Real-Time Implementation of the Optimal Predictor and Optimal Filter: Accuracy Versus Latency

SYED ASEEM UL ISLAM, ANKIT GOEL, and DENNIS S. BERNSTEIN

Summary

The Kalman filter uses noisy sensor measurements to estimate the states of a linear system that is subject to disturbances. An elegant implementation of this algorithm is the optimal two-step filter (OTSF), which involves a forecast step based on the dynamics and a data assimilation step based on sensor data. However, what happens if the order of the steps between state estimates is reversed? This variation of the OTSF, which is the optimal two-step predictor (OTSP), provides slightly earlier state estimates than the OTSF based on slightly older data. The goal of this article is to compare the accuracy of the OTSF and the OTSP to assess whether or not the delayed estimates of the OTSF are more accurate than the faster estimates of the OTSP.

Although the Kalman filter is often presented within a continuous-time context [1]–[4], the original derivation was provided in discrete time [5]. In practice, controllers and observers are invariably implemented digitally; thus discrete-time algorithms deserve special attention. This article focuses on the real-time implementation of the discrete-time Kalman filter and its relation to the discrete-time Kalman predictor. These algorithms differ in subtle ways, as highlighted in this article. Equations for the discrete-time Kalman filter and predictor are given in [6]. Unfortunately, the covariance matrix becomes indefinite after a few steps, which shows that those equations are erroneous. The present article corrects and extends the results from [6].

This derivation is based on necessary conditions and thus does not prove that the optimal one-step predictor (OOSP) is the globally optimal one. However, this derivation is succinct and uses minimal mathematics, making it efficient for classroom presentation. As shown in [2, p. 46], in the absence of Gaussian processes, this derivation yields the optimal linear predictor. Next, the OOSP is recast as an optimal two-step predictor (OTSP). The steps of the OTSP are then reversed to obtain the optimal two-step filter (OTSF), as described in "Summary." Finally, the steps of the OTSF are combined to obtain the optimal one-step filter (OOSF).

The OTSP and OTSF consist of assimilation and forecast updates. The OTSP begins with an assimilation update followed by a forecast update, whereas the OTSF begins with a forecast update followed by an assimilation update. An additional distinction between these two algorithms concerns their real-time implementation. Within the context of constraints on data collection and computation, the OTSP and OTSF use different data to produce state estimates that are available at different times. Overall, the OTSF uses more recent data but produces state estimates at a later time. It is thus expected that the OTSF estimates are more accurate than those of the OTSP, with latency the price paid for the enhanced accuracy of the filter.

To investigate the validity of the expected accuracy/latency tradeoff, the accuracy of the OTSP and OTSF is compared by means of a numerical example. Somewhat surprisingly, the numerical results show that the filter estimates are not uniformly better than the predictor estimates. This discrepancy is traced to the use of the initial data, suggesting a third algorithm: an optimal two-step filter that invokes an assimilation update at start-up (OTSFSU) to utilize the initial data. Aside from the initial start-up step, the OTSFSU is identical to the OTSF. We then revisit the accuracy/latency tradeoff by numerically comparing the accuracy of the OTSP and OTSFSU to determine whether the OTSFSU is uniformly more accurate than the OTSP. Spoiler alert: It is.

OPTIMAL ONE-STEP PREDICTOR

Consider the system

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \tag{1}$$
$$y_k = C_k x_k + D_k u_k + v_k, \tag{2}$$

Digital Object Identifier 10.1109/MCS.2019.2961588
Date of current version: 17 March 2020

1066-033X/20©2020IEEE

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

Controllers and observers are invariably implemented digitally; thus discrete-time algorithms deserve special attention.

where, for all $k \geq 0$, $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, $y_k \in \mathbb{R}^p$, $w_k \in \mathbb{R}^n$, and $v_k \in \mathbb{R}^p$, and the real matrices A_k, B_k, C_k, and D_k are of corresponding size. All stochastic processes are assumed to have finite second moments; the details of the densities are not relevant to the derivation below. Specifically, $w_k \in \mathbb{R}^n$ is a zero-mean, white-noise disturbance signal with the covariance $Q_k \triangleq \mathbb{E}[w_k w_k^{\mathrm{T}}] \in \mathbb{R}^{n\times n}$, and $v_k \in \mathbb{R}^p$ is a zero-mean, white sensor noise with the covariance $R_k \triangleq \mathbb{E}[v_k v_k^{\mathrm{T}}] \in \mathbb{R}^{p\times p}$. The cross-covariance between v_k and w_k is denoted by $S_k \triangleq \mathbb{E}[w_k v_k^{\mathrm{T}}] \in \mathbb{R}^{n\times p}$. The input u_k is assumed to be known; in practice, u_k is typically a control input. Consider the OOSP

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + K_k (y_k - \hat{y}_k), \tag{3}$$

$$\hat{y}_k = C_k \hat{x}_k + D_k u_k, \tag{4}$$

where $\hat{x}_k \in \mathbb{R}^n$ is an estimate of the state x_k; the optimal gain $K_k \in \mathbb{R}^{n\times p}$ is determined below.

Defining the state error

$$e_k \triangleq x_k - \hat{x}_k, \tag{5}$$

it follows that

$$\begin{aligned} e_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\ &= \tilde{A}_k e_k + \tilde{w}_k, \end{aligned} \tag{6}$$

where

$$\tilde{A}_k \triangleq A_k - K_k C_k, \tag{7}$$

$$\tilde{w}_k \triangleq w_k - K_k v_k. \tag{8}$$

Next, we define the cost

$$\begin{aligned} J_k(K_k) &\triangleq \mathbb{E}[e_{k+1}^{\mathrm{T}} e_{k+1}] \\ &= \mathrm{tr}\ \mathbb{E}[e_{k+1}^{\mathrm{T}} e_{k+1}] \\ &= \mathbb{E}[\mathrm{tr}\ (e_{k+1}^{T} e_{k+1})] \\ &= \mathbb{E}[\mathrm{tr}(e_{k+1} e_{k+1}^{\mathrm{T}})] \\ &= \mathrm{tr}\ \mathbb{E}[e_{k+1} e_{k+1}^{\mathrm{T}}] \\ &= \mathrm{tr}\ P_{k+1}, \end{aligned} \tag{9}$$

where the *error covariance* is defined by

$$P_{k+1} \triangleq \mathbb{E}[e_{k+1} e_{k+1}^{\mathrm{T}}] \in \mathbb{R}^{n\times n}. \tag{10}$$

It thus follows that

$$\begin{aligned} e_{k+1} e_{k+1}^{\mathrm{T}} &= (\tilde{A}_k e_k + \tilde{w}_k)(\tilde{A}_k e_k + \tilde{w}_k)^{\mathrm{T}} \\ &= \tilde{A}_k e_k e_k^{\mathrm{T}} \tilde{A}_k^{\mathrm{T}} + \tilde{A}_k e_k \tilde{w}_k^{\mathrm{T}} + \tilde{w}_k e_k^{\mathrm{T}} \tilde{A}_k^{\mathrm{T}} + \tilde{w}_k \tilde{w}_k^{\mathrm{T}}. \end{aligned} \tag{11}$$

Next, since w_k and v_k are white-noise sequences that affect e_{k+1} but not e_k, it follows that e_k and $\tilde{w}_k$ are uncorrelated. Furthermore, since w_k and v_k have a zero mean, it follows that $\tilde{w}_k$ also has a zero mean. Therefore, $\mathbb{E}[e_k \tilde{w}_k^{\mathrm{T}}] = \mathbb{E}[e_k]\mathbb{E}[\tilde{w}_k^{\mathrm{T}}] = 0$. Taking the expected value of (11) yields

$$\begin{aligned} P_{k+1} &= \mathbb{E}[e_{k+1} e_{k+1}^{\mathrm{T}}] \\ &= \mathbb{E}[\tilde{A}_k e_k e_k^{\mathrm{T}} \tilde{A}_k^{\mathrm{T}} + \tilde{A}_k e_k \tilde{w}_k^{\mathrm{T}} + \tilde{w}_k e_k^{\mathrm{T}} \tilde{A}_k^{\mathrm{T}} + \tilde{w}_k \tilde{w}_k^{\mathrm{T}}] \\ &= \tilde{A}_k \mathbb{E}[e_k e_k^{\mathrm{T}}]\tilde{A}_k^{\mathrm{T}} + \tilde{A}_k \mathbb{E}[e_k \tilde{w}_k^{\mathrm{T}}] + \mathbb{E}[\tilde{w}_k e_k^{T}]\tilde{A}_k^{\mathrm{T}} + \mathbb{E}[\tilde{w}_k \tilde{w}_k^{\mathrm{T}}] \\ &= \tilde{A}_k P_k \tilde{A}_k^{\mathrm{T}} + \mathbb{E}[\tilde{w}_k \tilde{w}_k^{\mathrm{T}}] \\ &= \tilde{A}_k P_k \tilde{A}_k^{\mathrm{T}} + \mathbb{E}[(w_k - K_k v_k)(w_k - K_k v_k)^{\mathrm{T}}] \\ &= \tilde{A}_k P_k \tilde{A}_k^{\mathrm{T}} + \mathbb{E}[w_k w_k^{\mathrm{T}} - w_k v_k^{\mathrm{T}} K_k^{\mathrm{T}} - K_k v_k w_k^{\mathrm{T}} + K_k v_k v_k^{\mathrm{T}} K_k^{\mathrm{T}}] \\ &= \tilde{A}_k P_k \tilde{A}_k^{\mathrm{T}} + \mathbb{E}[w_k w_k^{\mathrm{T}}] - \mathbb{E}[w_k v_k^{\mathrm{T}}]K_k^{\mathrm{T}} - K_k \mathbb{E}[v_k w_k^{\mathrm{T}}] \\ &\quad + K_k \mathbb{E}[v_k v_k^{\mathrm{T}}]K_k^{\mathrm{T}} \\ &= \tilde{A}_k P_k \tilde{A}_k^{\mathrm{T}} + Q_k + K_k R_k K_k^{\mathrm{T}} - \mathbb{E}[w_k v_k^{\mathrm{T}}]K_k^{\mathrm{T}} - K_k[\mathbb{E}[w_k v_k^{\mathrm{T}}]]^{\mathrm{T}}. \end{aligned} \tag{12}$$

The cost (9) is thus given by

$$\begin{aligned} J_k(K_k) &= trP_{k+1} \\ &= \mathrm{tr}[(A_k - K_k C_k)P_k(A_k - K_k C_k)^{\mathrm{T}} + Q_k + K_k R_k K_k^{\mathrm{T}} \\ &\quad - S_k K_k^{\mathrm{T}} - K_k S_k^{\mathrm{T}}] \\ &= \mathrm{tr}[A_k P_k A_k^{\mathrm{T}} - K_k C_k P_k A_k^{\mathrm{T}} - A_k P_k C_k^{\mathrm{T}} K_k^{\mathrm{T}} + K_k C_k P_k C_k^{\mathrm{T}} K_k^{\mathrm{T}} \\ &\quad + Q_k + K_k R_k K_k^{\mathrm{T}} - S_k K_k^{\mathrm{T}} - K_k S_k^{\mathrm{T}}] \\ &= \mathrm{tr}[K_k(C_k P_k C_k^{\mathrm{T}} + R_k)K_k^{\mathrm{T}} - K_k C_k P_k A_k^{\mathrm{T}} - A_k P_k C_k^{\mathrm{T}} K_k^{\mathrm{T}} \\ &\quad + A_k P_k A_k^{\mathrm{T}} + Q_k - S_k K_k^{\mathrm{T}} - K_k S_k^{\mathrm{T}}] \\ &= \mathrm{tr}[K_k(C_k P_k C_k^{\mathrm{T}} + R_k)K_k^{\mathrm{T}}] - 2\mathrm{tr}\ K_k(C_k P_k A_k^{\mathrm{T}} + S_k^{\mathrm{T}}) \\ &\quad + \mathrm{tr}(A_k P_k A_k^{\mathrm{T}} + Q_k). \end{aligned} \tag{13}$$

To minimize $J_k(K_k)$, note that

$$\frac{\mathrm{d}J_k(K_k)}{\mathrm{d}K_k} = 2(C_k P_k C_k^{\mathrm{T}} + R_k)K_k^{\mathrm{T}} - 2C_k P_k A_k^{\mathrm{T}} - 2S_k^{\mathrm{T}}. \tag{14}$$

Setting the derivative to zero yields the *optimal one-step predictor gain*

$$K_k = (A_k P_k C_k^{\mathrm{T}} + S_k)(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}. \tag{15}$$

By substituting (7) and (15) in (12), the *one-step predictor* is thus given by

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + (A_k P_k C_k^{\mathrm{T}} + S_k)(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1}(y_k - C_k \hat{x}_k), \tag{16}$$

$$\begin{aligned} P_{k+1} &= A_k P_k A_k^{\mathrm{T}} - (A_k P_k C_k^{\mathrm{T}} + S_k)(C_k P_k C_k^{\mathrm{T}} + R_k)^{-1} \\ &\quad \cdot (C_k P_k A_k^{\mathrm{T}} + S_k^{\mathrm{T}}) + Q_k. \end{aligned} \tag{17}$$

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

Optimal Two-Step Predictor and Two-Step Filter With Correlated Disturbance and Sensor Noise

This sidebar gives the update equations for the optimal two-step predictor (OTSP) and optimal two-step filter (OTSF) in the case where $S_k \neq 0$. For the OTSP, the *assimilation update* is given by

$$x_k^{\mathrm{a}} = x_k^{\mathrm{f}} + P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} (y_k - C_k x_k^{\mathrm{f}}), \tag{S1}$$

$$P_k^{\mathrm{a}} = P_k^{\mathrm{f}} - P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}}, \tag{S2}$$

which are identical to (20) and (21). The *forecast update* is given by

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{a}} + B_k u_k + S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} (y_k - C_k x_k^{\mathrm{f}}), \tag{S3}$$

$$\begin{aligned} P_{k+1}^{\mathrm{f}} = {} & A_k P_k^{\mathrm{a}} A_k^{\mathrm{T}} - S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}} A_k^{\mathrm{T}} \\ & - A_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} S_k^{\mathrm{T}} \\ & - S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} S_k^{\mathrm{T}} + Q_k, \end{aligned} \tag{S4}$$

which are (22) and (23) with additional terms involving S_k. Defining the *OTSP gain*

$$K_k^{\mathrm{f}} \triangleq P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{S5}$$

(S1) and (S2) can be written as

$$x_k^{\mathrm{a}} = x_k^{\mathrm{f}} + K_k (y_k - C_k x_k^{\mathrm{f}}), \tag{S6}$$

$$P_k^{\mathrm{a}} = (I - K_k C_k) P_k^{\mathrm{f}}, \tag{S7}$$

which are identical to (25) and (26).

For the OTSF, the *forecast update* is given by

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{a}} + B_k u_k + S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} (y_k - C_k x_k^{\mathrm{f}}), \tag{S8}$$

$$\begin{aligned} P_{k+1}^{\mathrm{f}} = {} & A_k P_k^{\mathrm{a}} A_k^{\mathrm{T}} - S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}} A_k^{\mathrm{T}} \\ & - A_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} S_k^{\mathrm{T}} \\ & - S_k (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} S_k^{\mathrm{T}} + Q_k, \end{aligned} \tag{S9}$$

which are (27) and (28) with additional terms involving S_k. The *assimilation update* is given by

$$\begin{aligned} x_{k+1}^{a} = {} & x_{k+1}^{f} + P_{k+1}^{f} C_{k+1}^{T} (C_{k+1} P_{k+1}^{f} C_{k+1}^{T} + R_{k+1})^{-1} \\ & \cdot (y_{k+1} - C_{k+1} x_{k+1}^{f}), \end{aligned} \tag{S10}$$

$$P_{k+1}^{\mathrm{a}} = P_{k+1}^{\mathrm{f}} - P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} C_{k+1} P_{k+1}^{\mathrm{f}}, \tag{S11}$$

which are identical to (29) and (30). In terms of the *OTSF gain*

$$K_{k+1}^{\mathrm{f}} \triangleq P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1}, \tag{S12}$$

(S10) and (S11) can be written as

$$x_{k+1}^{\mathrm{a}} = x_{k+1}^{\mathrm{f}} + K_{k+1} (y_{k+1} - C_{k+1} x_{k+1}^{\mathrm{f}}), \tag{S13}$$

$$P_{k+1}^{\mathrm{a}} = (I - K_{k+1} C_{k+1}) P_{k+1}^{\mathrm{f}}, \tag{S14}$$

which are identical to (32) and (33).

The OTSF in the traditional notation is

$$\begin{aligned} \hat{x}_{k+1|k} = {} & A_k \hat{x}_{k|k} + B_k u_k + S_k (C_k P_{k|k-1} C_k^T + R_k)^{-1} \\ & \cdot (y_k - C_k \hat{x}_{k|k-1}), \end{aligned} \tag{S15}$$

$$\begin{aligned} P_{k+1|k} = {} & A_k P_{k|k} A_k^T - S_k (C_k P_{k|k-1} C_k^T + R_k)^{-1} \\ & \cdot C_k P_{k|k-1} A_k^T - A_k P_{k|k-1} C_k^T (C_k P_{k|k-1} C_k^T + R_k)^{-1} S_k^T \\ & - S_k (C_k P_{k|k-1} C_k^T + R_k)^{-1} S_k^T + Q_k, \end{aligned} \tag{S16}$$

$$\begin{aligned} \hat{x}_{k+1|k+1} = {} & \hat{x}_{k+1|k} + P_{k+1|k} C_{k+1}^T (C_{k+1} P_{k+1|k} C_{k+1}^T + R_{k+1})^{-1} \\ & \cdot (y_{k+1} - C_{k+1} \hat{x}_{k+1|k}), \end{aligned} \tag{S17}$$

$$\begin{aligned} P_{k+1|k+1} = {} & P_{k+1|k} - P_{k+1|k} C_{k+1}^T (C_{k+1} P_{k+1|k} C_{k+1}^T + R_{k+1})^{-1} \\ & \cdot C_{k+1} P_{k+1|k}, \end{aligned} \tag{S18}$$

where (S15) and (S16) are (38) and (39) with additional terms involving S_k, and (S17) and (S18) are identical to (40) and (41). The OTSF gain (S12) can be written as

$$K_{k+1}^{\mathrm{f}} = P_{k+1|k}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1|k}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1}, \tag{S19}$$

and thus (S17) and (S18) can be written as

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}^{\mathrm{f}} (y_{k+1} - C_{k+1} \hat{x}_{k+1|k}), \tag{S20}$$

$$P_{k+1|k+1} = (I - K_{k+1}^{\mathrm{f}} C_{k+1}) P_{k+1|k}, \tag{S21}$$

which are identical to (43) and (44).

Note that the error-covariance propagation equation is independent of data. Furthermore, (16) and (17) can be written in terms of K_k as

$$\hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k + K_k (y_k - C_k \hat{x}_k), \tag{18}$$

$$P_{k+1} = A_k P_k A_k^{\mathrm{T}} - K_k (C_k P_k A_k^{\mathrm{T}} + S_k^{\mathrm{T}}) + Q_k. \tag{19}$$

OPTIMAL TWO-STEP PREDICTOR

As an alternative but equivalent implementation of the optimal predictor, (16) and (17) can be implemented as the OTSP. For simplicity in this and subsequent sections, consider the case where $S_k = 0$. The case where S_k is nonzero is discussed in "Optimal Two-Step Predictor and Optimal Two-Step Filter With Correlated Disturbance and Sensor Noise." The *assimilation update* is given by

$$x_k^{\mathrm{a}} = x_k^{\mathrm{f}} + P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} (y_k - C_k x_k^{\mathrm{f}}), \tag{20}$$

$$P_k^{\mathrm{a}} = P_k^{\mathrm{f}} - P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1} C_k P_k^{\mathrm{f}}, \tag{21}$$

and the *forecast update* is given by

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{a}} + B_k u_k, \tag{22}$$

$$P_{k+1}^{\mathrm{f}} = A_k P_k^{\mathrm{a}} A_k^{\mathrm{T}} + Q_k. \tag{23}$$

Figure 1 shows a timing diagram for the real-time implementation of the OTSP. By substituting (20) into (22) and (21) into (23), it can be seen that $x_k^{\mathrm{f}} = \hat{x}_k$ and $P_k^{\mathrm{f}} = P_k$. Defining the *OTSP gain*

$$K_k^{\mathrm{f}} \triangleq P_k^{\mathrm{f}} C_k^{\mathrm{T}} (C_k P_k^{\mathrm{f}} C_k^{\mathrm{T}} + R_k)^{-1}, \tag{24}$$

(20) and (21) can be written as

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

$$x_k^{\mathrm{a}} = x_k^{\mathrm{f}} + K_k(y_k - C_k x_k^{\mathrm{f}}), \tag{25}$$

$$P_k^{\mathrm{a}} = (I - K_k C_k) P_k^{\mathrm{f}}. \tag{26}$$

It is interesting to note that the computational requirements of the OTSP are more burdensome than the OOSP, and yet the OTSP and OOSP produce identical state estimates. Therefore, there is no implementation advantage of the OTSP over the OOSP. However, the value of the OTSP is in providing a framework for the OTSF, as discussed next.

OPTIMAL TWO-STEP FILTER

In contrast to the OTSP in (20)–(23), the OTSF is given by the *forecast update*

$$x_{k+1}^{\mathrm{f}} = A_k x_k^{\mathrm{a}} + B_k u_k, \tag{27}$$

$$P_{k+1}^{\mathrm{f}} = A_k P_k^{\mathrm{a}} A_k^{\mathrm{T}} + Q_k, \tag{28}$$

and the *assimilation update*

$$x_{k+1}^{\mathrm{a}} = x_{k+1}^{\mathrm{f}} + P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} \cdot (y_{k+1} - C_{k+1} x_{k+1}^{\mathrm{f}}), \tag{29}$$

$$P_{k+1}^{\mathrm{a}} = P_{k+1}^{\mathrm{f}} - P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} (C_{k+1} P_{k+1}^{\mathrm{f}} C_{k+1}^{\mathrm{T}} + R_{k+1})^{-1} C_{k+1} P_{k+1}^{\mathrm{f}}. \tag{30}$$

Note that (20) and (21) are slightly different from (29) and (30); specifically, (20) and (21) use the measurement y_k, whereas (29) and (30) use the measurement y_{k+1}. Moreover, the forecast and assimilation updates of the OTSF appear in reverse order compared to the forecast and assimilation updates of the OTSP. This reversal explains the index k in (20) and (21) and the index $k+1$ in (29) and (30).

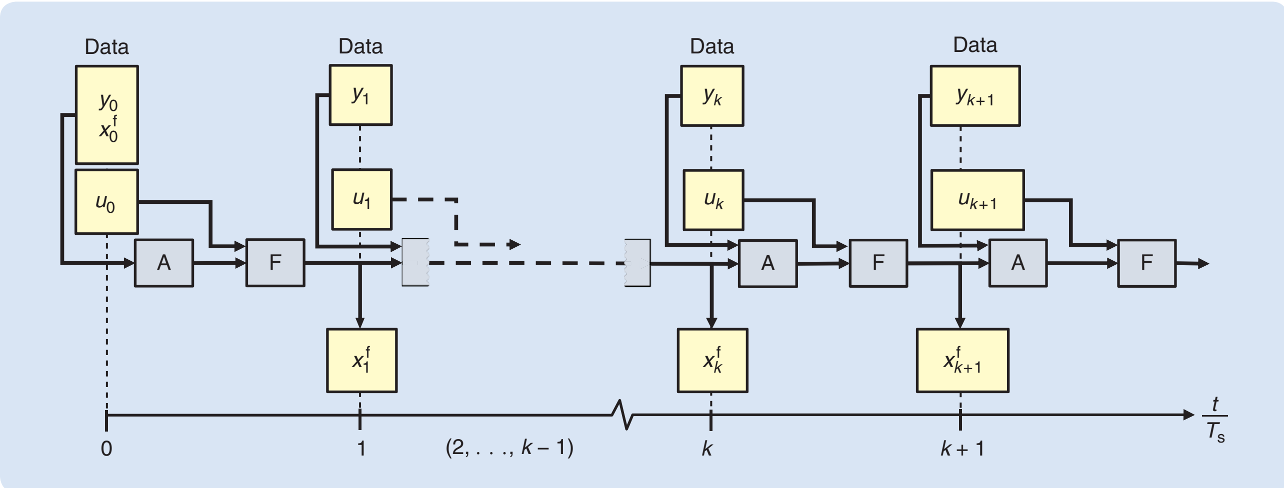

FIGURE 1 The timing diagram for the optimal two-step predictor (OTSP), where A denotes an assimilation update and F denotes a forecast update. The measurement y_k and applied control u_k are available at step k for computation. For all $k \geq 1$, the OTSP produces an estimate of x_k at step k without latency.

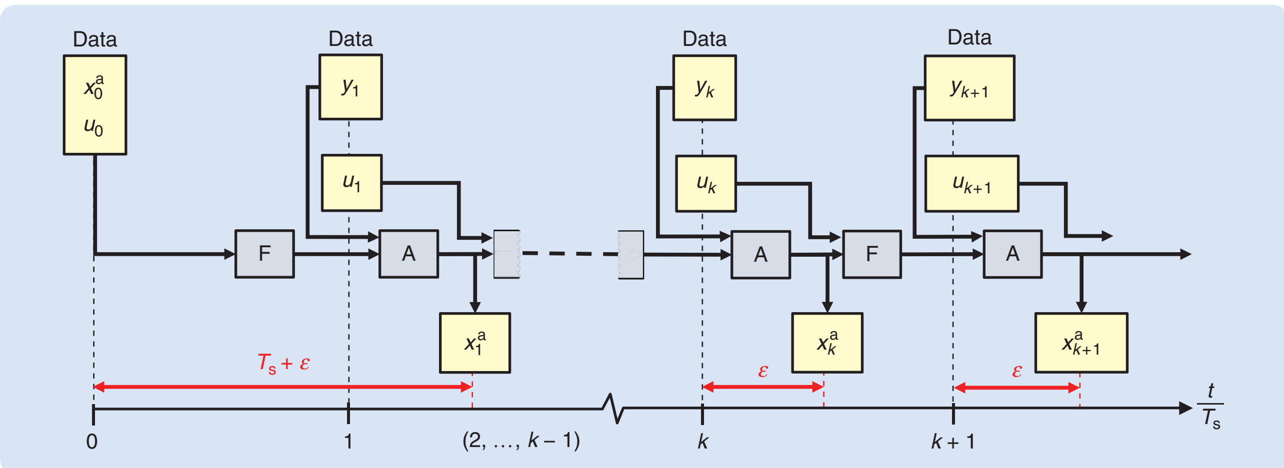

FIGURE 2 The timing diagram for the optimal two-step filter (OTSF), where A denotes an assimilation update and F denotes a forecast update. The measurement y_k and applied control u_k are available at step k for computation. For all $k \geq 1$, the OTSF produces an estimate of x_k with latency ε.

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

Next, note that, since x^{a}_{k+1} depends on y_{k+1} and the required computation cannot be performed instantaneously, the estimate x^{a}_{k+1} of x_{k+1} is not available at step $k+1$ but at some time afterward, where the latency depends on the computer speed and architecture. Therefore, if the latency of x^{a}_k as an estimate of x_k is critical in a real-time application, the OTSP may be a better choice than the OTSF. Figure 2 provides a timing diagram for the real-time implementation of the OTSF.

In terms of the *OTSF gain*

$$K^{\mathrm{f}}_{k+1} \triangleq P^{\mathrm{f}}_{k+1}C^{\mathrm{T}}_{k+1}(C_{k+1}P^{\mathrm{f}}_{k+1}C^{\mathrm{T}}_{k+1}+R_{k+1})^{-1}, \tag{31}$$

(29) and (30) can be written as

$$x^{\mathrm{a}}_{k+1} = x^{\mathrm{f}}_{k+1} + K_{k+1}(y_{k+1}-C_{k+1}x^{\mathrm{f}}_{k+1}), \tag{32}$$

$$P^{\mathrm{a}}_{k+1} = (I-K_{k+1}C_{k+1})P^{\mathrm{f}}_{k+1}. \tag{33}$$

Note that, aside from a shift in the index, the OTSF gain is identical to the OTSP gain.

OPTIMAL ONE-STEP FILTER

The OOSF can be obtained by substituting (27) and (28) into (29) and (30). Specifically, the OOSF is given by

$$\begin{aligned} x^{\mathrm{a}}_{k+1} = {} & A_k x^{\mathrm{a}}_k + B_k u_k + (A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k)C^{\mathrm{T}}_{k+1} \\ & \cdot [C_{k+1}(A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k)C^{\mathrm{T}}_{k+1} + R_{k+1}]^{-1} \\ & \cdot [y_{k+1} - C_{k+1}(A_k x^{\mathrm{a}}_k + B_k u_k)], \end{aligned} \tag{34}$$

$$\begin{aligned} P^{\mathrm{a}}_{k+1} = {} & A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k - (A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k) \\ & \cdot C^{\mathrm{T}}_{k+1}[C_{k+1}(A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k)C^{\mathrm{T}}_{k+1} + R_{k+1}]^{-1} \\ & \cdot C_{k+1}(A_k P^{\mathrm{a}}_k A^{\mathrm{T}}_k + Q_k). \end{aligned} \tag{35}$$

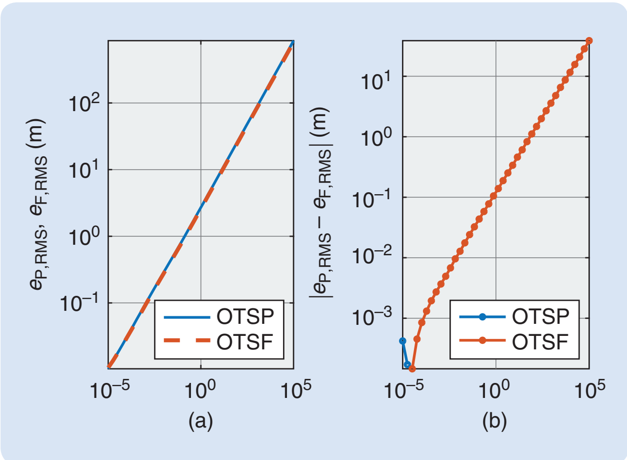

FIGURE 3 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter (OTSF) versus the disturbance-covariance scaling α_{d}. Part (a) shows the RMS position-estimation errors for the OTSP and OTSF, $e_{\mathrm{P,RMS}}$ and $e_{\mathrm{F,RMS}}$, respectively. Part (b) shows $|e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}|$. In all cases where $e_{\mathrm{P,RMS}} > e_{\mathrm{F,RMS}}$, the OTSF is more accurate than the OTSP; these are shown in red. In all cases where $e_{\mathrm{P,RMS}} < e_{\mathrm{F,RMS}}$, the OTSP is more accurate than the OTSF; these are shown in blue. Note that the OTSP is more accurate than the OTSF for $\alpha_{\mathrm{d}} < 10^{-4.5}$, whereas the OTSF is more accurate than the OTSP for $\alpha_{\mathrm{d}} > 10^{-4.5}$.

OPTIMAL TWO-STEP FILTER AND TRADITIONAL KALMAN FILTER

We now restate the OTSF in standard notation to show that it is precisely the Kalman filter. Defining

$$\hat{x}_{k|k-1} \triangleq x^{\mathrm{f}}_k, \quad P_{k|k-1} \triangleq P^{\mathrm{f}}_k, \tag{36}$$

$$\hat{x}_{k|k} \triangleq x^{\mathrm{a}}_k, \quad P_{k|k} \triangleq P^{\mathrm{a}}_k, \tag{37}$$

(27)–(30) can be written as

$$\hat{x}_{k+1|k} = A_k\hat{x}_{k|k} + B_k u_k, \tag{38}$$

$$P_{k+1|k} = A_k P_{k|k} A^{\mathrm{T}}_k + Q_k, \tag{39}$$

$$\begin{aligned} \hat{x}_{k+1|k+1} = {} & \hat{x}_{k+1|k} + P_{k+1|k}C^{\mathrm{T}}_{k+1}(C_{k+1}P_{k+1|k}C^{\mathrm{T}}_{k+1} + R_{k+1})^{-1} \\ & \cdot (y_{k+1} - C_{k+1}\hat{x}_{k+1|k}), \end{aligned} \tag{40}$$

$$\begin{aligned} P_{k+1|k+1} = {} & P_{k+1|k} - P_{k+1|k}C^{\mathrm{T}}_{k+1}(C_{k+1}P_{k+1|k}C^{\mathrm{T}}_{k+1} + R_{k+1})^{-1} \\ & \cdot C_{k+1}P_{k+1|k}, \end{aligned} \tag{41}$$

which is the standard notation for the Kalman filter [2]. Furthermore, the OTSF gain (31) can be written as

$$K^{\mathrm{f}}_{k+1} = P^{\mathrm{f}}_{k+1|k}C^{\mathrm{T}}_{k+1}\left(C_{k+1}P^{\mathrm{f}}_{k+1|k}C^{\mathrm{T}}_{k+1} + R_{k+1}\right)^{-1}, \tag{42}$$

and thus (40) and (41) can be written as

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K^{\mathrm{f}}_{k+1}(y_{k+1} - C_{k+1}\hat{x}_{k+1|k}), \tag{43}$$

$$P_{k+1|k+1} = (I - K^{\mathrm{f}}_{k+1}C_{k+1})P_{k+1|k}. \tag{44}$$

An alternative notation [1], [3], [4] is given by

$$\hat{x}^{+}_k \triangleq x^{\mathrm{f}}_k, \quad P^{+}_k \triangleq P^{\mathrm{f}}_k, \tag{45}$$

$$\hat{x}^{-}_k \triangleq x^{\mathrm{a}}_k, \quad P^{-}_k \triangleq P^{\mathrm{a}}_k. \tag{46}$$

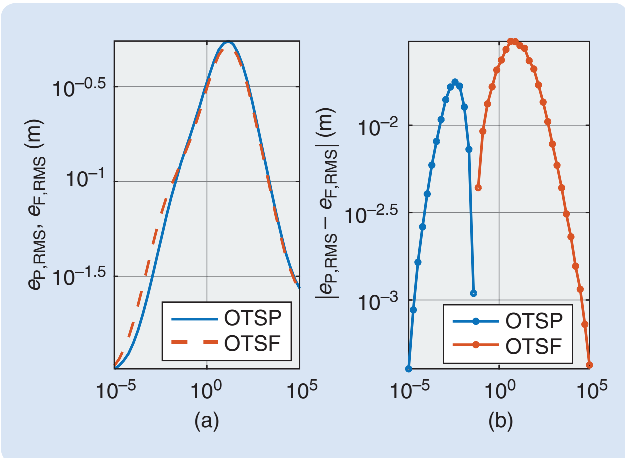

FIGURE 4 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter (OTSF) versus the sensor-noise covariance α_{sn}. Part (a) shows the RMS position-estimation errors for the OTSP and OTSF, $e_{\mathrm{P,RMS}}$ and $e_{\mathrm{F,RMS}}$, respectively. Part (b) shows that the OTSP is more accurate than the OTSF for $\alpha_{\mathrm{sn}} < 10^{-2}$, whereas the OTSF is more accurate than the OTSP for $\alpha_{\mathrm{sn}} > 10^{-2}$.

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

OPTIMAL TWO-STEP PREDICTOR VERSUS OPTIMAL TWO-STEP FILTER

In this section, we investigate whether or not the OTSF estimates are more accurate than the OTSP estimates. Consider an undamped oscillator with a mass of 4 kg and stiffness of 2 N/m, sampled at 4 Hz using a zero-order hold input. The velocity is measured, and the position is to be estimated. The sampled-data system representing this arrangement has the dynamics

$$\begin{bmatrix} x_{k+1} \\ \dot{x}_{k+1} \end{bmatrix} = \begin{bmatrix} 0.9844 & 0.2487 \\ -0.1243 & 0.9844 \end{bmatrix} \begin{bmatrix} x_k \\ \dot{x}_k \end{bmatrix} + \begin{bmatrix} 0.0078 \\ 0.0622 \end{bmatrix} u_k + w_k, \quad (47)$$

$$y_k = [0 \ \ 1] \begin{bmatrix} x_k \\ \dot{x}_k \end{bmatrix} + v_k, \quad (48)$$

where $x_k \in \mathbb{R}$ is the position in meters, $y_k = \dot{x}_k \in \mathbb{R}$ is the velocity in meters/second, u_k is the applied force in Newtons, $w_k \in \mathbb{R}^2$ is the disturbance, and $v_k \in \mathbb{R}$ is the sensor noise.

To investigate the effect of the disturbance w_k, sensor noise v_k, and initial condition x_0 on the estimates of the position produced by the OTSP and OTSF, let

$$u_k \sim \mathcal{N}(0,1), w_k \sim \mathcal{N}(0, \alpha_{\rm d} I), v_k \sim \mathcal{N}(0, \alpha_{\rm sn}), x_0 \sim \mathcal{N}(0, \alpha_{\rm ic} I), \quad (49)$$

where $\alpha_{\rm d}, \alpha_{\rm sn}, \alpha_{\rm ic}$ are varied individually, with the remaining variables fixed. Specifically, we first vary $\alpha_{\rm d} \in [10^{-5}, 10^5]$, with $\alpha_{\rm sn} = \alpha_{\rm ic} = 10^{-5}$ fixed. Next, we vary $\alpha_{\rm sn} \in [10^{-5}, 10^5]$, with $\alpha_{\rm d} = \alpha_{\rm ic} = 10^{-5}$ fixed. Finally, we vary $\alpha_{\rm ic} \in [10^{-5}, 10]$, with $\alpha_{\rm d} = \alpha_{\rm sn} = 10^{-5}$ fixed. In all cases, the initial states of the OTSP and OTSF are set to zero and $P_0^{\rm f} = I_2, P_0^{\rm a} = I_2$.

In each case and for each choice of $\alpha_{\rm d}, \alpha_{\rm sn}, \alpha_{\rm ic}$, 10,000 simulations are run with randomly generated values of u_k, w_k, and v_k and the initial condition x_0. For each simulation, the root mean square (RMS) position-estimation errors for the OTSP and OTSF are computed for $0 \le k \le 80$. The RMS position-estimation errors for the OTSP and OTSF are then averaged across the 10,000 simulations. The RMS position-estimation errors for the OTSP and OTSF are defined to be $e_{\rm P,RMS}(\alpha_{\rm d}, \alpha_{\rm sn}, \alpha_{\rm ic})$ and $e_{\rm F,RMS}(\alpha_{\rm d}, \alpha_{\rm sn}, \alpha_{\rm ic})$, respectively. Specifically, we compute

$$e_{\rm P,RMS} = \frac{1}{10{,}000} \sum_{j=1}^{10{,}000} \sqrt{\frac{1}{80} \sum_{i=1}^{80} (x_{i,j}^{\rm f} - x_{i,j})^2}, \quad (50)$$

$$e_{\rm F,RMS} = \frac{1}{10{,}000} \sum_{j=1}^{10{,}000} \sqrt{\frac{1}{80} \sum_{i=1}^{80} (x_{i,j}^{\rm a} - x_{i,j})^2}, \quad (51)$$

where $x_{i,j}^{\rm f}$ is the position estimate produced by the OTSP at the ith step for the jth simulation, $x_{i,j}^{\rm a}$ is the position estimate produced by the OTSF at the ith step for the jth simulation,

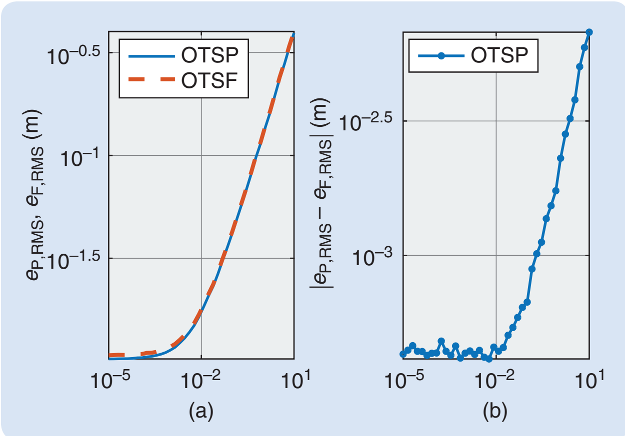

FIGURE 5 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter (OTSF) versus the initial-condition-covariance scaling, $\alpha_{\rm ic}$. Part (a) shows the RMS position-estimation errors for the OTSP and OTSF, $e_{\rm P,RMS}$ and $e_{\rm F,RMS}$, respectively. Part (b) shows that the OTSP is uniformly more accurate than the OTSF.

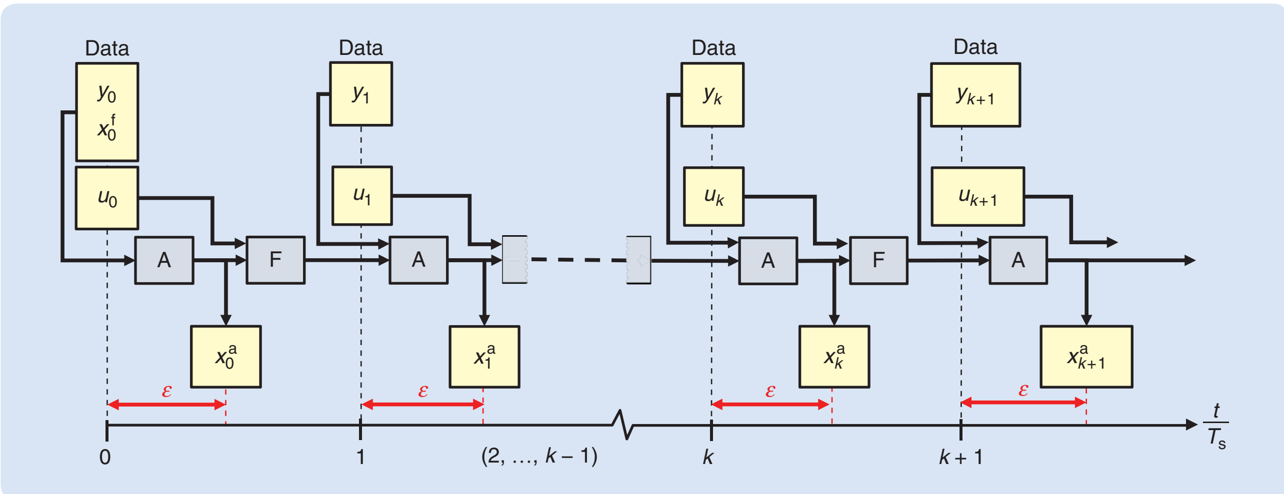

FIGURE 6 The timing diagram for the optimal two-step filter with start-up (OTSFSU), where A denotes an assimilation update and F denotes a forecast update. The measurement y_k and applied control u_k are available at step k for computation. For all $k \ge 0$, the OTSFSU produces an estimate of x_k with latency ε.

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

The filter can be realized only as a two-step algorithm, whereas the predictor can be realized as either a one- or two-step algorithm.

and $x_{i,j}$ is the true position at the ith step for the jth simulation. Furthermore, $|e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}|$ is plotted to show the accuracy of the OTSF position estimate relative to the OTSP position estimate in meters. Specifically, $e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}$ is positive in the case where the OTSF RMS position-estimation error is smaller than the OTSP RMS position-estimation error and vice versa.

For the three cases considered above, Figures 3(a), 4(a), and 5(a) show $e_{\mathrm{P,RMS}}$, $e_{\mathrm{F,RMS}}$ versus α_{d}, α_{sn}, α_{ic}, respectively. Figures 3(b), 4(b), and 5(b) graph $|e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}|$ versus α_{d}, α_{sn}, α_{ic}, respectively, where the values of $|e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}|$ are color coded based on the sign of $e_{\mathrm{P,RMS}} - e_{\mathrm{F,RMS}}$. Figure 5 shows that the OTSP position estimate is more accurate than the OTSF position estimate for all choices of α_{ic}. This situation is surprising since we expect the latency of the OTSF estimates to be offset by greater accuracy. We thus seek a variation of the OTSF that produces the expected improved accuracy in return for latency, as discussed in the next section.

TABLE 1 The initial estimate $\hat{x}_0$ and real-time data y and u used by the estimator to approximate x_k. The time at which the estimate of x_k becomes available is given in terms of the step k, sample time T_{s}, and latency ε.

Estimator	$\hat{x}_0$	y Data	u Data	$\hat{x}_k$	When Available?
OTSP	x_0^{f}	y_{k-1}	u_{k-1}	x_k^{f}	kT_{s}
OTSF	x_0^{a}	y_k	u_{k-1}	x_k^{a}	$kT_{\mathrm{s}} + \varepsilon$
OTSFSU	x_0^{f}	y_k	u_{k-1}	x_k^{a}	$kT_{\mathrm{s}} + \varepsilon$

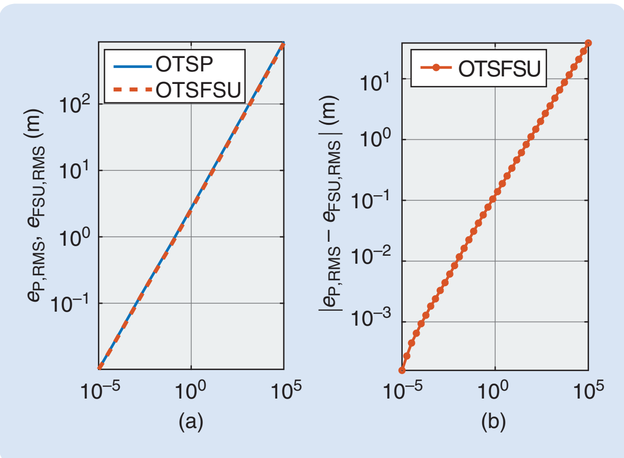

FIGURE 7 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter with start-up (OTSFSU) versus the disturbance-covariance scaling, α_{d}. Part (a) shows the RMS position-estimation errors for the OTSP and OTSFSU $e_{\mathrm{P,RMS}}$, and $e_{\mathrm{FSU,RMS}}$, respectively. Part (b) shows that the OTSFSU is uniformly more accurate than the OTSP.

OPTIMAL TWO-STEP FILTER WITH START-UP

As shown in Figure 5, the OTSP may be more accurate than the OTSF. This is surprising because the OTSF produces latency estimates and uses more recent data, leading to the expectation of more accurate state estimates. This phenomenon can be traced to the fact that OTSF does not use the measurement y_0. In fact, at step $k = 0$ (20) uses y_0, whereas at step $k = 0$ (29) does not use y_0. This situation suggests the possibility of a variant of the OTSF that employs an additional assimilation update before the initial forecast update; otherwise, all subsequent updates of the OTSFSU are identical to the OTSF. Figure 6 presents a timing diagram for the real-time implementation of the OTSFSU. The next section investigates the accuracy of the OTSFSU compared to the OTSP to determine if the OTSFSU is more accurate, as compensation for its inherent latency.

REAL-TIME IMPLEMENTATION

To clarify the timing of the computation required for the OTSP, OTSF, and OTSFSU, note that for the OTSP the state estimate is x_k^{f}, whereas for the OTSF and OTSFSU the state estimate is x_k^{a}. Figures 1, 2, and 6 show timing diagrams for the real-time implementation of the OTSP, OTSF, and OTSFSU. Note that the OTSP is the only estimator that produces an estimate of x_k at step k. In contrast, the OTSF and OTSFSU produce estimates of x_k at time $kT_{\mathrm{s}} + \varepsilon$, where T_{s} is the sample time of the digital implementation and $\varepsilon \in (0, T_{\mathrm{s}})$ is the latency (which is platform dependent). As pictured in Figure 2, the first state estimate produced by the OTSF is an estimate of x_1; this estimate is not available until $t = T_{\mathrm{s}} + \varepsilon$. In addition, this estimate of x_1 does not use the data y_0, as can be seen by the absence of y_0 in Figure 2. Alternately, as displayed in Figure 6, the OTSFSU uses y_0 to produce an estimate of x_0 at $t = \varepsilon$. Table 1 lists the data used by each algorithm and the associated latency.

OPTIMAL TWO-STEP PREDICTOR VERSUS OPTIMAL TWO-STEP FILTER WITH START-UP

We now compare the OTSP and OTSFSU using the same procedure employed to compare the OTSP and OTSF. For the OTSFSU, define $e_{\mathrm{FSU,RMS}}(\alpha_{\mathrm{d}}, \alpha_{\mathrm{sn}}, \alpha_{\mathrm{ic}})$, and plot $|e_{\mathrm{P,RMS}} - e_{\mathrm{FSU,RMS}}|$. Specifically, compute

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.

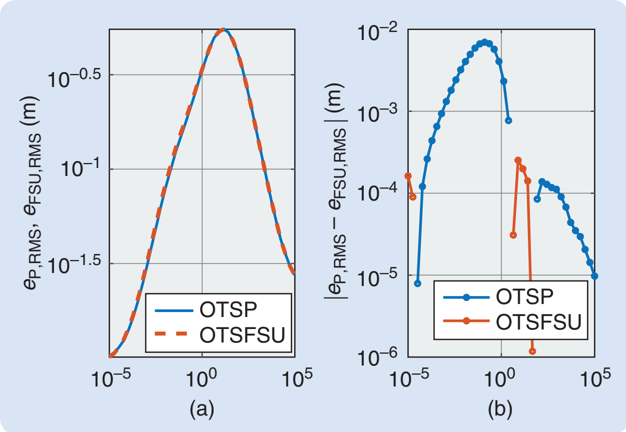

FIGURE 8 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter with start-up (OTSFSU) versus the sensor-noise-covariance scaling, α_{sn}. Part (a) shows the RMS position-estimation errors for the OTSP and OTSFSU, $e_{\text{P,RMS}}$ and $e_{\text{FSU,RMS}}$, respectively. Part (b) shows that the OTSP is more accurate than the OTSFSU for $10^{-4.5} < \alpha_{\text{sn}} < 10^{-2}$ and $\alpha_{\text{sn}} > 10^{2}$.

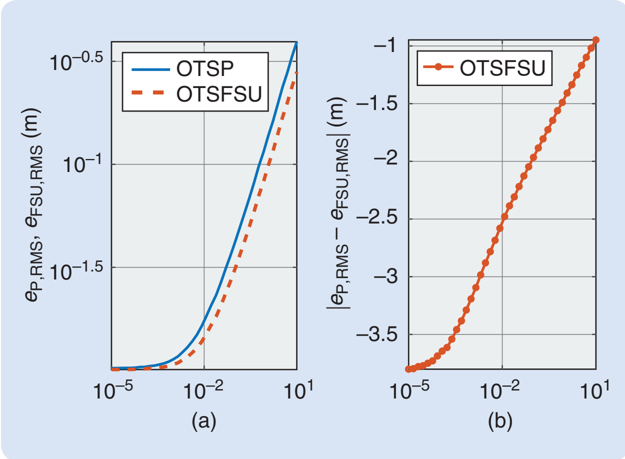

FIGURE 9 The root-mean-square (RMS) position-estimation errors for the optimal two-step predictor (OTSP) and optimal two-step filter with start-up (OTSFSU) versus the initial-condition-covariance scaling, α_{ic}. Part (a) shows the RMS position-estimation errors for the OTSP and OTSFSU, $e_{\text{P,RMS}}$ and $e_{\text{FSU,RMS}}$, respectively. Part (b) shows that the OTSFSU is uniformly more accurate than the OTSP.

$$e_{\text{FSU,RMS}} = \frac{1}{10{,}000}\sum_{j=1}^{10{,}000}\sqrt{\frac{1}{80}\sum_{i=1}^{80}(x_{i,j}^{\text{a}} - x_{i,j})^2}, \tag{52}$$

where $x_{i,j}^{\text{a}}$ is the position estimate produced by OTSFSU at the ith step for the jth simulation, and $x_{i,j}$ is the true position at the ith step for the jth simulation. Figures 7–9 show that the OTSFSU position estimate is more accurate than the OTSP position estimate for all choices of α_{d} and α_{ic}. Furthermore, as displayed in Figure 8, the OTSP shows no distinct advantage over the OTSFSU for variations of α_{sn}.

CONCLUSIONS

The goal of this article was to untangle the relationship between the OTSP and OTSF. A fundamental distinction between the filter and the predictor is that the filter can be realized only as a two-step algorithm, whereas the predictor can be realized as either a one- or two-step algorithm. An analysis by a timing diagram shows that the estimates produced by the OTSF take advantage of more recent data but have latency due to the time needed for computation. Partially contradicting the expected latency/accuracy tradeoff, a numerical example showed that the OTSF is not uniformly better than the OTSF. This discrepancy was traced to the nonuse of initial data by the OTSF, which suggested a variation of the OTSF that implements an additional assimilation update at start-up; otherwise the OTSF and the OTSFSU are identical. Revisiting the numerical example showed that the OTSFSU is uniformly more accurate than the OTSP. As with the OTSF, this additional accuracy comes at the price of latency.

ACKNOWLEDGMENTS

This research was partially supported by the Air Force Office of Scientific Research under Dynamic Data-Driven Applications Systems grant FA9550-16-1-0071 (http://www.1dddas.org/).

AUTHOR INFORMATION

Syed Aseem Ul Islam (aseemisl@umich.edu) received the B.S. degree in aerospace engineering from the Institute of Space Technology, Islamabad, Pakistan, and he is pursuing the Ph.D. degree in flight dynamics and control from the University of Michigan, Ann Arbor. His research interests include data-driven adaptive control for aerospace applications.

Ankit Goel received the B.E. degree in mechanical engineering from the Delhi College of Engineering, India, and the Ph.D. degree from the University of Michigan, Ann Arbor. His research interests include the data-driven estimation and control of high-dimensional complex systems.

Dennis S. Bernstein received the Sc.B. degree from Brown University, Providence, Rhode Island, and the Ph.D. degree from the University of Michigan, Ann Arbor, where he is a professor in the Aerospace Engineering Department. His research interests include the identification, estimation, and control for aerospace applications. He is the author of *Scalar, Vector, and Matrix Mathematics*, published by Princeton University Press.

REFERENCES

[1] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Elsevier, 1970.

[2] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.

[3] J. Crassidis and J. Junkins, *Optimal Estimation of Dynamic Systems*. Boca Raton, FL: CRC Press, 2004.

[4] D. Simon, *Optimal State Estimation: Kalman, H-Infinity, and Nonlinear Approaches*. Hoboken, NJ: Wiley, 2006.

[5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960. doi: 10.1115/1.3662552.

[6] B. Teixeira, "Kalman filters," *IEEE Control Syst. Mag.*, vol. 28, no. 2, pp. 16–18, Apr. 2008. doi: 10.1109/MCS.2008.4472371.

Authorized licensed use limited to: University of Michigan Library. Downloaded on March 19,2020 at 16:59:33 UTC from IEEE Xplore. Restrictions apply.