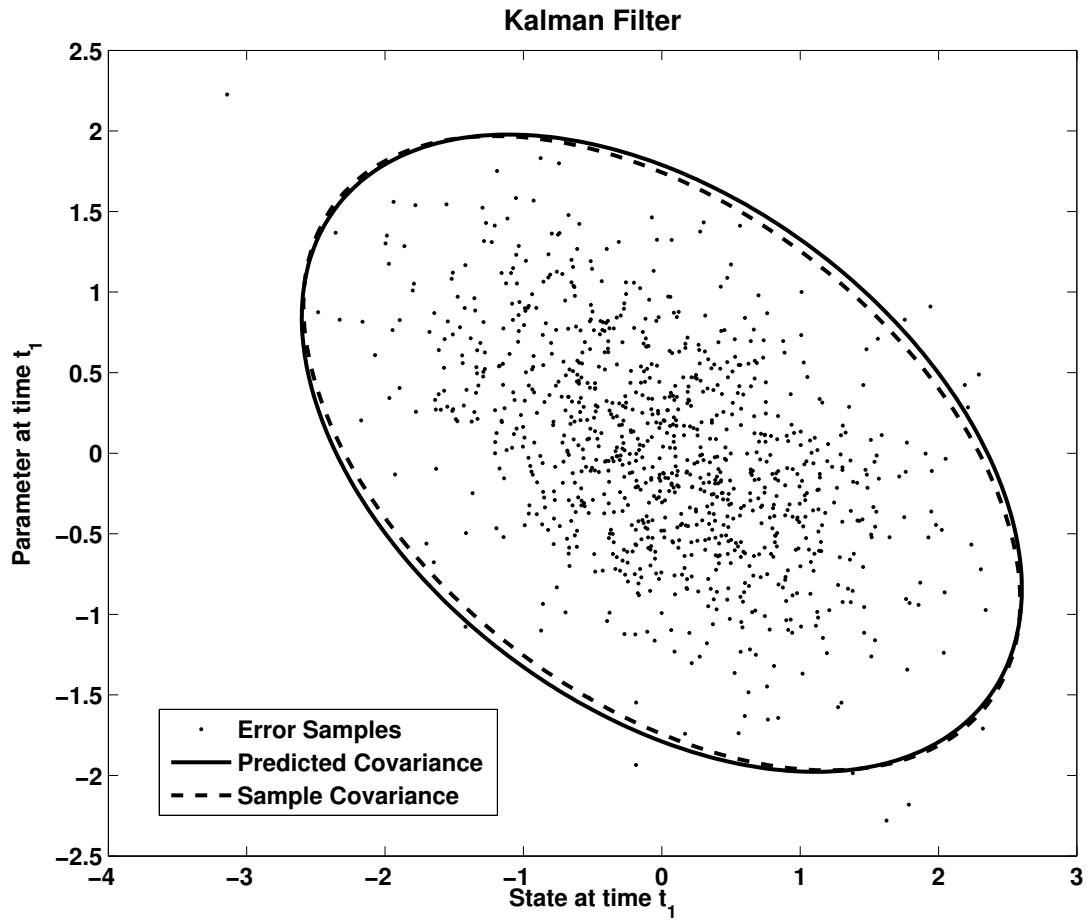


Figure 1. Kalman Filter Monte Carlo Results

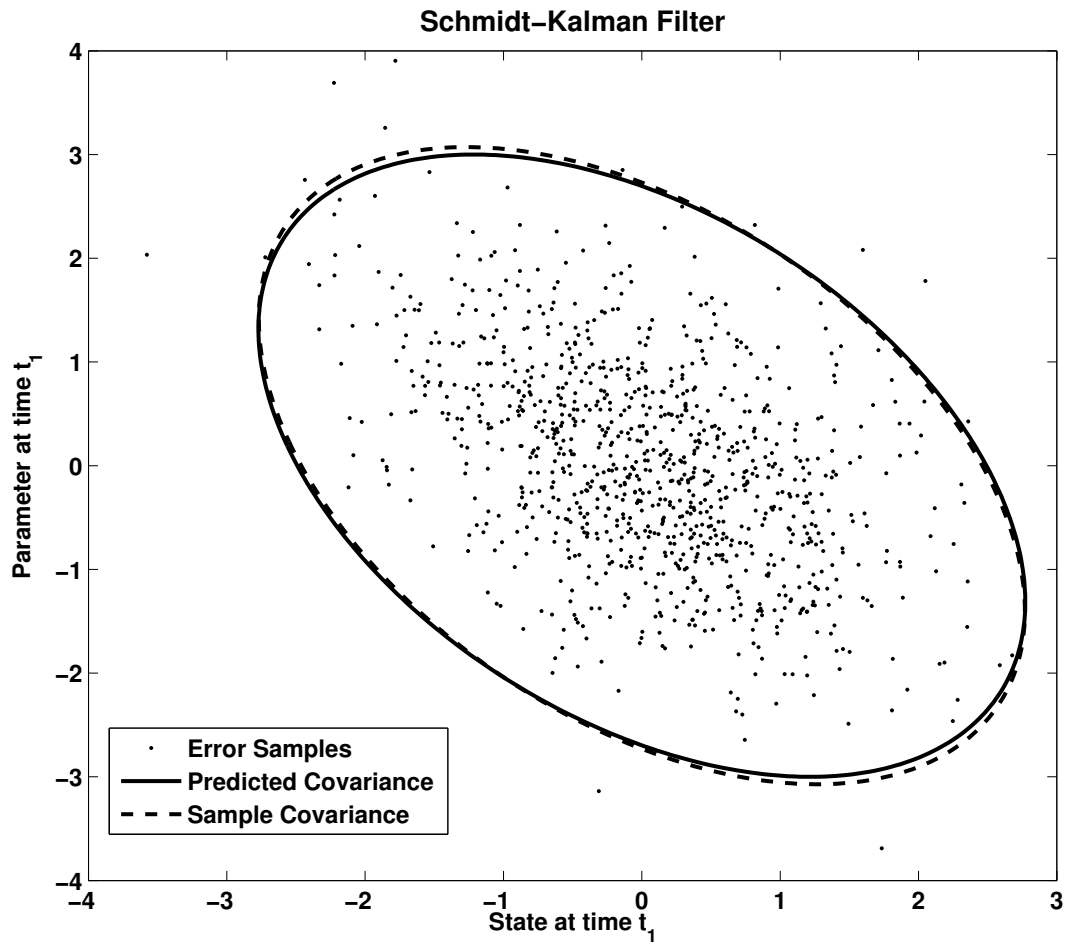


Figure 2. Schmidt-Kalman Filter Monte Carlo Results

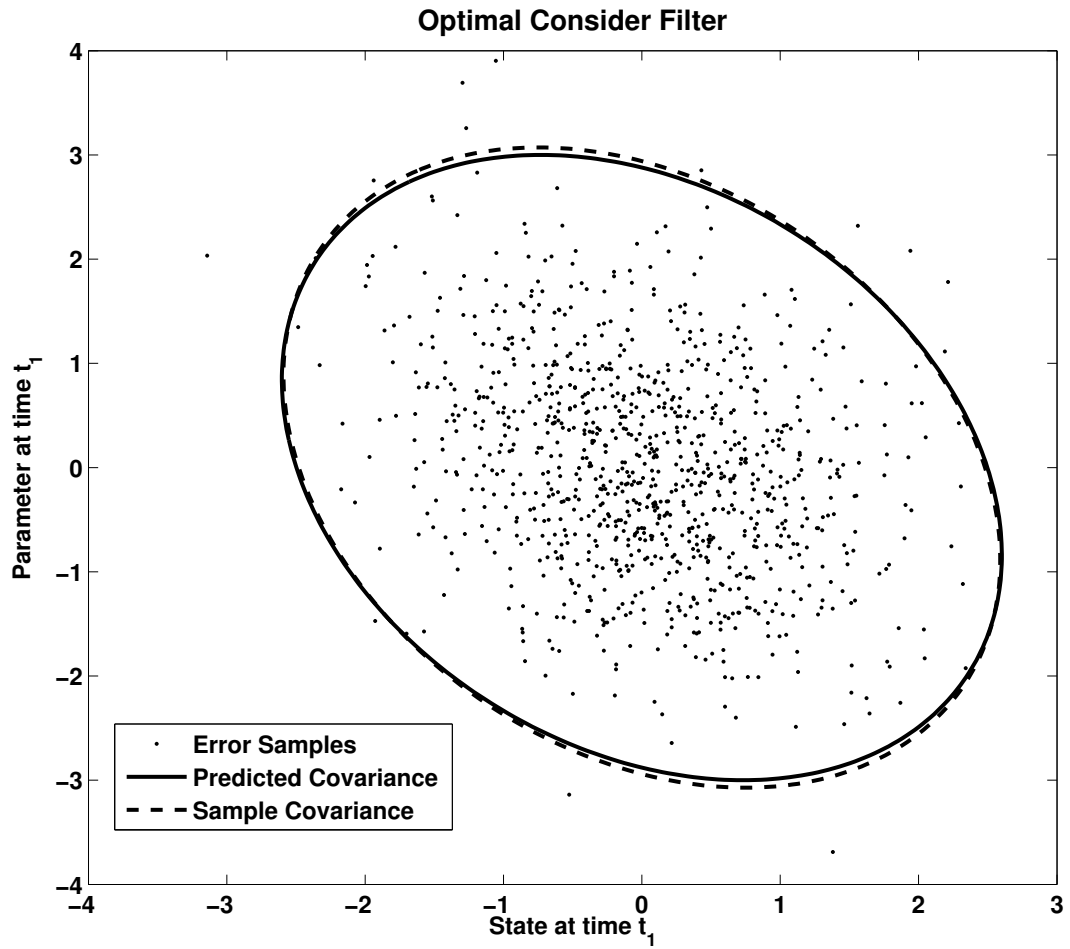


Figure 3. Optimal Consider Filter Monte Carlo Results

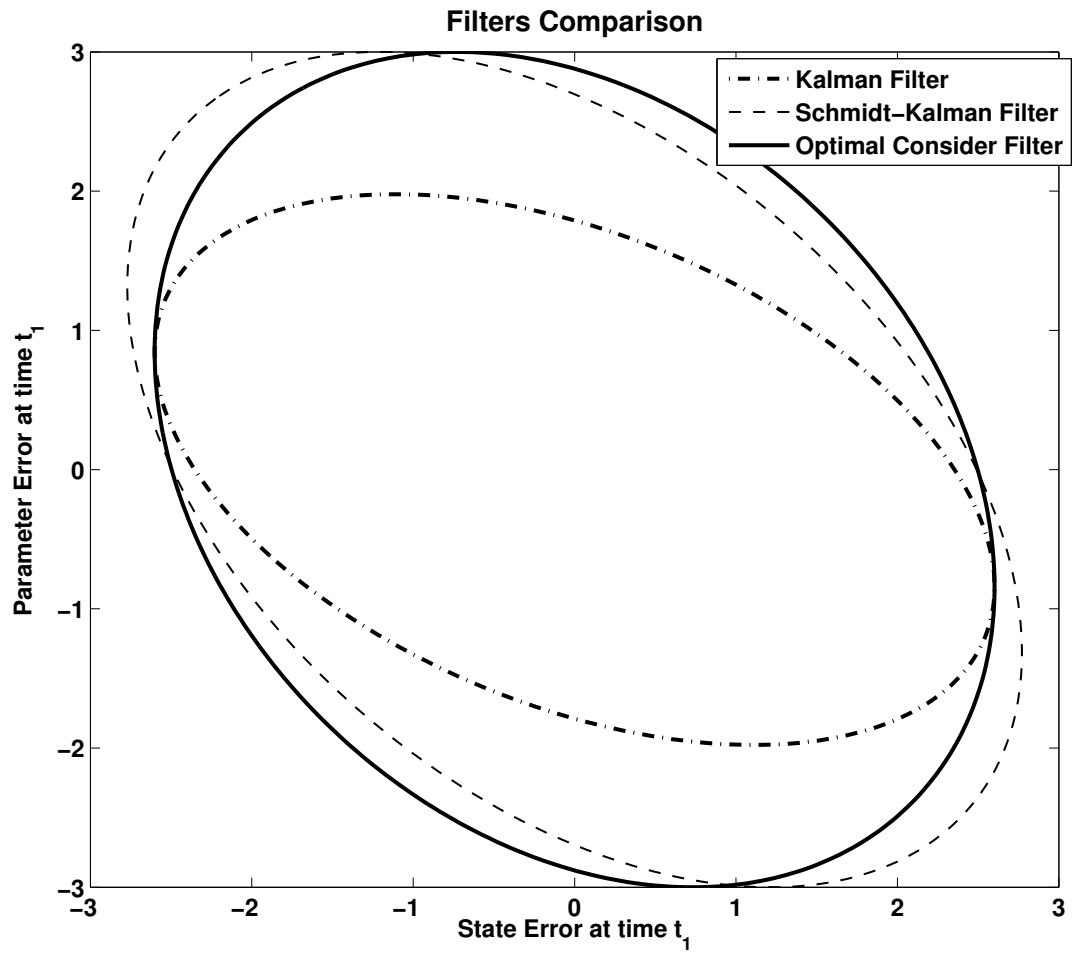


Figure 4. Comparison of Filters Estimation Errors

- [10] M. E. Hough, "Orbit Determination with Improved Covariance Fidelity, Including Sensor Measurement Biases," *Journal of Guidance, Control, and Dynamics*, Vol. 34, May–June 2011, pp. 903–911.
- [11] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*. New York: Dover Publications, 2006.
- [12] W. Agee and R. Turner, "Triangular Decomposition of a Positive Definite Matrix Plus a Symmetric Dyad with Application to Kalman Filtering," *White Sands Missile Range Tech. Rep. No. 38*, 1972.
- [13] N. Carlson, "Fast Triangular Factorization of the Square Root Filter," *AIAA Journal*, Vol. 11, No. 9, 1973, pp. 1259–1265.
- [14] R. G. Brown and P. Y. Hwang, *Introduction To Random Signals And Applied Kalman Filtering*. John Wiley and Sons, third ed., 1997.