Diverse Derivation Methods and Expressions of Discrete-Time Finite Memory Structure Filter

Pyung Soo Kim,

Abstract—**State estimation using a filter is one of the fundamental problems encountered in the study of control theory and signal processing for noisy dynamics systems. In this review paper, the finite memory structure (FMS) filter for state estimation is expressed in a variety of forms, such as the iterative form, matrix form, summation form, and smoothing form, using diverse derivation methods for noisy discrete-time state-space models. These diverse derivation methods and expressions can provide a comprehensive understanding of FMS filtering algorithms. Various aspects of FMS filters are characterized. Through discussions, it is demonstrated that the choice of window length can be considered a key design parameter for optimizing the performance of the FMS filter. Simulation results for a noisy discrete-time system indicate that an FMS filter can be better than an infinite memory structure filter for temporarily uncertain systems.**

Index Terms—**Finite memory structure filter, Infinite memory structure filter, Iterative form, Matrix form, Sigma form, Smoothing form.**

I. Introduction

THE Kalman filter, used to estimate precise states in dynamic systems, has been widely applied in time series analysis in fields such as control theory, signal processing, statistics, and econometrics[1]-[11]. However, given their recursive formulation and infinite memory structure (IMS), Kalman filters may exhibit performance degradation, and even divergence, in severe cases of mismodeling and temporary uncertainties[12]-[14].

Therefore, as an alternative to the Kalman filter, the finite memory structure (FMS) filter has been designed for state estimation. The FMS filter has been demonstrated to be inherently more robust against temporary uncertainties[15]-[26]. It has been applied successfully in various engineering applications such as signal processing, mobile target tracking, computer networks, electromagnetic systems, fault detection, wireless sensor networks, automotive suspension systems, and electric motor systems[27]-[34].

To provide a comprehensive understanding of FMS filtering algorithms, this review paper provides several expressions, such as the iterative form, matrix form, summation form, and smoothing form, of the FMS filter by using diverse derivation methods for the discrete-time state-space model in white Gaussian noise environments. FMS filters are evaluated in terms of a variety of aspects, such as handling of window initial state, handling of noise covariance, processing manner, delay tolerance, and inversion computation of system matrix. It is shown that the window length can be considered as a useful design parameter to optimize the filtering performance of the FMS filter. It is demonstrated via computer simulations for a sinusoidal signal system and an electrical motor system that the FMS filter can outperform the IMS filter for temporarily uncertain systems.

This paper has the following structure. In Section II, the discrete-time state-space model and IMS filter are described. In Section III, diverse derivation ways and expressions of the FMS filter are provided and compared from a variety of views. In Section IV, how to choose the window length is discussed. In Section V, computer simulations are performed. Then, concluding remarks are given in Section VI.

Manuscript received January 22, 2021; revised March 31, 2021.

Pyung Soo Kim is a professor at Department of Electronic Engineering of Korea Polytechnic University, 237 Sangidaehak-ro, Siheung-si, Gyeonggi-do, 15073, KOREA(e-mail: pskim@kpu.ac.kr)

II. Infinite Memory Structure Filter

A. Discrete-Time State-Space Model

The state-space approach is a generalized time domain method for modeling, analyzing and designing a wide range of dynamic systems and is particularly well suited to digital computational technique. Discrete-time systems are either inherently discrete (e.g. models of bank accounts, national economy growth models, population growth models, digital words) or they are obtained as a result of sampling (discretization) of continuous-time systems. Thus, various discretized systems can be modeled by a following discrete-time state-space form with noises as well as input terms:

$$\begin{aligned} x_{i+1} &= Ax_i + Bu_i + Gw_i, \\ z_i &= Cx_i + v_i, \end{aligned} \tag{1}$$

where $x_i \in \Re^n$ is the unknown state, $u_i \in \Re^p$ is the control input, $z_i \in \Re^q$ is the known measurement. At the initial time i_0 of system, the state x_{i_0} is a random variable with a mean $\bar{x}_{i_0}$ and a covariance Σ_{i_0}. The system noise $w_i \in \Re^p$ and the measurement noise $v_i \in \Re^q$ are zero-mean white Gaussian and mutually uncorrelated. The covariances of w_i and v_i are denoted by positive definite matrices Q and R, respectively.

B. IMS Filter : Kalman Filter

The infinite memory structure(IMS) filter such as the well-known Kalman filter[1]-[11] provides a minimum variance state estimate $\hat{x}_i$, called the one-step predicted estimate of the system state x_i for the discrete-time state space model (1) as follow:

$$\begin{aligned} \hat{x}_{i+1} &= A\hat{x}_i + \Big[A\Sigma_i C^T(R + C\Sigma_i C^T)^{-1}\Big](z_i - C\hat{x}_i) \\ &\quad + Bu_i \\ &= A(I + \Sigma_i C^T R^{-1} C)^{-1}(\hat{x}_i + \Sigma_i C^T R^{-1} z_i) \\ &\quad + Bu_i, \end{aligned} \tag{2}$$

$$\begin{aligned} \Sigma_{i+1} &= A\Sigma_i A^T + GQG^T \\ &\quad - A\Sigma_i C^T(R + C\Sigma_i C^T)^{-1} C\Sigma_i A^T \\ &= A(I + \Sigma_i C^T R^{-1} C)^{-1}\Sigma_i A^T + GQG^T, \end{aligned} \tag{3}$$

where $\hat{x}_{i_0} = \bar{x}_{i_0}$ and Σ_i is the error covariance of the estimate $\hat{x}_i$ with initial value Σ_{i_0}. The IMS filter such as the Kalman filter has been a standard choice for the state estimation and thus a beautiful reference for diverse engineering areas for removing noise from a contaminated signal to help reveal important signal features and components. The IMS filter has the recursive formulation for computational efficiency. However, since the IMS filter processes all past measurements, it tends to accumulate estimation errors during its implementation. Therefore, the IMS filter has been known to show performance degradation and even divergence phenomena for mismodeling and temporary uncertainties.

III. Diverse Derivation Methods and Expressions of Finite Memory Structure Filter

To solve shortcomings of the IMS filter in some cases, the finite memory structure (FMS) filter using finite measurements on the most recent window $[i-M(\triangleq i_M), i]$ has been developed[12]-[26]. The FMS filter has been known to have inherent good properties such as unbiasedness, deadbeat, and intrinsic robustness, etc., and thus applied successfully for various applications[27]-[34]. The FMS filter is formulated over a fixed window of length M whose size does not increase with time. The window of past measurements moves forward in time at each sampling time when a new measurement is available. The FMS filter discards past measurements outside the window $[i_M, i]$.

A. Iterative Form

The FMS filter can be represented by the iterative form[15][16]. The iterative form of the FMS filter is obtained by combining the information form of the Kalman filter with the moving window formulation. The information form of the Kalman filter, called simply the information filter or the inverse-covariance filter, allows optimal, unbiased, recursive state estimation without an initial state estimate in many of its realizations[35]. The information form of the Kalman filter on the most recent window $[i_M, i]$ can be represented by

$$\begin{aligned}\hat{x}_{i_M+j+1} &= A(\Sigma_{i_M+j}^{-1} + C^T R^{-1} C)^{-1} \Sigma_{i_M+j}^{-1} \hat{x}_{i_M+j} \\ &\quad + A(\Sigma_{i_M+j}^{-1} + C^T R^{-1} C)^{-1} C^T R^{-1} z_{i_M+j} \\ &\quad + B u_{i_M+j}, \end{aligned} \tag{4}$$

where the error covariance Σ_{i_M+j} is given by

$$\begin{aligned}\Sigma_{i_M+j+1}^{-1} &= \Big[I + A^{-T}(\Sigma_{i_M+j}^{-1} \\ &\quad + C^T R^{-1} C) A^{-1} G Q G^T\Big]^{-1} \\ &\quad A^{-T}(\Sigma_{i_M+j}^{-1} + C^T R^{-1} C) A^{-1}. \end{aligned} \tag{5}$$

The window initial state $\hat{x}_{i_M}$ in (4) is assumed to be unknown and thus must have an infinite covariance $\Sigma_{i_M} = \infty$, which means $\Sigma_{i_M}^{-1} = 0$. Therefore, the error covariance Σ_{i_M+j} (5) can be represented by the time-invariant equation as follows:

$$\begin{aligned}\Sigma_{j+1}^{-1} &= \Big[I + A^{-T}(\Sigma_j^{-1} + C^T R^{-1} C) A^{-1} G Q G^T\Big]^{-1} A^{-T} \\ &\quad (\Sigma_j^{-1} + C^T R^{-1} C) A^{-1}. \end{aligned} \tag{6}$$

Using following definitions

$$\Omega_j \triangleq \Sigma_j^{-1}, \ \hat{\theta}_j \triangleq \Omega_j \hat{x}_{i_M+j},$$

the intermediate state estimate, denoted by $\hat{\theta}_j$, is defined from the information form (4) of the Kalman filter on the moving measurement window $[i_M, i]$ as follows:

$$\begin{aligned}\hat{\theta}_{j+1} &= \Big[I + A^{-T}(\Omega_j + C^T R^{-1} C) A^{-1} G Q G^T\Big]^{-1} A^{-T} \\ &\quad \Big[\hat{\theta}_j + C^T R^{-1} z_{i_M+j} \\ &\quad + (\Omega_j + C^T R^{-1} C) A^{-1} B u_{i_M+j}\Big], \\ \hat{\theta}_0 &= 0, \ 0 \le j \le M-1, \end{aligned} \tag{7}$$

where the error covariance Ω_j is obtained from (6) as follows:

$$\begin{aligned}\Omega_{j+1} &= \Big[I + A^{-T}(\Omega_j + C^T R^{-1} C) A^{-1} G Q G^T\Big]^{-1} A^{-T} \\ &\quad (\Omega_j + C^T R^{-1} C) A^{-1}, \end{aligned} \tag{8}$$

with $\Omega_0 = 0$. Then, the ultimate state estimate $\hat{x}_i$ at the current time i can be represented as follows[15]:

$$\hat{x}_i = \Omega_M^{-1} \hat{\theta}_M. \tag{9}$$

Since the iterative form of the FMS filter (9) is derived from the well-known information form of the Kalman filter with the moving window formulation, general readers might find it easy to understand the derivation of filtering algorithms. However, the infinite covariance, $\Sigma_{i_M} = \infty$, of the window initial state is sometimes awkward and seems to have no physical meaning. So, the iterative form of the FMS filter (9) might be so difficult to understand its optimality. In addition, the system matrix A is required to be nonsingular because the inverse of the system matrix appear in filtering algorithms (7) and (8).

B. Matrix Form

The FMS filter can be represented by the matrix form by solving optimization problems[17]-[23]. From the discrete-time state space model (1), finite measurements Z_i and inputs U_i on the most recent window $[i_M, i]$ can be expressed by the following regression form with the current state x_i:

$$Z_i - \Xi U_i = \Gamma x_i + \Lambda W_i + V_i, \tag{10}$$

where Z_i and U_i are defined by

$$Z_i \triangleq \begin{bmatrix} z_{i_M} \\ z_{i_M+1} \\ \vdots \\ z_{i-2} \\ z_{i-1} \end{bmatrix}, \ U_i \triangleq \begin{bmatrix} u_{i_M} \\ u_{i_M+1} \\ \vdots \\ u_{i-2} \\ u_{i-1} \end{bmatrix}, \tag{11}$$

and W_i, V_i have the same form as (11) for w_i, v_i, respectively. Matrices Γ, Ξ, and Λ are defined by

$$\Gamma \triangleq \begin{bmatrix} CA^{-M} \\ CA^{-M+1} \\ \vdots \\ CA^{-2} \\ CA^{-1} \end{bmatrix},$$

$$\Xi \triangleq -\begin{bmatrix} CA^{-1}B & CA^{-2}B & \cdots & CA^{-M}B \\ 0 & CA^{-1}B & \cdots & CA^{-M+1}B \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & CA^{-1}B \end{bmatrix},$$

$$\Lambda \triangleq - \begin{bmatrix} CA^{-1}G & CA^{-2}G & \cdots & CA^{-M}G \\ 0 & CA^{-1}G & \cdots & CA^{-M+1}G \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & CA^{-1}G \end{bmatrix}. \quad (12)$$

The noise term $\Lambda W_i + V_i$ in (10) is zero-mean white Gaussian with covariance Π given by

$$\Pi \triangleq \Lambda\Big[\mathrm{diag}(\overbrace{Q\ Q\ \cdots\ Q}^{M})\Big]\Lambda^T + \Big[\mathrm{diag}(\overbrace{R\ R\ \cdots\ R}^{M})\Big], \quad (13)$$

where $\mathrm{diag}(Q\ Q\ \cdots\ Q)$ and $\mathrm{diag}(R\ R\ \cdots\ R)$ denote block-diagonal matrices with M elements of Q and R, respectively.

Based on different estimation approaches, three kinds of matrix forms for the FMS filter have been developed as follows.

1) Best Linear Unbiased Estimation: The first matrix form is developed from *best linear unbiased estimation* approach in [35]. The matrix form of the FMS filter $\hat{x}_i$ is assumed to be obtained from

$$\hat{x}_i \triangleq \mathcal{H}\Big(Z_i - \Xi U_i\Big), \quad (14)$$

where $\mathcal{H}$ is the gain matrix. Taking the expectation both sides of (14), the following relation is obtained:

$$\mathbf{E}\big[\hat{x}_i\big] = \mathbf{E}\Big[\mathcal{H}\Big(Z_i - \Xi U_i\Big)\Big] = \mathcal{H}\Gamma\mathbf{E}\big[x_i\big].$$

Then, with the following constraint:

$$\mathcal{H}\Gamma = I, \quad (15)$$

$\hat{x}_i$ is unbiased, i.e., $\mathbf{E}\big[\hat{x}_i\big] = \mathbf{E}\big[x_i\big]$. Thus, the constraint (15) can be called the *unbiasedness constraint* for the matrix form of the FMS filter $\hat{x}_i$.

The objective is now to obtain the gain matrix $\mathcal{H}_*$, subject to the unbiasedness constraint (15), in such a way that the error of $\hat{x}_i$ has a minimum variance as follows:

$$\mathcal{H}_* = \arg\min_{\mathcal{H}} \mathbf{E}\Big[(x_i - \hat{x}_i)^T(x_i - \hat{x}_i)\Big]. \quad (16)$$

Using the approach of best linear unbiased estimation in [35], the matrix form of the FMS filter $\hat{x}_i$ is obtained by the solution of (16) as follows[12][17][18]:

$$\hat{x}_i = \mathcal{H}\Big(Z_i - \Xi U_i\Big), \quad (17)$$

where

$$\mathcal{H} = (\Gamma^T\Pi^{-1}\Gamma)^{-1}\Gamma^T\Pi^{-1}.$$

2) Maximum Likelihood Estimation: The second matrix form is developed from *maximum likelihood estimation* approach in [35]. A maximum likelihood estimation approach is introduced to obtain the noise suppressed state estimate $\hat{x}_i$ of the current state x_i. The connotation of *maximum likelihood* is a setting in which nothing is known *a priori* about the unknown state, but there is *a priori* information on the measurement process itself. The noise term $\Lambda W_i + V_i$ in (10) has the following multivariate Gaussian density function:

$$f\Big(\Lambda W_i + V_i\Big) = \frac{1}{\sqrt{(2\pi)^M|\Pi|}}\, e^{\left[-\frac{1}{2}\left(\Lambda W_i + V_i\right)^T \Pi^{-1}\left(\Lambda W_i + V_i\right)\right]}.$$

It is noted that linear transformation on, and linear combinations of, Gaussian random processes are themselves Gaussian random processes. For this reason, it is clear from the equation (10) that when $\Lambda W_i + V_i$ is Gaussian, $Z_i - \Xi U_i$ is as well. The multivariate Gaussian density function of $Z_i - \Xi U_i$ is derived from a shifted version of $f\Big(\Lambda W_i + V_i\Big)$ as follows:

$$f\Big(\big\{Z_i - \Xi U_i\big\}|x_i\Big) = f\Big(Z_i - \Xi U_i - \Gamma x_i\Big) = \frac{1}{\sqrt{(2\pi)^M|\Pi|}}\, e^{\left[-\frac{1}{2}\left(Z_i - \Xi U_i - \Gamma x_i\right)^T \Pi\left(Z_i - \Xi U_i - \Gamma x_i\right)\right]},$$

called the *likelihood function*. The maximum likelihood filter is obtained from the maximizing of this likelihood function with respect to x_i. To maximize $f\big(\big\{Z_i - \Xi U_i\big\}|x_i\big)$ with respect to x_i is equivalent to the minimization problem of the following cost function:

$$J = \frac{1}{2}\Big(Z_i - \Xi U_i - \Gamma x_i\Big)^T \Pi\Big(Z_i - \Xi U_i - \Gamma x_i\Big). \quad (18)$$

Differentiating both sides of (18) gives the following maximum likelihood criterion:

$$\frac{\partial J}{\partial x_i} = \Gamma^T\Pi^{-1}\Big(Z_i - \Xi U_i - \Gamma x_i\Big) = 0, \quad (19)$$

called the *likelihood equation*. Assume that $\{A, C\}$ is observable and $M \geq n$, the matrix form of the FMS filter $\hat{x}_i$ is then given by the solution of the likelihood equation (19) as follows [19][20]:

$$\hat{x}_i = \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big(Z_i - \Xi U_i\Big). \quad (20)$$

This matrix form (20) is equivalent to (17).

3) Bayesian Estimation: The third matrix form is developed from *Bayesian estimation* approach in [35]. As shown in Bayesian estimation filtering[21][22], the FMS filter can be interested in the Gaussian probability density function that is conditional on finite measurements Z_i and inputs U_i on the most recent window $[i_M, i]$. The most recent window $[i_M, i]$ becomes the averaging window of M points. To develop an alternative matrix form of the FMS filter, the conditional density of current state x_i given finite measurements Z_i and inputs U_i is derived. On the most recent window $[i_M, i]$, $Z_i - \Xi U_i$ (10) can be expressed by

$$\Gamma x_i = Z_i - \Xi U_i - \Big(\Lambda W_i + V_i\Big), \quad (21)$$

with the noise term $\Lambda W_i + V_i$. Then, multiplying both sides of (21) by $\Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma_i^T\Pi^{-1}$ leads to

$$\begin{aligned} x_i &= \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big[Z_i - \Xi U_i - \Big(\Lambda W_i + V_i\Big)\Big] \\ &= \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big(Z_i - \Xi U_i\Big) \\ &\quad -\Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big(\Lambda W_i + V_i\Big). \end{aligned} \quad (22)$$

Hence, for given finite measurements Z_i and inputs U_i, the equation (22) clearly means that the current state x_i is a multi-variate Gaussian with its mean

$$\hat{x}_i = \mathbf{E}[x_i] = \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big(Z_i - \Xi U_i\Big),$$

and covariance

$$\begin{aligned}\Sigma &= \Big[\Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big]\Pi_i\Big[\Pi^{-1}\Gamma\Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Big]\\ &= \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Big(\Gamma^T\Pi^{-1}\Gamma\Big)\Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\\ &= \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}.\end{aligned}$$

Therefore, from the linearity described in (10), the conditional density of current state x_i given finite measurements Z_i and inputs U_i has the following expression:

$$p\Big(x_i|Z_i - \Xi U_i\Big) = \mathcal{N}\Big(x_i; \hat{x}_i, \Sigma\Big). \quad (23)$$

The mean value $\hat{x}_i$ from the conditional density (23) of current state x_i is adopted as the FMS filter and thus the following matrix form

$$\hat{x}_i = \Big(\Gamma^T\Pi^{-1}\Gamma\Big)^{-1}\Gamma^T\Pi^{-1}\Big(Z_i - \Xi U_i\Big) \quad (24)$$

provides the state estimate $\hat{x}_i$ conditional on finite measurements Z_i and inputs U_i. This matrix form (24) is equivalent to (17) and (20).

4) Forgetting Factor Least Square Estimation: The fourth matrix form is developed from *least squares estimation* using a forgetting factor when there is no *a priori* information about noise covariances Q and R[23]. Given given finite measurements Z_i and inputs U_i on the window $[i_M, i]$, the FMS filter $\hat{x}_i$ is obtained from the following forgetting factor least squares criterion:

$$\hat{x}_i = \arg\min_{x_i}\ \Big[Z_i - \Xi U_i - \Gamma x_i\Big]^T\tilde{\Pi}\Big[Z_i - \Xi U_i - \Gamma x_i\Big], \quad (25)$$

where $\tilde{\Pi}$ is a diagonal matrix as follows:

$$\tilde{\Pi} \triangleq diag[\tilde{\pi}^{M-1}I\ \tilde{\pi}^{M-2}I\ \cdots\ \tilde{\pi}^0 I], \quad 0 < \tilde{\pi} < 1,$$

where $\tilde{\pi}$ is called the forgetting factor. Note that a main role of the forgetting factor $\tilde{\pi}$ is to account for the fact that the discrete-time state-space model (1) is not perfect to globally model the observed phenomenon, thus is to make the model that is locally well modeling the observations by concentrating on observations on the most recent window $[i_M, i]$. Then, when $\{A, C\}$ is observable and $M \geq n$, the solution of (25) is given by

$$\begin{aligned}\hat{x}_i &= (\Gamma^T\tilde{\Pi}^{\frac{T}{2}}\tilde{\Pi}^{\frac{1}{2}}\Gamma)^{-1}\Gamma^T\tilde{\Pi}^{\frac{T}{2}}\tilde{\Pi}^{\frac{1}{2}}\big(Z_i - \Xi U_i\big)\\ &= A^M(\tilde{\Gamma}^T\tilde{\Gamma})^{-1}\tilde{\Gamma}^T\big(\tilde{Z}_i - \tilde{\Xi}\tilde{U}_i\big)\\ &= A^M\check{\Gamma}\tilde{\Gamma}^T\big(\tilde{Z}_i - \tilde{\Xi}\tilde{U}_i\big) \quad (26)\end{aligned}$$

where

$$\begin{aligned}\check{\Gamma} &= (\tilde{\Gamma}^T\tilde{\Gamma})^{-1},\\ \tilde{\Gamma} &= \tilde{\Pi}^{\frac{1}{2}}\bar{\Gamma} = \begin{bmatrix}\tilde{\Gamma}_*\\ CA^{M-1}\end{bmatrix} = \begin{bmatrix}\tilde{\pi}^{\frac{M-1}{2}}C\\ \tilde{\Gamma}_* A\end{bmatrix},\\ \tilde{Z}_i &= \tilde{\Pi}^{\frac{1}{2}}Z_i = \begin{bmatrix}\tilde{Z}_i^U\\ z_{i-1}\end{bmatrix} = \begin{bmatrix}\tilde{\pi}^{\frac{M-1}{2}}z_{i_M}\\ \tilde{Z}_i^L\end{bmatrix},\end{aligned}$$

where $\tilde{\Gamma}_*$ is upper $M-1$ rows of $\tilde{\Gamma}$, and $\tilde{Z}_{i-1}^U$ and $\tilde{Z}_{i-1}^L$ are upper $M-1$ rows and lower $M-1$ rows of $\tilde{Z}_{i-1}$, respectively. Note that the gain matrix $A^M\check{\Gamma}\tilde{\Gamma}^T$ requires computation only on the interval $[0, M]$ once and is time-invariant for all windows. This means that the FMS filter is time-invariant.

C. Sigma Form

The FMS filter can be represented by the sigma form(also known as the summation form), which might be similar to the finite impulse response (FIR) form used widely in digital signal processing areas[24]. The standard Kalman filter (2) and (3) can be modified by $\hat{x}_{i_M+j}$ on the window $[i_M, i]$ as follows:

$$\begin{aligned}\hat{x}_{i_M+j+1} &= A(I + \Sigma_{i_M+j}C^TR^{-1}C)^{-1}(\hat{x}_{i_M+j}\\ &\quad + \Sigma_{i_M+j}C^TR^{-1}z_{i_M+j}) + Bu_{i_M+j}, \quad (27)\end{aligned}$$

where the error covariance Σ_{i_M+j} is given by

$$\begin{aligned}\Sigma_{i_M+j+1} &= A(I + \Sigma_{i_M+j}C^TR^{-1}C)^{-1}\Sigma_{i_M+j}A^T\\ &\quad + GQG^T, \quad (28)\end{aligned}$$

and $0 \leq j \leq M-1$. It is noted in (27) that $\hat{x}_{i_M+j}$ for $i_M + j < i$ is an intermediate variable to compute $\hat{x}_i$. That is, only $\hat{x}_i$ is used as the ultimate estimate of x_i and can be represented by the sigma form with the window initial condition as well as the control input as follows:

$$\begin{aligned}\hat{x}_i &= \Phi_M\hat{x}_{i_M} + \sum_{j=0}^{M-1}\Phi_{M-j}\Sigma_{i_M+j}C^TR^{-1}z_{i_M+j}\\ &\quad + \sum_{j=0}^{M-1}\Phi_{M-j}Bu_{i_M+j}, \quad (29)\end{aligned}$$

where the transition matrix Φ_j is given by

$$\Phi_{j+1} = \Phi_j A\big[I + \Sigma_{i_M+M-j-1}C^TR^{-1}C\big]^{-1},\ \Phi_0 = I.$$

At this point, *a posteriori* knowledge about the window initial condition $\{\hat{x}_{i_M}, \Sigma_{i_M}\}$ on the window $[i_M, i]$ is required for (28) and (29).

With the window initial state x_{i_M}, finite measurements Z_i and inputs U_i on the most recent window $[i_M, i]$ can be expressed by the following regression form

$$Z_i - \bar{\Xi}U_i = \bar{\Gamma}x_{i_M} + \bar{\Lambda}W_i + V_i, \quad (30)$$

where matrices $\bar{\Gamma}$, $\bar{\Xi}$ and $\bar{\Lambda}$ are defined by

$$\bar{\Gamma} \triangleq \begin{bmatrix}C\\ CA\\ \vdots\\ CA^{M-2}\\ CA^{M-1}\end{bmatrix},$$

$$\bar{\Xi} \triangleq \begin{bmatrix}0 & 0 & \cdots & 0 & 0\\ CB & 0 & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ CA^{M-3}B & CA^{M-4}B & \cdots & 0 & 0\\ CA^{M-2}B & CA^{M-3}B & \cdots & CB & 0\end{bmatrix},$$

$$\bar{\Lambda} \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CG & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{M-3}G & CA^{M-4}G & \cdots & 0 & 0 \\ CA^{M-2}G & CA^{M-3}G & \cdots & CG & 0 \end{bmatrix}. \quad (31)$$

The noise term $\bar{\Lambda}W_i+\bar{V}_i$ in (30) is zero-mean white Gaussian with covariance $\bar{\Pi}$ given by

$$\bar{\Pi} \triangleq \bar{\Lambda}\Big[\mathrm{diag}(\overbrace{Q\ Q\ \cdots\ Q}^{M})\Big]\bar{\Lambda}^T + \Big[\mathrm{diag}(\overbrace{R\ R\ \cdots\ R}^{M})\Big].$$

Then, using the approach of best linear unbiased estimation in [35], the window initial condition $\hat{x}_{i_M}$ is obtained by

$$\hat{x}_{i_M} = \bar{\mathcal{H}}\Big(Z_i - \Xi U_i\Big) = \bar{\mathcal{H}}Z_i + \bar{\mathcal{B}}U_i, \quad (32)$$

where

$$\begin{aligned} \bar{\mathcal{H}} &\triangleq [\bar{\mathcal{H}}_0\ \bar{\mathcal{H}}_1\ \cdots\ \bar{\mathcal{H}}_{M-1}] \\ &= \big(\bar{\Gamma}^T\bar{\Pi}^{-1}\bar{\Gamma}\big)^{-1}\bar{\Gamma}^T\bar{\Pi}^{-1}, \quad (33) \\ \bar{\mathcal{B}} &\triangleq [\bar{\mathcal{B}}_0\ \bar{\mathcal{B}}_1\ \cdots\ \bar{\mathcal{B}}_{M-1}] \\ &= -\bar{\mathcal{H}}\Xi = -\big(\bar{\Gamma}^T\bar{\Pi}^{-1}\bar{\Gamma}\big)^{-1}\bar{\Gamma}^T\bar{\Pi}^{-1}\Xi. \quad (34) \end{aligned}$$

In addition, the window initial condition Σ_{i_M} is obtained by the error covariance of $\hat{x}_{i_M}$ as follows:

$$\begin{aligned} \Sigma_{i_M} &= \mathbf{E}\big[\big(x_{i_M} - \bar{x}_{i_M}\big)\big(x_{i_M} - \bar{x}_{i_M}\big)^T\big] \\ &= \mathbf{E}\big[\big(x_{i_M} - \bar{\mathcal{H}}Z_i - \bar{\mathcal{B}}U_i\big)\big(x_{i_M} - \bar{\mathcal{H}}Z_i - \bar{\mathcal{B}}U_i\big)^T\big] \\ &= \big(\bar{\Gamma}^T\bar{\Pi}^{-1}\bar{\Gamma}\big)^{-1}. \quad (35) \end{aligned}$$

Therefore, $a\ posteriori$ knowledge about the window initial condition $\{\hat{x}_{i_M}, \Sigma_{i_M}\}$ in (32) and (35) on the window $[i_M, i]$ is given for (28) and (29) in the *unbiasedness* sense. As shown in (35), the window initial condition Σ_{i_M} is constant value. Thus, the error covariance Σ_{i_M+j} (28) defined on the window $[i_M, i]$ can be rewritten as follows:

$$\begin{aligned} \Sigma_{j+1} &= A(I+\Sigma_j C^T R^{-1} C)^{-1}\Sigma_j A^T + GQG^T, \\ &\quad 0 \le j \le M-1, \quad (36) \end{aligned}$$

with the window initial condition $\Sigma_0 = \Sigma_{i_M} = (\bar{\Gamma}^T\bar{\Pi}^{-1}\bar{\Gamma})^{-1}$.

Using (11), (33), (34), (36), the FMS filter (29) with the window initial condition $\hat{x}_{i_M} = \bar{\mathcal{H}}Z_i + \bar{\mathcal{B}}U_i$ can be rewritten by

$$\begin{aligned} \hat{x}_i &= \Phi_M \hat{x}_{i_M} + \sum_{j=0}^{M-1} \Phi_{M-j}\Sigma_j C^T R^{-1} z_{i_M+j} \\ &\quad + \sum_{j=0}^{M-1} \Phi_{M-j} B u_{i_M+j} \\ &= \Phi_M\Big(\bar{\mathcal{H}}Z_i + \bar{\mathcal{B}}U_i\Big) \\ &\quad + \sum_{j=0}^{M-1} \Phi_{M-j}\Sigma_j C^T R^{-1} z_{i_M+j} \\ &\quad + \sum_{j=0}^{M-1} \Phi_{M-j} B u_{i_M+j}, \quad (37) \end{aligned}$$

where the transition matrix Φ_j is given by

$$\Phi_{j+1} = \Phi_j A\big[I + \Sigma_{M-j-1}C^T R^{-1} C\big]^{-1},\ \Phi_0 = I. \quad (38)$$

Therefore, the sigma form of the FMS filter (37) is given as follow:

$$\begin{aligned} \hat{x}_i &= \sum_{j=0}^{M-1}\Big(\Phi_M\bar{\mathcal{H}}_j + \Phi_{M-j}\Sigma_j C^T R^{-1}\Big) z_{i_M+j} \\ &\quad + \sum_{j=0}^{M-1}\Big(\Phi_M\bar{\mathcal{B}}_j + \Phi_{M-j}B\Big) u_{i_M+j}. \quad (39) \end{aligned}$$

In the sigma form of the FMS filter (39), the system matrix A is not required to be nonsingular unlike both iterative and matrix forms of the FMS filter because the inverse of the system matrix disappears in filtering algorithms (31), (36), and (38).

D. Smoothing Form

Meanwhile, because the FMS filter is a causal filter providing estimates for states at given times based only on the relative past, the estimates exhibit a delay. Hence, the FMS smoothing filter has been developed for estimation problems where there is a fixed delay between a measurement and the availability of its estimate[13][25][26]. This fixed delay is associated only with the availability of the estimate - not with an error in the actual estimates, as is the case with the FMS filter. Although FMS smoothing filters in [13][25][26] have their own unique features, they have the following common advantages. The smoothing filter generally utilizes more measurement information than the filter to provide state estimates, which can give more accurate estimation performance than the filter. In addition, since the smoothing filter provides state estimates at the delayed time using measurement information up to the current time, measurement information can be reflected in advance in the presence of the state change, which can give more fast convergence than the filter.

The FMS smoothing filter to estimate the state x_{i-d} at the lagged time $i-d$ is developed under a weighted least square criterion using only finite measurements as well as inputs on the most recent window $[i-M, i]$. The lagged time $i-d$ means there is a fixed delay between the measurement and the availability of its estimate. The positive integer d is the delay length satisfying $0 \le d < M$ and equal to the number of discrete time steps between the lagged time $i-d$ at which the state is to be estimated and the current time i of the last measurement used in estimating it.

From the discrete-time state-space model (1), the state x_{i-d} at the lagged time $i-d$ is represented by

$$x_{i-d} = A^{M-d}x_{i_M} + \tilde{\Xi}U_i + \tilde{\Lambda}W_i, \quad (40)$$

where

$$\begin{aligned} \tilde{\Xi} &\triangleq \Big[A^{M-d-1}B\ \cdots\ AB\ B\ \overbrace{0\ 0\ \cdots\ 0}^{d}\Big], \\ \tilde{\Lambda} &\triangleq \Big[A^{M-d-1}G\ \cdots\ AG\ G\ \overbrace{0\ 0\ \cdots\ 0}^{d}\Big]. \quad (41) \end{aligned}$$

Therefore, with applying (40) to (30), the following regression form can be expressed in terms with x_{i-d} at the lagged time $i-d$ as follows:

$$Z_i - \Xi_s U_i = \Gamma_s x_{i-d} + \Lambda_s W_i + V_i, \qquad (42)$$

where

$$\begin{aligned} \Gamma_s &\triangleq \bar{\Gamma} A^{-(M-d)}, \ \Lambda_s \triangleq \bar{\Lambda} - \bar{\Gamma} A^{-(M-d)} \tilde{\Lambda}, \\ \Xi_s &\triangleq \bar{\Xi} - \bar{\Gamma} A^{-(M-d)} \tilde{\Xi}. \end{aligned} \qquad (43)$$

The noise term $\Lambda_s W_i + V_i$ in (42) is zero-mean white Gaussian with covariance Π given by

$$\begin{aligned} \Pi_s &\triangleq \Lambda_s \Big[\mathrm{diag}(\overbrace{Q \ Q \ \cdots \ Q}^{M}) \Big] \Lambda_s^T \\ &\quad + \Big[\mathrm{diag}(\overbrace{R \ R \ \cdots \ R}^{M}) \Big]. \end{aligned} \qquad (44)$$

Now, to get the FMS smoothing filter $\hat{x}_{i-d}$ given finite measurements Z_i and inputs U_i on the most recent window $[i_M, i]$, the following weighted least square cost function must be minimized:

$$\begin{aligned} &\Big\{ \Big(Z_i - \Xi_s U_i \Big) - \Gamma_s x_{i-d} \Big\}^T \Pi_s^{-1} \\ &\quad \Big\{ \Big(Z_i - \Xi_s U_i \Big) - \Gamma_s x_{i-d} \Big\}. \end{aligned} \qquad (45)$$

Taking a derivation of (45) with respect to x_{i-d} and setting it to zero, the FMS smoother $\hat{x}_{i-d}$ is given by following simple matrix form:

$$\hat{x}_{i-d} = \Big(\Gamma_s^T \Pi^{-1} \Gamma_s \Big)^{-1} \Gamma_s^T \Pi^{-1} \Big(Z_i - \Xi_s U_i \Big). \qquad (46)$$

When there is no delay, that is $d = 0$, the FMS smoothing filter is shown to be equivalent to the FMS filter. With $d = 0$, matrices of (41) become

$$\begin{aligned} \tilde{\Xi} &= \Big[A^{M-1}B \ A^{M-2}B \ \cdots \ AB \ B \Big], \\ \tilde{\Lambda} &= \Big[A^{M-1}G \ A^{M-2}G \ \cdots \ AG \ G \Big]. \end{aligned} \qquad (47)$$

Thus, matrices Γ_s, Λ_s, Ξ_s and Π_s of (43) and (44) become the same as matrices Γ, Λ, Ξ and Π of (12) and (13) as follows:

$$\begin{aligned} \Gamma_s &= \bar{\Gamma} A^{-M} = \begin{bmatrix} CA^{-M} \\ CA^{-M+1} \\ \vdots \\ CA^{-2} \\ CA^{-1} \end{bmatrix} = \Gamma, \\ \Lambda_s &= \bar{\Lambda} - \bar{\Gamma} A^{-M} \tilde{\Lambda} \\ &= - \begin{bmatrix} CA^{-1}G & CA^{-2}G & \cdots & CA^{-M}G \\ 0 & CA^{-1}G & \cdots & CA^{-M+1}G \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & CA^{-2}G \\ 0 & 0 & \cdots & CA^{-1}G \end{bmatrix} \\ &= \Lambda, \end{aligned}$$

$$\begin{aligned} \Xi_s &= \bar{\Xi} - \bar{\Gamma} A^{-M} \tilde{\Xi} \\ &= - \begin{bmatrix} CA^{-1}B & CA^{-2}B & \cdots & CA^{-M}B \\ 0 & CA^{-1}B & \cdots & CA^{-M+1}B \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & CA^{-2}B \\ 0 & 0 & \cdots & CA^{-1}B \end{bmatrix} \\ &= \Xi, \\ \Pi_s &= \Lambda \Big[\mathrm{diag}(\overbrace{Q \ Q \ \cdots \ Q}^{M}) \Big] \Lambda^T + \Big[\mathrm{diag}(\overbrace{R \ R \ \cdots \ R}^{M}) \Big] \\ &= \Pi. \end{aligned} \qquad (48)$$

Then, the FMS smoothing filter (46) with $d = 0$ can be represented by

$$\hat{x}_i = \Big(\Gamma^T \Pi^{-1} \Gamma \Big)^{-1} \Gamma^T \Pi^{-1} \Big(Z_i - \Xi U_i \Big), \qquad (49)$$

with matrices (48). Therefore, the FMS smoothing filter (46) with $d = 0$ is equivalent to the existing FMS filter.

IV. Discussion about Window Length

The window length M can be a useful design parameter for the FMS filter. Thus, the important issue here is how to choose an appropriate window length M that makes the FMS filter's performance as good as possible. The noise suppression of the FMS filter might be closely related to the window length M, and it can have greater noise suppression as the window length M increases, which improves the FMS filter's performance. That is, choosing a larger M generally results in better performance, since more measurements are taken into account. However, at the same time, a larger M increases the convergence time of an FMS filtered estimate. Therefore, there might be FMS filter's compromise between noise suppression and tracking ability.

In addition, the window length may be related to the real-time computational complexity. Since the gain matrix for the FMS filter requires computation only on the interval $[0, M]$ once and is time-invariant for all windows, the on-line computation requires only filter updates. Thus, quite a large M can be chosen without worrying about computational burden. However, although the computational complexity of the FMS filter is $\mathcal{O}(M)$ and thus linear in the size of the window length M, the online computational complexity becomes larger as the window length increases.

Therefore, from an engineering perspective, there is trade-off that regulates the choice of the size of the window M. Since window length M is an integer, fine adjustment of the properties with M is difficult. Moreover, it is difficult to determine the window length systematically. In applications, one method of determining the window length is to take the appropriate value that can provide sufficient noise suppression. Therefore, it can be stated from the above discussions that the window length M can be considered a useful parameter to make the residual performance of the FMS filter as good as possible.

A heuristic would be to start the standard Kalman filtering estimation (2) by using all available measurements z_i, and determine M on-the-fly based on the error covariance Σ_M in (3). When Σ_M falls below a certain threshold, M is set and the FMS filter with moving window of measurements Z_i is started. Since the error covariance generally decreases

TABLE I: Comparison of FMS filters

	Iterative form	Matrix form				Sigma form	Smoothing form
		Best linear unbaised	Maximum likelihood	Bayesian	Least squares		
Handling of window initial state	Infinite covariance	Not required	Not required	Not required	Not required	Estimated	Not required
Handling of noise covariance	Required	Required	Required	Required	Not required	Required	Required
Processing manner	Iterative processing	Batch processing	Batch processing	Batch processing	Batch processing	Batch processing	Batch processing
Delay tolerance	No	No	No	No	No	No	Yes
Inversion computation of system matrix	Required	Required	Required	Required	Required	Not required	Required

over time, this heuristic allows us to choose of M in terms of the Kalman filtering error. An alternative heuristic would be to determine window length M in advance based on the error covariance Ω_j in (8) or Σ_j in (36). When the $L2$ norm of error covariance matrix Ω_M^{-1} or Σ_M falls below a certain threshold, M in the vicinity is set the window length M. Since the error covariance generally decreases over time, this heuristic allows to choose of M in terms of the $L2$ norm of the error covariance. This heuristic will be verified through an application example in the next section.

V. Application Examples with Computer Simulations

To illustrate the validity of the FMS filter and to compare with the recursive IMS filter, computer simulations are performed. Even if various dynamic systems and signal systems are represented in state-space model accurately on a long time scale, it may undergo unpredictable changes, such as jumps in frequency, phase, and velocity. Because these effects typically occur over a short time horizon, they are called temporary uncertainties[12]-[14]. As representative temporary uncertainties, there are a model uncertainty, an unknown input, and incomplete measurement information, etc. The state estimation filter for dynamic systems should be robust to diminish the effects of these temporary uncertainties. As representative temporary uncertainties, there are a model uncertainty, an unknown input, and incomplete measurement information, etc. The state estimation filter for dynamic systems should be robust to diminish the effects of these temporary uncertainties.

To deal with a temporary uncertainty, a couple of noisy discrete-time systems, sinusoidal signal system and electrical motor system, are assumed to have an uncertain model parameter. The state-space approach is commonly used when real physical systems and processes can be approximated with a reasonable number of states. The approximation implies model uncertainty that may cause an estimator to be biased and/or diverge. That is, due to concerns for model misspecification, there can be model uncertainty. From (1), the discrete-time state space model with the model uncertainty can be represented by

$$\begin{aligned} x_{i+1} &= (A+\Delta A)x_i + Bu_i + Gw_i, \\ z_i &= (C+\Delta C)x_i + v_i. \end{aligned} \tag{50}$$

Although FMS and IMS filters are designed by the nominal discrete-time state-space model (1), actual measured outputs for the estimation filtering are obtained from the uncertain system (50).

A. Sinusoidal Signal System

The noisy sinusoidal signal system is considered by

$$\begin{aligned} A &= \begin{bmatrix} cos(\pi/32) & sin(\pi/32) \\ -sin(\pi/32) & cos(\pi/32) \end{bmatrix}, \\ G &= \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ C = \begin{bmatrix} 1 \ 0 \end{bmatrix}, \end{aligned}$$

where its model uncertainty is assumed by

$$\Delta A = \begin{bmatrix} \delta_i & 0 \\ 0 & \delta_i \end{bmatrix}, \ \Delta C = \begin{bmatrix} 0.2\delta_i \ 0.2\delta_i \end{bmatrix},$$

with the uncertain model parameter δ_i as follows:

$$\delta_i = \begin{cases} 0.08 & \text{if } 100 \le i \le 150, \\ 0 & \text{otherwise.} \end{cases}$$

System and measurement noise covariances are taken by $Q = 0.05^2$ and $R = 0.05^2$, respectively.

In previous section, a heuristic to determine window length M was mentioned briefly. Before actual simulations, the $L2$ norm of error covariance matrix Ω_M^{-1} or Σ_M is computed from the error covariance equation (8) or (36) in order to determine the optimal window length M that can provide enough noise suppression. The $L2$ norm of error covariance matrix Ω_M^{-1} or Σ_M is obtained from the Matlab function `norm(X)`. This function returns the 2-norm or maximum singular value of matrix X, which can be also implemented approximately using another Matlab function `max(svd(X))`. The $L2$ norm of error covariance matrix Ω_M^{-1} or Σ_M is plotted according to increasing window lengths in Fig. 1. It can be seen that the $L2$ norm of error covariance matrix reduces as the window length grows and converges when the window length is around $M = 15$. Of course, the $L2$ norm of error covariance matrix can be more reduced when $M > 15$. However, in this case, the real-time application is somewhat difficult due to the computational load. Thus, the optimal window length can be taken by $M = 15$.

B. DC Motor System

The discretized direct current (DC) motor system is considered. The DC motor is the most commonly used electrical motor in the control systems due to their features such cost-efficiency, ease of use, high performance, longevity and quiet operation. The discretized DC motor system is modeled by

$$\begin{aligned} A &= \begin{bmatrix} 0.8178 & -0.0011 \\ 0.0563 & 0.3678 \end{bmatrix}, \ B = \begin{bmatrix} 0.1813 \\ 0.0069 \end{bmatrix}, \\ G &= \begin{bmatrix} 0.0006 & 0 \\ 0 & 0.0057 \end{bmatrix}, \ C = \begin{bmatrix} 1 \ 0 \end{bmatrix}, \end{aligned}$$

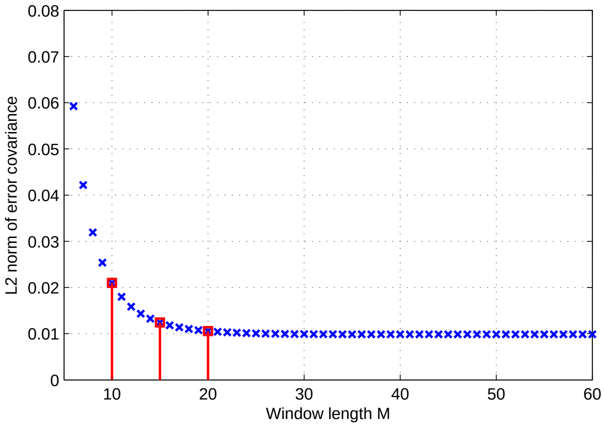

Fig. 1: $L2$ norms of error covariance matrix according to window length for sinusoid signal system.

where its model uncertainty is assumed by

$$\Delta A = \begin{bmatrix} \delta_i & 0 \\ 0 & \delta_i \end{bmatrix}, \ C = \begin{bmatrix} 0.2\delta_i & 0.2\delta_i \end{bmatrix},$$

with the uncertain model parameter δ_i as follows:

$$\delta_i = \begin{cases} 0.1 & \text{if } 150 \le i \le 300, \\ 0 & \text{otherwise.} \end{cases}$$

System and measurement noise covariances are taken by $Q = 0.02^2$ and $R = 0.03^2$, respectively. Before actual simulations, in the same way as in the sinusoidal signal example, the $L2$ norm of error covariance matrix Ω_M^{-1} or Σ_M is computed in order to determine the optimal window length M. It can be seen from Fig. 2 that the $L2$ norm of error covariance matrix reduces as the window length grows and converges when the window length is around $M = 20$. Of course, the L2 norm of error covariance matrix can be more reduced when $M > 20$. Thus, the optimal window length can be taken by $M = 20$.

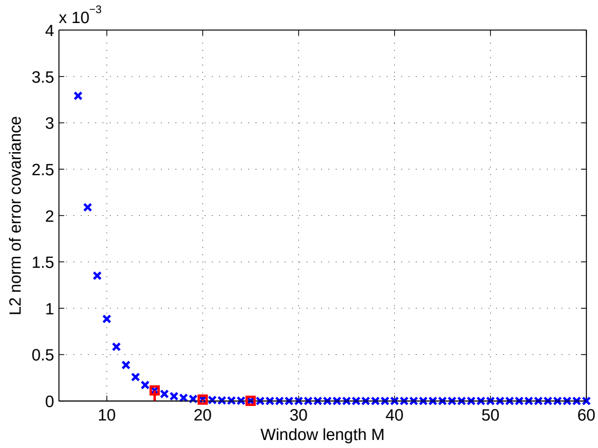

Fig. 2: $L2$ norms of error covariance matrix according to window length for DC motor system.

C. Simulations Results

The FMS filter and the recursive IMS filter are compared for the temporarily uncertain system. For consider diverse situations, three kinds of window lengths are set by $M = 10$, $M = 15$, and $M = 20$ for sinusoidal signal system and set by $M = 15$, $M = 20$, and $M = 25$ for DC motor system. Fig. 3 and 4 show simulation results according to diverse window lengths for two discrete-time systems. Left plots of figures show root-mean-square (RMS) estimation errors for 20 simulations. In addition, right plots of figures also show estimation errors for one of 20 simulations. The estimation error of the FMS filter is smaller than that of the recursive IMS filter on the interval where modeling uncertainty exist for all cases. In addition, the convergence time of estimation error is much shorter than that of the recursive IMS filter after temporary modeling uncertainty disappears. In addition, the FMS filter can be comparable to the recursive IMS filter after the effect of temporary modeling uncertainty completely disappears. Therefore, the FMS filter can be more robust than the recursive IMS filter when applied to temporarily uncertain systems, although the FMS filter is designed with no consideration for robustness. Moreover, it can be known that the larger window length may yield the longer convergence time of the estimation error, which can degrade the performance of the FMS filter. Therefore, it can be stated that the window length can be used as an effective design parameter to make the best possible performance of the FMS filter.

VI. Conclusion

This review paper has provided various expressions for an FMS filter, such as the iterative form, matrix form, summation form, and smoothing form, using various derivation methods for a noisy discrete-time state-space model. These diverse derivation methods and expressions can provide a comprehensive understanding of FMS filtering algorithms. Various factors that affect FMS filters have been evaluated. Through discussions about the choice of window length, it has been demonstrated that the window length can be considered a useful design parameter to improve the performance of the FMS filter as much as possible. Simulation results for the noisy sinusoidal signal model and the DC motor model have proven that the FMS filter is more suitable for temporarily uncertain systems than the IMS filter.

References

[1] M. Grewal, "Applications of Kalman filtering in aerospace 1960 to the present," *IEEE Control Systems Magazine*, vol. 30, no. 3, pp. 69–78, 2010.

[2] R. Faragher, "Understanding the basis of the Kalman filter via a simple and intuitive derivation," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 128–132, 2012.

[3] F. Auger, M. Hilairet, J. M. Guerrero, E. Monmasson, T. Orlowska-Kowalska, and S. Katsura, "Industrial applications of the Kalman filter: A review," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 12, pp. 5458–5471, 2013.

[4] M. Rhudy, R. Salguero, and K. Holappa, "A Kalman filtering tutorial for undergraduate students," *International Journal of Computer Science & Engineering Survey*, vol. 8, no. 1, pp. 1–18, 2017.

[5] M. Roth, G. Hendeby, C. Fritsche, and F. Gustafsson, "The ensemble Kalman filter: a signal processing perspective," *EURASIP Journal on Advances in Signal Processing*, vol. 2017, p. 56, 2017.

[6] A. Barrau and S. Bonnabel, "Invariant Kalman filtering," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, no. 1, pp. 237–257, 2018.

[7] T. O. Ting, K. L. Man, C.-U. Lei, and C. Lu, "Equivalent circuit model of lead-acid battery in energy storage power station and its state-of-charge estimation based on extended Kalman filtering method," *Engineering Letters*, vol. 26, no. 4, pp. 504–517, 2018.

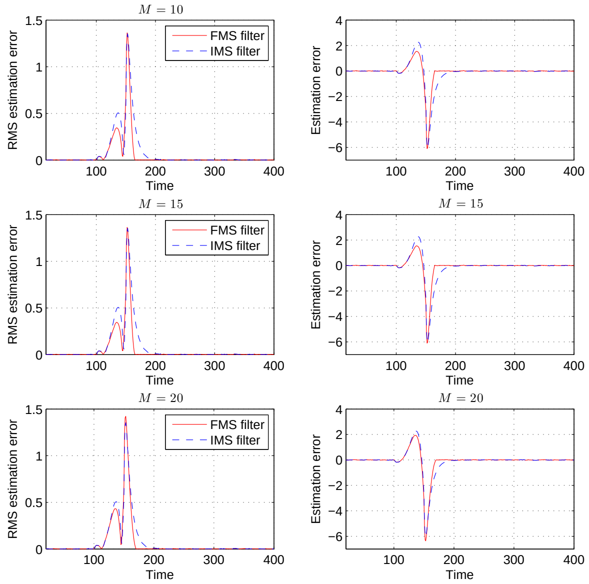

Fig. 3: Simulation results for sinusoid signal system.

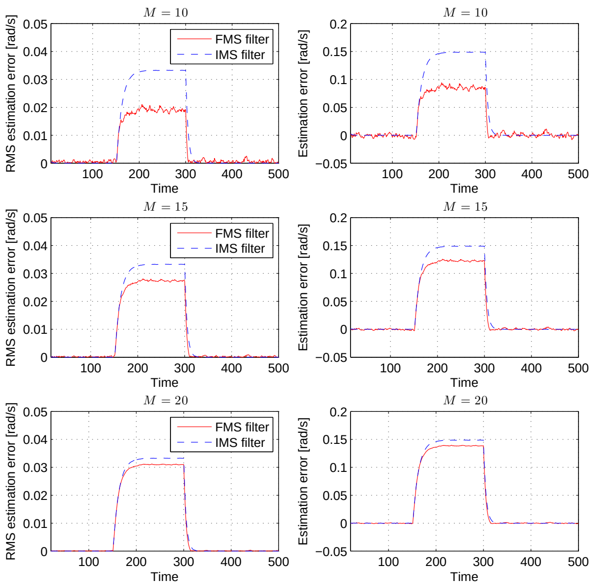

Fig. 4: Simulation results for DC motor system.

[8] M. Zhang, K. Li, B. Hu, and C. Meng, "Comparison of Kalman filters for inertial integrated navigation," *Sensors*, vol. 19, no. 6, p. 1426, 2019.

[9] M. Jahja, D. Farrow, R. Rosenfeld, and R. J. Tibshirani, "Kalman filter, sensor fusion, and constrained regression: equivalences and insights," *Advances in Neural Information Processing Systems*, vol. 32, pp. 13 187–13 196, 2019.

[10] C. Wei, X. Lian, and F. Yuan, "Estimation for constantinides-ingersol model with small levy noises from discrete observations," *IAENG International Journal of Applied Mathematics*, vol. 50, no. 1, pp. 1–6, 2020.

[11] X. Zeng and R. Ghanem, "Dynamics identification and forecasting of COVID-19 by switching Kalman filters," *Computation Mechanics*, vol. 66, pp. 1179–1193, 2020.

[12] Y. S. Shmaliy, S. Zhao, and C. K. Ahn, "Unbiased finite impulse response filtering: An iterative alternative to Kalman filtering ignoring noise and initial conditions," *IEEE Control Systems Magazine*, vol. 37, no. 5, pp. 70–89, 2017.

[13] Y. S. Shmaliy, Y. Neuvo, and S. Khan, "Review of unbiased FIR filters, smoothers, and predictors for polynomial signals," *Frontiers in Signal Processing*, vol. 2, no. 1, pp. 1–29, 2018.

[14] P. S. Kim and M. H. Kim, "Various forms of finite memory structure filter for discrete-time state-space model," in *Proc. International Conference on Artificial Intelligence in information and communication (ICAIIC 2020)*, Fukuoka, Japan, 2020, pp. 631–636.

[15] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Trans. Automat. Contr.*, vol. 44, no. 9, pp. 1787–1791, 1999.

[16] Y. S. Shmaliy and D. Simon, "Iterative unbiased FIR state estimation: A review of algorithms," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 113, 2013.

[17] W. H. Kwon, P. S. Kim, and S. H. Han, "A receding horizon unbiased FIR filter for discrete-time state space models," *Automatica*, vol. 38, no. 3, pp. 545–551, 2002.

[18] S. Zhao, Y. S. Shmaliy, B. Huang, and F. Liu, "Minimum variance unbiased FIR filter for discrete time-variant systems," *Automatica*, vol. 53, no. 2, pp. 355–361, 2015.

[19] P. S. Kim, "Maximum likelihood FIR filter for state space signal models," *IEICE Trans. Commun.*, vol. E85-B, no. 12, pp. 1604–1607, 2002.

[20] J. M. Pak, C. K. Ahn, Y. H. Mo, M. T. Lim, and M. K. Song, "Maximum likelihood FIR filter for visual object tracking," *Neurocomput.*, vol. 216, no. C, pp. 543–553, 2016.

[21] S. Zhao, B. Huang, and Y. Shmaliy, "Bayesian state estimation on finite horizons: The case of linear statespace model," *Automatica*, vol. 85, pp. 91–99, 2017.

[22] P. S. Kim, "Selective finite memory structure filtering using the chi-square test statistic for temporarily uncertain systems," *Applied Sciences*, vol. 9, no. 20, p. 4257, 2019.

[23] J. M. Pak, P. S. Kim, S. H. You, S. S. Lee, and M. K. Song, "Extended least square unbiased FIR filter for target tracking using the constant velocity motion model," *International Journal of Control, Automation and Systems*, vol. 15, no. 2, pp. 947–951, 2017.

[24] P. S. Kim, "An alternative FIR filter for state estimation in discrete-time systems," *Digital Signal Processing*, vol. 20, no. 3, pp. 935–943, 2010.

[25] P. S. Kim, "A finite memory structure smoother with recursive form using forgetting factor," *Mathematical Problems in Engineering*, vol. 2017, pp. 1–6, 2017.

[26] A. Vicente, S. Zhao, and B. Huang, "Forward-backward smoothers with finite impulse response structure," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–10, 2018.

[27] B.-K. Kwon, S. Han, and S. Han, "Improved receding horizon fourier analysis for quasi-periodic signals," *Journal of Electrical Engineering & Technology*, vol. 12, no. 1, pp. 378–384, 2017.

[28] Y. Shmaliy, S. H. Khan, S. Zhao, and O. Ibarra-Manzano, "General unbiased FIR filter with applications to GPS-based steering of oscillator frequency," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 3, pp. 1141–1148, 2017.

[29] P. S. Kim, E. H. Lee, M. S. Jang, and S. Y. Kang, "A finite memory structure filtering for indoor positioning in wireless sensor networks with measurement delay," *International Journal of Distributed Sensor Networks*, vol. 13, no. 1, pp. 1–8, 2017.

[30] S. Zhao, Y. Ma, and B. Huang, "Robust FIR state estimation of dynamic processes corrupted by outliers," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 1, pp. 139–147, 2018.

[31] Y. Xu, Y. S. Shmaliy, Y. Li, X. Chen, and H. Guo, "Indoor INS/LiDAR-based robot localization with improved robustness using cascaded FIR filter," *IEEE Access*, vol. 7, pp. 34 189–34 197, 2019.

[32] P. S. Kim, "Finite memory structure filtering and smoothing for target tracking in wireless network environments," *Applied Sciences*, vol. 9, no. 14, p. 2872, 2019.

[33] M. Vazquez-Olguin, Y. S. Shmaliy, and O. G. Ibarra-Manzano, "Distributed UFIR filtering over WSNs with consensus on estimates," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1645–1654, 2020.

[34] Y. Xu, Y. S. Shmaliy, C. K. Ahn, T. Shen, and Y. Zhuang, "IEEE Internet of Things Journal," *Tightly-coupled integration of INS and UWB using fixed-lag extended UFIR smoothing for quadrotor localization*, 2020.

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

Pyung Soo Kim He received the B.S. degree in Electrical Engineering from Inha University, Incheon, Korea, in 1994. He received the M.S. degree in Control and Instrumentation Engineering and the Ph.D. degree at the School of Electrical Engineering and Computer Science from Seoul National University, Seoul, Korea, in 1996 and 2001, respectively. From 2001 to 2005, he was a senior researcher at the Digital Media R&D Center of Samsung Electronics Co. Ltd. Since 2005, he has been a professor at Department of Electronics Engineering of Korea Polytechnic University, Shiheung, Korea. His main research interests are in the areas of system software solutions, statistical signal processing, wireless mobile networks, next generation network system design, and various industrial applications.