Embedded Unbiasedness: Effect on Optimal FIR Filtering Estimates

SNUNYI ZHAO
Jiangnan University
Institute of Automation
214122, Wuxi
P.R.CHINA

YURIY SHMALIY
Universidad de Guanajuato
Electronics Engineering Dept.
36885, Salamanca
MEXICO

SANOWAR KHAN
City University London
Electronics Engineering Dept.
EC1V 0HB, London
UK

Abstract: In this paper, we give an analysis of the embedded unbiasedness (EU) on optimal finite impulse response (OFIR) estimates. By minimizing the mean square error (MSE) constrained by the unbiasedness condition, a new OFIR-EU filter is derived. We show that the OFIR-EU filter does not require the initial conditions, and occupies an intermediate place between the UFIR and OFIR filters. It is also shown that the MSEs of the OFIR-EU and OFIR filters diminish with the estimation horizon. A numerical example has demonstrated that the OFIR-UE filter has better robustness against temporary model uncertainties than the OFIR and Kalman filters.

Key–Words: Optimal FIR filter, unbiasedness, state space, mean square error, Kalman filter

1 Introduction

The finite impulse response (FIR) filter is a device or algorithm that utilizes N most resent neighbouring measurements to obtain an estimate in real time. Compared to the infinite impulse response (IIR) filtering structures, the FIR filter exhibits some useful engineering features such as the bounded input/bounded output (BIBO) stability [1], robustness against temporary model uncertainties and round-off errors [2], and immunity to errors in the noise statistics [3].

Due to these advantages, the interest to FIR estimators has grown in last decades [4–27] . A significant progress was achieved in receding horizon FIR filtering. Kwon, Kim and Park have combined in [18] the receding horizon strategy with the Kalman filter (KF) strategy. An optimal FIR filter with embedded unbiasedness was proposed for discrete-time system model in [19], and a fixed-lag FIR smoother developed in [20] for continuous-time models. There were also important developments in real-time optimal FIR filtering.

The unbiased FIR (UFIR) filter was derived by Shmaliy in [21] for real-time state space models. Further, the p-shift optimal FIR (OFIR) estimator was obtained in [3] for time-variant state space model. Using in part the results obtained in [3], the Kalman-like UFIR estimator was derived in [7] for the time-variant case. Recently, the Kalman-like OFIR algorithm was proposed in [27] for time-invariant case. In [17], a suboptimal UFIR estimator was developed by using the extended KF strategy. Moreover, unified forms for KF and UFIR filter and smoother were shown and investigated in [22]. These results have opened new horizons in optimal and robust estimation of linear and nonlinear models [23–25].

In this paper, we investigate effect of the embedded unbiasedness on OFIR estimates. We derive the OFIR filter with embedded unbiasedness, consider mean square errors (MSEs) in different OFIR structures, and investigate the trade-off between OFIR filter with and without the embedded unbiasedness.

2 State-Space Model

Let us consider a linear discrete-time model given with the state and observation equations:

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{w}_k\,, \qquad (1)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{v}_k\,, \qquad (2)$$

in which k is the discrete time index, $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector, and $\mathbf{y}_k \in \mathbb{R}^p$ is the measurement vector. Matrices $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times u}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$ and $\mathbf{D} \in \mathbb{R}^{p\times v}$ are time-invariant and known. We suppose that the process noise $\mathbf{w}_k \in \mathbb{R}^u$ and the measurement noise $\mathbf{v}_k \in \mathbb{R}^v$ are zero mean, $E\{\mathbf{w}_k\} = \mathbf{0}$ and $E\{\mathbf{v}_k\} = \mathbf{0}$, mutually uncorrelated, and have arbitrary distributions and known covariances $\mathbf{Q}(i,j) = E\{\mathbf{w}_i\mathbf{w}_j^T\}$, $\mathbf{R}(i,j) = E\{\mathbf{v}_i\mathbf{v}_j^T\}$ for all i and j, to mean that $\mathbf{w}_k$ and $\mathbf{v}_k$ are not obligatorily white Gaussian.

Following [3], the model (1) and (2) can be extended on a horizon of N points $[l,k]$ as [28]

$$\begin{aligned}\mathbf{X}_{k,l} &= \mathbf{A}_{k-l}\mathbf{x}_l + \mathbf{B}_{k-l}\mathbf{W}_{k,l}\,, &(3)\\ \mathbf{Y}_{k,l} &= \mathbf{C}_{k-l}\mathbf{x}_l + \mathbf{H}_{k-l}\mathbf{W}_{k,l} + \mathbf{D}_{k-l}\mathbf{V}_{k,l}\,, &(4)\end{aligned}$$

where $l = k - N + 1$ is a start point of the averaging horizon. The time-variant state vector $\mathbf{X}_{k,l} \in \mathbb{R}^{Nn\times 1}$, observation vector $\mathbf{Y}_{k,l} \in \mathbb{R}^{Np\times 1}$, process noise vector $\mathbf{W}_{k,l} \in \mathbb{R}^{Nu\times 1}$, and observation noise vector $\mathbf{V}_{k,l} \in \mathbb{R}^{Nv\times 1}$ are specified as, respectively,

$$\begin{aligned}\mathbf{X}_{k,l} &= \left[\,\mathbf{x}_k^T\,\mathbf{x}_{k-1}^T\cdots\mathbf{x}_l^T\,\right]^T, &(5)\\ \mathbf{Y}_{k,l} &= \left[\,\mathbf{y}_k^T\,\mathbf{y}_{k-1}^T\cdots\mathbf{y}_l^T\,\right]^T, &(6)\\ \mathbf{W}_{k,l} &= \left[\,\mathbf{w}_k^T\,\mathbf{w}_{k-1}^T\cdots\mathbf{w}_l^T\,\right]^T, &(7)\\ \mathbf{V}_{k,l} &= \left[\,\mathbf{v}_k^T\,\mathbf{v}_{k-1}^T\cdots\mathbf{v}_l^T\,\right]^T. &(8)\end{aligned}$$

The extended model matrix $\mathbf{A}_{k-l} \in \mathbb{R}^{Nn\times n}$, process noise matrix $\mathbf{B}_{k-l} \in \mathbb{R}^{Nn\times Nu}$, observation matrix $\mathbf{C}_{k-l} \in \mathbb{R}^{Np\times n}$, auxiliary matrix $\mathbf{H}_{k-l} \in \mathbb{R}^{Np\times Nu}$, and measurement noise matrix $\mathbf{D}_{k-l} \in \mathbb{R}^{Np\times Nv}$ are all time-invariant and N-dependent. Model (1) and (2) suggests that these matrices can be written as, respectively

$$\mathbf{A}_i = \left[(\mathbf{A}^i)^T(\mathbf{A}^{i-1})^T\cdots\mathbf{A}^T\,\mathbf{I}\right]^T, \quad (9)$$

$$\mathbf{B}_i = \begin{bmatrix}\mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{i-1}\mathbf{B} & \mathbf{A}^i\mathbf{B}\\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{A}^{i-2}\mathbf{B} & \mathbf{A}^{i-1}\mathbf{B}\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B} & \mathbf{AB}\\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B}\end{bmatrix}, \quad (10)$$

$$\begin{aligned}\mathbf{C}_i &= \bar{\mathbf{C}}_i\mathbf{A}_i\,, &(11)\\ \mathbf{H}_i &= \bar{\mathbf{C}}_i\mathbf{B}_i\,, &(12)\\ \mathbf{D}_i &= \mathbf{diag}(\underbrace{\mathbf{D}\,\mathbf{D}\cdots\mathbf{D}}_{\mathbf{i+1}})\,, &(13)\\ \bar{\mathbf{C}}_i &= \mathbf{diag}(\underbrace{\mathbf{C}\,\mathbf{C}\cdots\mathbf{C}}_{\mathbf{i+1}})\,. &(14)\end{aligned}$$

For mode detail, see [27].

The FIR filtering estimate can be written as [2]

$$\hat{\mathbf{x}}_{k|k} = \mathbf{K}_k\mathbf{Y}_{k,l}\,, \quad (15)$$

where $\hat{\mathbf{x}}_{k|k}$ is the estimate, and $\mathbf{K}_k$ is the FIR filter gain determined using a given cost criterion. A distinctive difference between the FIR filter and KF is that the latter requires only one nearest past measurement to produce the estimate, while the convolution-based batch FIR filter requires N most recent measurements.

The estimate (15) will be unbiased if to obey the following unbiasedness condition,

$$E\{\mathbf{x}_k\} = E\{\hat{\mathbf{x}}_{k|k}\}\,, \quad (16)$$

in which $\mathbf{x}_k$ can be specified as

$$\mathbf{x}_k = \mathbf{A}^{N-1}\mathbf{x}_l + \bar{\mathbf{B}}_{k-l}\mathbf{W}_{k,l} \quad (17)$$

if to combine (3) and (4). Here $\bar{\mathbf{B}}_{k-l}$ is the first vector row in $\mathbf{B}_{k-l}$. By substituting (15) and (17) into (16), replacing the term $\mathbf{Y}_{k,l}$ with (4), and providing the averaging, one arrives at the unbiasedness constraint

$$\mathbf{A}^{N-1} = \mathbf{K}_k\mathbf{C}_{k-l} \quad (18)$$

or the deadbeat constraint. Provided $\hat{\mathbf{x}}_{k|k}$, the instantaneous estimation error $\mathbf{e}_k$ can be defined as

$$\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k}\,. \quad (19)$$

Given (1) and (2), below we derive an OFIR filter with embedded unbiasedness (EU), called OFIR-EU, by minimizing the variance as

$$\begin{aligned}\mathbf{K}_k^{\mathrm{OEU}} &= \arg\min_{\mathbf{K}_k} E\{\mathbf{e}_k\mathbf{e}_k^T\} \quad (20)\\ &\quad \text{subject to } (18)\,.\end{aligned}$$

In the next section, we derive and analyze the OFIR filter with embedded unbiasedness.

3 OFIR-EU Filter

The following lemma will be used to derive the OFIR-EU filter.

Lemma 1 *The trace optimization problem is given by*

$$\begin{aligned}&\arg\min_{\mathbf{K}}\ \mathrm{tr}\left[(\mathbf{KF}-\mathbf{G})\mathbf{H}(\mathbf{KF}-\mathbf{G})^T\right.\\ &\left.+(\mathbf{KL}-\mathbf{M})\mathbf{P}(\mathbf{KL}-\mathbf{M})^T + \mathbf{KSK}^T\right], \quad (21)\\ &\textit{subject to } \mathfrak{L}_{\{\mathbf{KU}=\mathbf{Z}\}|\theta}\end{aligned}$$

where $\mathbf{H} = \mathbf{H}^T > \mathbf{0}$, $\mathbf{P} = \mathbf{P}^T > \mathbf{0}$, $\mathbf{S} = \mathbf{S}^T > \mathbf{0}$, $\mathrm{tr}\,\mathbf{M}$ *is the trace of* $\mathbf{M}$, θ *denotes the constraint indication parameter such that* $\theta = 1$ *if the constraint exists and* $\theta = 0$ *otherwise. Here,* $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, $\mathbf{L}$, $\mathbf{M}$, $\mathbf{P}$, $\mathbf{S}$, $\mathbf{U}$, *and* $\mathbf{Z}$ *are constant matrices of appropriate dimensions. The solution to (21) is*

$$\mathbf{K} = \begin{bmatrix}\mathbf{Z}\\ \mathbf{G}\\ \mathbf{M}\end{bmatrix}^T \begin{bmatrix}\theta\left(\mathbf{U}^T\boldsymbol{\Xi}^{-1}\mathbf{U}\right)^{-1}\mathbf{U}^T\boldsymbol{\Xi}^{-1}\\ \mathbf{HF}^T\boldsymbol{\Xi}^{-1}\boldsymbol{\Pi}\\ \mathbf{PL}^T\boldsymbol{\Xi}^{-1}\boldsymbol{\Pi}\end{bmatrix}, \quad (22)$$

where $\mathbf{\Pi} = \mathbf{I} - \theta\mathbf{U}\big(\mathbf{U}^T\mathbf{\Xi}^{-1}\mathbf{U}\big)^{-1}\mathbf{U}^T\mathbf{\Xi}^{-1}$ *and*

$$\mathbf{\Xi} = \begin{cases} \mathbf{LPL}^T + \mathbf{S}, & \textit{if } \mathbf{F} = \mathbf{U},\ \mathbf{G} = \mathbf{Z},\ \textit{and } \theta = 1 \\ \mathbf{FHF}^T + \mathbf{S}, & \textit{if } \mathbf{L} = \mathbf{U},\ \mathbf{M} = \mathbf{Z},\ \textit{and } \theta = 1 \\ \mathbf{FHF}^T + \mathbf{LPL}^T + \mathbf{S}, & \textit{if } \theta = 0 \end{cases} \quad (23)$$

Proof: Represent the performance criterion in (21) as

$$\phi = \mathrm{tr}\,[(\mathbf{KF}-\mathbf{G})\mathbf{H}(\mathbf{KF}-\mathbf{G})^T + (\mathbf{KL}-\mathbf{M})\mathbf{P} \times(\mathbf{KL}-\mathbf{M})^T + \mathbf{KSK}^T]\,. \quad (24)$$

By partitioning $\mathbf{K}$ as $\mathbf{K}^T = [\mathbf{k}_1\mathbf{k}_2\cdots\mathbf{k}_m]$, where m is the dimension of $\mathbf{K}$, rewrite ϕ as

$$\phi = \sum_{i=1}^{m}\phi_i\,, \quad (25)$$

in which

$$\phi_i = \left(\mathbf{k}_i^T\mathbf{F} - \mathbf{g}_i^T\right)\mathbf{H}(\cdots)^T + \left(\mathbf{k}_i^T\mathbf{L} - \mathbf{m}_i^T\right) \times\mathbf{P}(\cdots)^T + \mathbf{k}_i^T\mathbf{S}\mathbf{k}_i\,, \quad (26)$$

where $\mathbf{g}_i$ and $\mathbf{m}_i$ are the ith column vector of $\mathbf{G}$ and $\mathbf{M}$, respectively, and $i = 1,2,\ldots,m$. Reasoning similarly, the ith constraint can be specified by

$$\mathfrak{L}^i_{\{\mathbf{U}^T\mathbf{k}_i=\mathbf{z}_i\}|\theta} = \begin{cases} \mathbf{U}^T\mathbf{k}_i = \mathbf{z}_i, & \text{if } \theta = 1 \\ \mathbf{0}, & \text{if } \theta = 0 \end{cases}\,. \quad (27)$$

Now note that ϕ_i and $\mathfrak{L}^i_{\{\mathbf{U}^T\mathbf{k}_i=\mathbf{z}_i\}|\theta}$ are independent on $\mathbf{k}_j$, $j \neq i$, and the problem (21) can be reduced to m independent optimization problems as

$$\min_{\mathbf{k}_i}\phi_i \quad \text{subject to} \quad \mathfrak{L}^i_{\{\mathbf{U}^T\mathbf{k}_i=\mathbf{z}_i\}|\theta}\,, \quad (28)$$

where $i = 1,2,\ldots,m$. Now, define $\varphi_{i|\theta}$ as

$$\varphi_{i|\theta} = \phi_i + \theta\lambda_i^T(\mathbf{U}^T\mathbf{k}_i - \mathbf{z}_i)\,, \quad (29)$$

where λ_i denotes the ith vector of the Lagrange multiplier. Note that $\varphi_{i|\theta}$ depends on θ which governs the existing of constraint. Setting $\theta = 1$, first consider a general case of $\mathbf{F} \neq \mathbf{U}$, $\mathbf{L} \neq \mathbf{U}$, $\mathbf{G} \neq \mathbf{Z}$ and $\mathbf{M} \neq \mathbf{Z}$ which is denoted as case (a). Taking the derivative of $\varphi_{i|a}$ with respect to $\mathbf{k}_i$ and λ_i respectively and making them equal to zero lead to

$$\frac{\partial\varphi_{i|a}}{\partial\mathbf{k}_{i|a}} = 2\mathbf{\Xi}_a\mathbf{k}_{i|a} - 2\left(\mathbf{FHg}_i + \mathbf{LPm}_i\right) + \mathbf{U}\lambda_i = 0\,, \quad (30)$$

which can further be rewritten as

$$\mathbf{k}_{i|a} = \mathbf{\Xi}_a^{-1}(\mathbf{FHg}_i + \mathbf{LPm}_i - 0.5\mathbf{U}\lambda_i)\,. \quad (31)$$

where $\mathbf{\Xi}_a \triangleq \mathbf{FHF}^T + \mathbf{LPL}^T + \mathbf{S}$, $\mathbf{H} > 0$, $\mathbf{P} > 0$, and $\mathbf{S} > 0$. By multiplying the both sides of (31) with $\mathbf{U}^T$ from the left-hand side, using the constraint (27), and arranging the terms, arrive at

$$\lambda_i = 2(\mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{U})^{-1}(\mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{FHg}_i + \mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{LPm}_i - \mathbf{z}_i)\,. \quad (32)$$

Substituting (32) into (31) and taking into account that $\mathbf{H} = \mathbf{H}^T$, $\mathbf{P} = \mathbf{P}^T$, $\mathbf{S} = \mathbf{S}^T$ and $\mathbf{\Xi}_a = \mathbf{\Xi}_a^T$, transforms $\mathbf{k}_i^T$ to

$$\mathbf{k}_{i|a}^T = (\mathbf{g}_i^T\mathbf{HF}^T + \mathbf{m}_i^T\mathbf{PL}^T)\mathbf{\Xi}_a^{-1} + [\mathbf{z}_i^T - (\mathbf{g}_i^T\mathbf{HF}^T + \mathbf{m}_i^T\mathbf{PL}^T) \times\mathbf{\Xi}_a^{-1}\mathbf{U}](\mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_a^{-1}\,. \quad (33)$$

At this point, reconstruct $\mathbf{K}_a$ as

$$\mathbf{K}_a = (\mathbf{GHF}^T\mathbf{\Xi}_a^{-1} + \mathbf{MPL}^T\mathbf{\Xi}_a^{-1}) \times(\mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_a^{-1}) + \mathbf{Z}(\mathbf{U}^T\mathbf{\Xi}_a^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_a^{-1}\,. \quad (34)$$

In the case of $\theta = 1$, $\mathbf{F} = \mathbf{U}$ and $\mathbf{H} = \mathbf{Z}$ which is denoted as case (b) or $\theta = 1$, $\mathbf{G} = \mathbf{U}$ and $\mathbf{M} = \mathbf{Z}$ which is denoted as case (c), the solutions can be obtained similarly to case (a), respectively,

$$\mathbf{K}_b = \mathbf{MPL}^T\mathbf{\Xi}_b^{-1}(\mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{\Xi}_b^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_b^{-1}) + \mathbf{Z}(\mathbf{U}^T\mathbf{\Xi}_b^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_b^{-1}\,, \quad (35)$$

$$\mathbf{K}_c = \mathbf{GHF}^T\mathbf{\Xi}_c^{-1}(\mathbf{I} - \mathbf{U}(\mathbf{U}^T\mathbf{\Xi}_c^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_c^{-1}) + \mathbf{Z}(\mathbf{U}^T\mathbf{\Xi}_c^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}_c^{-1}\,, \quad (36)$$

with

$$\mathbf{\Xi}_b = \mathbf{LPL}^T + \mathbf{S}\,, \quad (37)$$

$$\mathbf{\Xi}_c = \mathbf{FHF}^T + \mathbf{S}\,. \quad (38)$$

Note that (35) and (36) are equal to the results found in [2] for the receding horizon FIR filtering via prediction state model.

In the case of $\theta = 0$ which is denoted as case (d), the derivative of $\varphi_{i|d}$ with respect to $\mathbf{k}_{i|d}$ becomes

$$\frac{\partial\varphi_{i|d}}{\partial\mathbf{k}_{i|d}} = 2\mathbf{\Xi}_d\mathbf{k}_{i|d} - 2\left(\mathbf{FHg}_i + \mathbf{LPm}_i\right) = 0\,, \quad (39)$$

where $\mathbf{\Xi}_d = \mathbf{\Xi}_a$, and yields

$$\mathbf{k}_{i|d}^T = (\mathbf{g}_i^T\mathbf{HF}^T + \mathbf{m}_i^T\mathbf{PL}^T)\mathbf{\Xi}_d^{-1}\,. \quad (40)$$

Then $\mathbf{K}_d$ can be found to be

$$\mathbf{K}_d = (\mathbf{GHF}^T + \mathbf{MPL}^T)\mathbf{\Xi}_d^{-1}. \quad (41)$$

Finally, by observing that

$$\begin{aligned}\mathbf{GHF}^T\mathbf{\Xi}^{-1}(\mathbf{I}-\mathbf{U}(\mathbf{U}^T\mathbf{\Xi}^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}^{-1}) &= \mathbf{0}\,,\\ \mathbf{MPL}^T\mathbf{\Xi}^{-1}(\mathbf{I}-\mathbf{U}(\mathbf{U}^T\mathbf{\Xi}^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}^{-1}) &= \mathbf{0}\,,\end{aligned}$$

when $\mathbf{F}=\mathbf{U}$ and $\mathbf{L}=\mathbf{U}$, and using θ as an indicating parameter of the constraint, matrices $\mathbf{K}_a$, $\mathbf{K}_b$, $\mathbf{K}_c$, and $\mathbf{K}_d$ can be unified with

$$\begin{aligned}\mathbf{K} &= (\mathbf{GHF}^T\mathbf{\Xi}^{-1}+\mathbf{MPL}^T\mathbf{\Xi}^{-1})\\ &\quad\times(\mathbf{I}-\theta\mathbf{U}(\mathbf{U}^T\mathbf{\Xi}^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}^{-1})\\ &\quad+\mathbf{Z}\theta(\mathbf{U}^T\mathbf{\Xi}^{-1}\mathbf{U})^{-1}\mathbf{U}^T\mathbf{\Xi}^{-1}\,, \qquad (42)\end{aligned}$$

where $\mathbf{\Xi}$ is specified by (23). An equivalent form of (42) is (22) and the proof is complete.

3.1 The Gain for OFIR-EU Filter

Using the trace operation, the optimization problem (20) can be rewritten as

$$\begin{aligned}\mathbf{K}_k^{\mathrm{OEU}} &= \arg\min_{\mathbf{K_k}} E\left\{\mathrm{tr}\left[\mathbf{e}_k\mathbf{e}_k^T\right]\right\}\\ &= \arg\min_{\mathbf{K}_k} E\left\{\mathrm{tr}\left[\left(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k}\right)(\cdots)^T\right]\right\} \qquad (43)\end{aligned}$$

subject to (18), where $(\cdots)$ denotes the term that is equal to the relevant preceding term. By substituting $\mathbf{x}_k$ with (17) and $\hat{\mathbf{x}}_{k|k}$ with (15), the cost function becomes

$$\begin{aligned}\mathbf{K}_k^{\mathrm{OEU}} &= \arg\min_{\mathbf{K}_k} E\left\{\mathrm{tr}\left[\left(\mathbf{A}^{N-1}\mathbf{x}_l+\bar{\mathbf{B}}_{k-l}\mathbf{W}_{k,l}\right.\right.\right.\\ &\quad\left.\left.\left.-\ \mathbf{K}_k\mathbf{Y}_{k,l}\right)(\cdots)^T\right]\right\}\,. \qquad (44)\end{aligned}$$

If we take into account constraint (18), provide the averaging, and rearrange the terms, (44) can be transformed to

$$\begin{aligned}\mathbf{K}_k^{\mathrm{OEU}} &= \arg\min_{\mathbf{K}_k} E\left\{\mathrm{tr}\left[\left(\bar{\mathbf{B}}_{k-l}\mathbf{W}_{k,l}\right.\right.\right.\\ &\quad\left.\left.\left.-\mathbf{K}_k\left(\mathbf{H}_{k-l}\mathbf{W}_{k,l}+\mathbf{D}_{k-l}\mathbf{V}_{k,l}\right)\right)(\cdots)^T\right]\right\}\\ &= \arg\min_{\mathbf{K}_k} E\left\{\mathrm{tr}\left[\left(\left(\mathbf{K}_k\mathbf{H}_{k-l}-\bar{\mathbf{B}}_{k-l}\right)\mathbf{W}_{k,l}\right.\right.\right.\\ &\quad\left.\left.\left.+\mathbf{K}_k\mathbf{D}_{k-l}\mathbf{V}_{k,l}\right)(\cdots)^T\right]\right\}\\ &= \arg\min_{\mathbf{K}_k} \mathrm{tr}\left[\left(\mathbf{K}_k\mathbf{H}_{k-l}-\bar{\mathbf{B}}_{k-l}\right)\mathbf{\Theta}_w(\cdots)^T\right.\\ &\quad\left.+\mathbf{K}_k\mathbf{\Delta}_v\mathbf{K}_k^T\right]\,, \qquad (45)\end{aligned}$$

where the fact is invoked that the system noise vector $\mathbf{W}_{k,l}$ and the measurement noise vector $\mathbf{V}_{k,l}$ are pairwise independent. The auxiliary matrices are

$$\mathbf{\Theta}_w = E\left\{\mathbf{W}_{k,l}\mathbf{W}_{k,l}^T\right\}\,, \qquad (46)$$

$$\mathbf{\Delta}_v = \mathbf{D}_{k-l}E\left\{\mathbf{V}_{k,l}\mathbf{V}_{k,l}^T\right\}\mathbf{D}_{k-l}^T\,. \qquad (47)$$

Referring to Lemma 1 with $\theta=1$, the solution to the optimization problem (45) can be obtained by neglecting $\mathbf{L}$, $\mathbf{M}$, and $\mathbf{P}$ and using the replacements: $\mathbf{F}\leftarrow\mathbf{H}_{k-l}$, $\mathbf{G}\leftarrow\bar{\mathbf{B}}_{k-l}$, $\mathbf{H}\leftarrow\mathbf{\Theta}_w$, $\mathbf{U}\leftarrow\mathbf{C}_{k-l}$, $\mathbf{Z}\leftarrow\mathbf{A}^{N-1}$, and $\mathbf{S}\leftarrow\mathbf{\Delta}_v$. We thus have

$$\mathbf{K}_k^{\mathrm{OEU}} = \mathbf{K}_k^{\mathrm{OEUa}}+\mathbf{K}_k^{\mathrm{OEUb}}\,, \qquad (48)$$

where

$$\mathbf{K}_k^{\mathrm{OEUa}} = \mathbf{A}^{N-1}(\mathbf{C}_{k-l}^T\mathbf{\Delta}_{w+v}^{-1}\mathbf{C}_{k-l})^{-1}\mathbf{C}_{k-l}^T\mathbf{\Delta}_{w+v}^{-1}\,, \qquad (49)$$

$$\mathbf{K}_k^{\mathrm{OEUb}} = \bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\mathbf{\Delta}_{w+v}^{-1}(\mathbf{I}-\mathbf{\Omega}_{k-l})\,, \qquad (50)$$

in which

$$\mathbf{\Omega}_{k-l} = \mathbf{C}_{k-l}(\mathbf{C}_{k-l}^T\mathbf{\Delta}_{w+v}^{-1}\mathbf{C}_{k-l})^{-1}\mathbf{C}_{k-l}^T\mathbf{\Delta}_{w+v}^{-1}\,, \qquad (51)$$

$$\mathbf{\Delta}_{w+v} = \mathbf{\Delta}_w+\mathbf{\Delta}_v\,, \qquad (52)$$

$$\mathbf{\Delta}_w = \mathbf{H}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\,. \qquad (53)$$

The OFIR-EU filter structure can now be summarized in the following theorem.

Theorem 2 *Given the discrete time-invariant state space model (1) and (2) with zero mean mutually independent and uncorrelated noise vectors* $\mathbf{w}_k$ *and* $\mathbf{v}_k$*, the OFIR-EU filter utilizing measurements from* l *to* k *is stated by*

$$\hat{\mathbf{x}}_{k|k} = \mathbf{K}_k^{\mathrm{OEU}}\mathbf{Y}_{k,l}\,, \qquad (54)$$

where $\mathbf{K}_k^{\mathrm{OEU}}=\mathbf{K}_k^{\mathrm{OEUa}}+\mathbf{K}_k^{\mathrm{OEUb}}$, $\mathbf{Y}_{k,l}\in\mathbb{R}^{Np\times 1}$ *is the measurement vector given by (6), and* $\mathbf{K}_k^{\mathrm{OEUa}}$ *and* $\mathbf{K}_k^{\mathrm{OEUb}}$ *are given by (49) and (50) with* $\mathbf{C}_{k-l}$ *and* $\bar{\mathbf{B}}_{k-l}$ *specified by (11) and the first row vector of (10), respectively.*

Proof: The proof is provided by (43)-(53).

Note that the horizon length N for (54) should be chosen such that the first inverse in (49) exists. In general, N can be set as $N\geqslant n$, where n is the number of the model states.

4 Estimation Errors

In what follows, we investigate MSEs in the OFIR and OFIR-EU filters.

4.1 Mean Square Errors

The MSE $\mathbf{J}_k$ at time k can be calculated as

$$\begin{aligned}\mathbf{J}_k &= E\left\{\mathbf{e}_k\mathbf{e}_k\right\}\\ &= E\left\{\left(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k}\right)\left(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k}\right)^T\right\}\\ &= E\{\mathbf{x}_k\mathbf{x}_k^T\}+E\{\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k|k}^T\}-2E\{\mathbf{x}_k\hat{\mathbf{x}}_{k|k}^T\}\,, \qquad (55)\end{aligned}$$

which means that the MSE can be decomposed via the squared bias and variance. Assuming that the $\mathbf{x}_k$ is unbiased, we write $E\{\mathbf{x}_k\mathbf{x}_k^T\} = \mathrm{Var}(\mathbf{x}_k)$ and $E\{\hat{\mathbf{x}}_{k|k}\hat{\mathbf{x}}_{k|k}^T\} = \mathrm{Bias}^2(\hat{\mathbf{x}}_{k|k}) + \mathrm{Var}(\hat{\mathbf{x}}_{k|k})$. We further decompose the estimate $\hat{\mathbf{x}}_{k|k}$ as $\hat{\mathbf{x}}_{k|k} = \mathrm{Bias}(\hat{\mathbf{x}}_{k|k}) + \tilde{\mathbf{x}}_{k|k}$, where $\tilde{\mathbf{x}}_{k|k}$ is a random part of $\hat{\mathbf{x}}_{k|k}$, get

$$\begin{aligned} E\{\mathbf{x}_k\hat{\mathbf{x}}_{k|k}^T\} &= E\{\mathbf{x}_k[\mathrm{Bias}(\hat{\mathbf{x}}_{k|k}) + \tilde{\mathbf{x}}_{k|k}]^T\} \\ &= E\{\mathbf{x}_k\}\mathrm{Bias}^T(\hat{\mathbf{x}}_{k|k}) + E\{\mathbf{x}_k\tilde{\mathbf{x}}_{k|k}^T\} \\ &= \mathrm{Cov}(\mathbf{x}_k\hat{\mathbf{x}}_{k|k}^T) \end{aligned}$$

and change (55) to

$$\begin{aligned} \mathbf{J}_k = {} & \mathrm{Bias}^2\left(\hat{\mathbf{x}}_{k|k}\right) + \mathrm{Var}\left(\mathbf{x}_k\right) + \mathrm{Var}\left(\hat{\mathbf{x}}_{k|k}\right) \\ & -2\mathrm{Cov}\left(\mathbf{x}_k, \hat{\mathbf{x}}_{k|k}\right), \end{aligned} \tag{56}$$

where the state variance $\mathrm{Var}\left(\mathbf{x}_k\right)$ is specified by

$$\mathrm{Var}\left(\mathbf{x}_k\right) = \bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\bar{\mathbf{B}}_{k-l}^T \tag{57}$$

and, for unbiased estimate, we have

$$\mathrm{Bias}\left(\hat{\mathbf{x}}_{k|k}\right) = \mathbf{0}\,. \tag{58}$$

Based upon (55), below we specify the MSEs for the above considered FIR filters. Accordingly, the MSE in the UFIR filter becomes

$$\begin{aligned} \mathbf{J}_k^{\mathrm{U}} = {} & \bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\bar{\mathbf{B}}_{k-l}^T + \mathbf{K}_k^{\mathrm{U}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{U}}\right)^T \\ & -2\bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\left(\mathbf{K}_k^{\mathrm{U}}\right)^T, \end{aligned} \tag{59}$$

where $\mathbf{K}_k^{\mathrm{U}}$ the filter gain of the UFIR filter.

4.2 MSE in the OFIR-EU Estimate

In the OFIR-EU filter, $\mathrm{Var}(\hat{\mathbf{x}}_{k|k})$ and $\mathrm{Cov}(\mathbf{x}_k, \hat{\mathbf{x}}_{k|k})$ are given by, respectively,

$$\mathrm{Var}(\hat{\mathbf{x}}_{k|k}) = \mathbf{K}_k^{\mathrm{OEU}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{OEU}}\right)^T, \tag{60}$$

$$\mathrm{Cov}(\mathbf{x}_k, \hat{\mathbf{x}}_{k|k}) = \bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\left(\mathbf{K}_k^{\mathrm{OEU}}\right)^T. \tag{61}$$

From (48) we have $\mathbf{K}_k^{\mathrm{OEU}} = \mathbf{K}_k^{\mathrm{U}} + \mathbf{K}_k^{\mathrm{b}}$ and get

$$\begin{aligned} \mathrm{Var}(\hat{\mathbf{x}}_{k|k}) = {} & \mathbf{K}_k^{\mathrm{U}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{U}}\right)^T \\ & +2\mathbf{K}_k^{\mathrm{U}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{b}}\right)^T \\ & +\mathbf{K}_k^{\mathrm{b}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{b}}\right)^T, \end{aligned} \tag{62}$$

$$\begin{aligned} \mathrm{Cov}(\mathbf{x}_k, \hat{\mathbf{x}}_{k|k}) = {} & \bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\left(\mathbf{K}_k^{\mathrm{U}}\right)^T \\ & +\bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T\left(\mathbf{K}_k^{\mathrm{b}}\right)^T. \end{aligned} \tag{63}$$

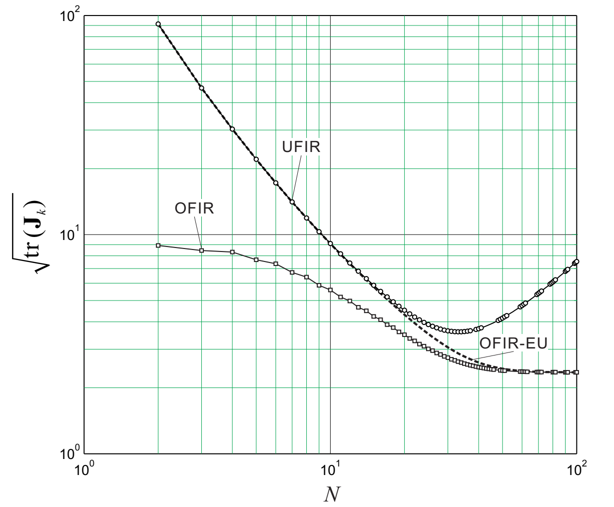

Figure 1: Typical RMSEs as functions of N for different filters with $\sigma_w^2 = 1$.

Next, substituting (61), (62) and (63) into (55) and rearranging the terms yield

$$\begin{aligned} \mathbf{J}_k^{\mathrm{OEU}} = {} & \mathbf{J}_k^{\mathrm{U}} + \mathbf{K}_k^{\mathrm{b}}\mathbf{\Delta}_{w+v}\left(\mathbf{K}_k^{\mathrm{b}}\right)^T - 2(\bar{\mathbf{B}}_{k-l}\mathbf{\Theta}_w\mathbf{H}_{k-l}^T \\ & -\mathbf{K}_k^{\mathrm{U}}\mathbf{\Delta}_{w+v})\left(\mathbf{K}_k^{\mathrm{b}}\right)^T, \end{aligned} \tag{64}$$

where $\mathbf{J}_k$ is the MSE of the UFIR filter.

5 Simulations

In this section, we show effect of the embedded unbiasedness on optimal estimates in more detail. In doing so, we run the UFIR, OFIR-EU, and OFIR filters in different noise environments using the two-state polynomial model specified with

$$\mathbf{A} = \begin{bmatrix} 1 & 0.05 \\ 0 & 1 \end{bmatrix},$$

$\mathbf{C} = [\,1\ 0\,]$, and $\mathbf{B}$ and $\mathbf{D}$ identity of proper dimensions.

5.1 Accurate Model – Ideal Case

In an ideal case, the model represents a process accurately and the noise statistics are known exactly. The goal then is to show effect of the horizon length N on the FIR estimates. We set the measurement noise variance as $\sigma_v^2 = 10$, and the initial states as $x_{10} = 1$ and $x_{20} = 0.01/s$. We then compute the root MSE (RMSE) of the estimate by $\mathrm{tr}\,\mathbf{J}_k$ as a function of N. The results are illustrated in Fig. 1 for $\sigma_w^2 = 1$ and in Fig. 2 for $\sigma_w^2 = 0.1$.

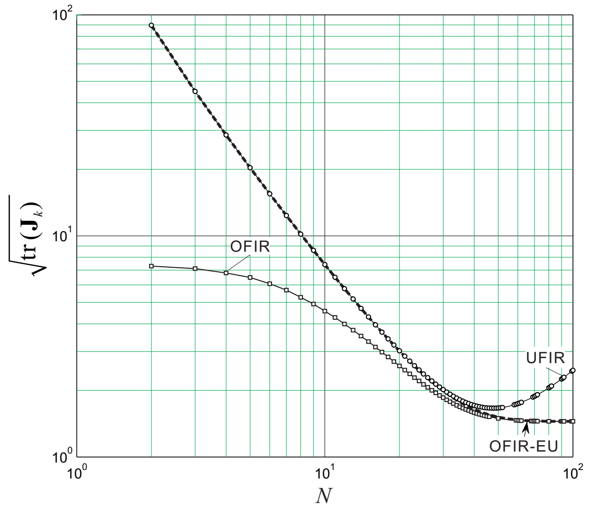

Figure 2: Typical RMSEs as functions of N for different filters with $\sigma_w^2 = 0.1$.

What we can see in Fig. 1 and Fig. 2 is that the MSE function of the UFIR filter is traditionally concave on N with a minimum at N_{opt} [26]: with $N < N_{\text{opt}}$, noise reduction is inefficient and, if $N > N_{\text{opt}}$, the bias error dominates. The following generalizations can also be made:

- The embedded unbiasedness puts the OFIR-EU filter error in between the UFIR and OFIR filters: the *OFIR-EU filter becomes* essentially the *UFIR filter when* $N < N_{\text{opt}}$ and the *OFIR filter if* $N > N_{\text{opt}}$.
- An increase in N_{opt} diminishes the error difference between the OFIR and UFIR filters (compare Fig. 1 with $N_{\text{opt}} = 33$ and Fig. 2 with $N_{\text{opt}} = 47$).
- Because MSEs in the *OFIR* and *OFIR-EU filters* reduce with N, these filters *are full-horizon* [3].

Referring to the fast that the ideal conditions are not the case in practice, we further investigate effect of temporary model uncertainties on the FIR estimates.

5.2 Filtering with Model Uncertainties

To learn effect of temporary model uncertainties on the filtering accuracy, we next set $\tau = 0.1$ s when $160 \leqslant k \leqslant 180$ and $\tau = 0.05$ s otherwise. The noise variances are allowed to be $\sigma_{w1}^2 = 1$, $\sigma_{w2}^2 = 1/\text{s}^2$, and $\sigma_v^2 = 10$. We also introduce a correction coefficient p and substitute the noise covariances with $p^2\mathbf{Q}_k$ and $\mathbf{R}_k/p^2$ in all the algorithms. The process is simulated at 400 subsequent points.

Typical estimates are sketched in Fig. 3. As can

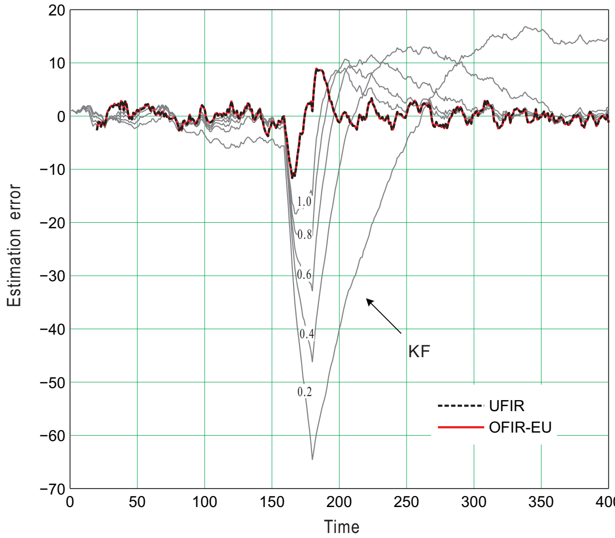

Figure 3: Instantaneous estimation errors caused by the temporary model uncertainties with $p < 1$ for the KF, UFIR filter, and OFIR-EU filter.

be seen, the OFIR-EU filter (case $p = 0.2$) and the UFIR filter produce almost equal errors and demonstrate good robustness against temporary model uncertainties. In contrast, the KF demonstrates much worse robustness for any $p \leqslant 1$ and we conclude that FIR filtering is more robust in real world than Kalman filtering.

6 Conclusions

Unbiasedness imbedded to the OFIR filter instills into it several useful properties. Unlike the OFIR filter, the OFIR-EU filter completely ignores the initial conditions. The OFIR-EU filter is equivalent to the MVU FIR filter. In terms of accuracy, the OFIR-EU filter is in between the UFIR and OFIR filters. Unlike in the UFIR filter which MSE is minimized by N_{opt}, MSEs in the OFIR-EU and OFIR filters diminish with N and these filters are thus full-horizon.

The performance of OFIR-EU filter is developed by varying the horizon N around N_{opt} or ranging the correction coefficient p around $p = 1$. Accordingly, the OFIR-EU filter in general demonstrates higher immunity against errors in the noise statistics and better robustness against temporary model uncertainties than the OFIR filter and KF.

This investigation was supported by the Royal Academy of Engineering under the Newton Research Collaboration Programme NRCP/1415/140.

References:

[1] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. (Academic, New York, 1970)

[2] W. H. Kwon and S. Han, *Receding Horizon Control: Model Predictive Control for State Models*. Springer: London, 2005.

[3] Y. S. Shmaliy and O. I. Manzano, "Time-variant linear optimal finite impulse response estimator for discrete state-space models," *Int. J. Adapt. Contrl Signal Process.*, vol. 26, no. 2, pp. 95–104, Sep. 2012.

[4] C. T. Mullis and R. A. Roberts, "Finite-memory problems and agorithms," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 440–445, Jul. 1974.

[5] K. V. Ling and K. W. Lim, "Receding horizon recursive state estimation," *IEEE Trans. Autom. Control*, vol. 44, no. 9, pp. 1750–1753, Sep. 1999.

[6] D. F. Crouse, P. Willet, and Y. Bar-Shalom, "A low-complexity sliding-window Kalman FIR smoother for discrete-time models," *IEEE Trans. Signal Process. Lett.*, vol. 17, no. 2, pp. 177–180, Feb. 2010.

[7] Y. S. Shmaliy, "An iterative Kalman-like algorithm ignoring noise and initial conditions," *IEEE Trans. Signal Process.*, vol. 59, pp. 2465–2473, 2011.

[8] L. Chang, "On Kalman-like finite impulse response filters," arXiv:1501.07132, 2015.

[9] S. Zhao, Y. S. Shmaliy, B. Huang, and F. Liu, "Minimum variance unbiased FIR filter for discrete time-variant systems, *Automatica*, vol.53, pp. 355–361, Mar. 2015.

[10] S. Zhao, Y. S. Shmaliy, and F. Liu, "Fast computation of discrete optimal FIR estimates in white Gaussian noise," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 718–722, 2015.

[11] J. Fu, J. Sun, S. Lu, and Y. Zhang, "Maneuvering target tracking with modified unbiased FIR filter," *J. Beijing Univ. Aeronautics and Astronautics*, vol. 41, no. 1, pp. 72–82, 2015.

[12] Y. Kou, Y. Jiao, D. Xu, M. Zhang, Y. Liu, X. Li, Low-cost precise measurement of oscillator frequency instability based on GNSS carrier observation, Advances in Space Research 51(6)(2013) 969–977.

[13] J. Yuan and J. A. Stuller, "Order-Recursive FIR smoothers," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1242–1246, May 1994.

[14] S. C. Pei and J. J. Shyu, "Fast design of 2-D linear phase complex FIR digital filters by analytical least squares methods," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 3157–3161, Dec. 1996.

[15] C. K. Ahn and P. S. Kim, "Fixed-lag maximum likelihood FIR smoother for state-space models," *IEICE Electron. Express*, vol. 5, no. 1, pp. 11–16, 2008.

[16] C. K. Ahn, "Strictly passive FIR filtering for state-space models with external disturbance," *Int. J. Electron. Commun.*, vol. 66, no. 11, pp. 944–948, 2012.

[17] Y. S. Shmaliy, "Suboptimal FIR filtering of nonlinear models in additive white Gaussian noise," *IEEE Trans. Signal Process.*, vol. 60, pp. 5519–5527, 2012.

[18] W. H. Kwon, P. S. Kim, and P. Park, "A receding horizon Kalman FIR filter for discrete time-invariant systems," *IEEE Trans. Autom. Control*, vol. 99, pp. 1787–1791, 1999.

[19] W. H. Kwon, P. S. Kim, and S. Han, "A receding horizon unbiased FIR filter for discrete-time state space models," *Automatica*, vol. 38, pp. 545–551, 2002.

[20] B. K. Kwon, S. Han, O. K. Kim, and W. H. Kwon, "Minimum variance FIR smoothers for discrete-time state space models," *IEEE Trans. Signal Process. Lett.*, vol. 14, pp. 557–560, 2007.

[21] Y. S. Shmaliy, "Unbiased FIR filtering of discrete-time polynomial state-space models," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1241–1249, Apr. 2009.

[22] D. Simon and Y. S. Shmaliy, "Unified forms for Kalman and finite impulse response filtering and smoothing," *Automatica*, vol. 49, pp. 1892–1899, 2013.

[23] P. K. Kitanidis, "Unbiased minimum-variance linear state estimation," *Automatica*, vol. 23, pp. 775–778, Nov. 1987.

[24] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems," *Automatica*, vol. 43, pp. 111–116, Jan. 2007.

[25] D. F. Crouse, P. Willett, and Y. Bar-Shalom, "A low-complexity sliding-windwon Kalman FIR smoother for discrete-time models," *IEEE Trans. Signal Process. Lett.*, vol. 17, pp. 177–180, 2010.

[26] F. Ramirez-Echeverria, A. Sarr, Y. S. Shmaliy, "Optimal memory for discrete-time FIR filters in state-space," *IEEE Trans. Signal Process.*, vol. 62, pp. 557–561, 2014.

[27] S. Zhao, Y. S. Shmaliy, and F. Liu, "Fast Kalman-like optimal unbiased FIR filtering with applications," *IEEE Trans. on Signal Process.* (to be published)

[28] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1994.