Unbiased, optimal, and in-betweens: the trade-off in discrete finite impulse response filtering

Shunyi Zhao[1,2], *Yuriy S. Shmaliy*[3] ✉, *Fei Liu*[1], *Sanowar H. Khan*[4]

[1]*Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, People's Republic of China*

[2]*Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, Canada*

[3]*Department of Electronics Engineering, Universidad de Guanajuato, Salamanca 36885, Mexico*

[4]*School of Mathematics, City University London, London, UK*

✉ *E-mail: shmaliy@ugto.mx*

Abstract: In this survey, the authors examine the trade-off between the unbiased, optimal, and in-between solutions in finite impulse response (FIR) filtering. Specifically, they refer to linear discrete real-time invariant state-space models with zero mean noise sources having arbitrary covariances (not obligatorily delta shaped) and distributions (not obligatorily Gaussian). They systematically analyse the following batch filtering algorithms: unbiased FIR (UFIR) subject to the unbiasedness condition, optimal FIR (OFIR) which minimises the mean square error (MSE), OFIR with embedded unbiasedness (EU) which minimises the MSE subject to the unbiasedness constraint, and optimal UFIR (OUFIR) which minimises the MSE in the UFIR estimate. Based on extensive investigations of the polynomial and harmonic models, the authors show that the OFIR-EU and OUFIR filters have higher immunity against errors in the noise statistics and better robustness against temporary model uncertainties than the OFIR and Kalman filters.

1 Introduction

From the beginning works by Gauss [1], unbiasedness plays a role of the necessary condition that is used to derive linear and non-linear estimators [2]. In statistics, the ordinary least squares (OLS) estimator proposed by Gauss in 1795 is an unbiased estimator. By the Gauss–Markov theorem [3], this estimator is also the best linear unbiased estimator [4] if noise is white and if it has the same variance at each time step [5]. The unbiasedness is obeyed by a condition $E\{\hat{\boldsymbol{x}}_k\} = E\{\boldsymbol{x}_k\}$ which means that the average of the estimate $\hat{\boldsymbol{x}}_k$ is equal to that of the model $\boldsymbol{x}_k$. It leads to the unbiased finite impulse response (UFIR) estimator [6, 7]. Of practical importance is that neither OLS nor UFIR require the noise statistics which are not always known to the engineer [8]. However, the unbiasedness condition does not guarantee 'good estimate' [9]. Therefore, the sufficient condition – the minimum noise variance – is often applied along to produce different kinds of estimators which are optimal in the minimum mean square error (MSE) sense or suboptimal: Bayesian, maximum likelihood, minimum variance unbiased, and so on. In recent decades, a new class of estimators having finite impulse response (FIR) (filters, smoothers, and predictors) has been developed to have optimal or suboptimal properties.

The FIR filter differs from the structures having the infinite impulse response (IIR) such as the Kalman filter (KF) and other recursive filters [10–13]: it utilises N discrete measurements over the most recent time interval (horizon). Accordingly, FIR filter exhibits some useful features such as the bounded input/bounded output stability [14], robustness against mismodelling as well as temporary model uncertainties and round-off errors [15], and better immunity against errors in the noise statistics [16]. Early results on optimal FIR (OFIR) theory can be found in [17–19]. At that time, FIR filters were not ones commonly used for state estimation due to large computation complexity. A tremendous progress in computation technology has changed the things and nowadays one can find efficient solutions on FIR filtering [20–24], smoothing [25–28], and prediction [29–31]. Many practical applications were also reported [32–35].

In state estimation, different models can be used. The prediction state model $\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{w}_k$ is basic in control [36, 37]. Here, $\boldsymbol{x}_k$ is the state vector, $\boldsymbol{w}_k$ is the system noise vector, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are some matrices. For this model, OFIR filtering has been developed by W. H. Kwon and his followers that has resulted in the theory of receding horizon control [21] and several key solutions [19–21, 25, 28, 38].

The real-time state model $\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{w}_k$ serves better when prediction is not required [39, 40]. Employing this model, the OFIR filter was derived by Shmaliy in [16] for time-invariant models. This filter was later extended to time-variant models in [41]. The unbiased (UFIR) filter and smoother were proposed in [27, 42] for polynomial systems. In [16], a p-shift UFIR estimator was derived as a special case of the p-shift OFIR filter. Soon after, the UFIR filter [16] was extended to time-variant systems [22]. For non-linear models, an extended UFIR filter was proposed in [43] and unified forms for FIR filtering and smoothing were discussed in [44]. An important advantage of the UFIR filter against OFIR filter is that the noise statistics are not required and noise reduction is provided by averaging. Therefore, if $N \gg 1$, the UFIR filter becomes as successful in accuracy as any optimal filter. Note that the FIR solutions derived from the prediction and real-time state-space models are convertible only if noise is white Gaussian and has the same variance at each time step.

In this survey, we examine and develop in part the batch OFIR, UFIR, OFIR with embedded unbiasedness (OFIR-EU), and optimal UFIR (OUFIR) filters. The trade-off between these filters is analysed systematically and we also learn their properties in a comparison to the KF. The most noticeable properties are that the OUFIR filter is equivalent to the OFIR-EU filter which occupies an intermediate position between the UFIR and OFIR filters in term of accuracy, and the MSEs of OFIR-EU and OFIR filters decrease with the increasing of estimation horizon N. The remaining parts of this paper are organised as follows. Section 2 describes the model and gives some preliminaries. Section 3 presents the UFIR filter. Section 4 presents the OFIR filter. The OFIR-EU filter is discussed in Section 5. Here, we also consider a unified form for different kinds of OFIR filters. Section 6 presents the OUFIR filter. The MSEs are provided and compared in Section 7, and simulations are provided based on the polynomial and harmonic models in Section 8. Finally, conclusions are made in Section 9.

2 State-space model and preliminaries

A linear discrete time-invariant state-space model can be described by

$$\boldsymbol{x}_k = \boldsymbol{A}\boldsymbol{x}_{k-1} + \boldsymbol{B}\boldsymbol{w}_k, \tag{1}$$

$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{D}\boldsymbol{v}_k, \tag{2}$$

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

Following [16], the models (1) and (2) can be represented in a batch form on a discrete time interval $[l, k]$ with recursively computed forward-in-time solutions as

$$\boldsymbol{X}_{k,l} = \boldsymbol{A}_{k-l}\boldsymbol{x}_l + \boldsymbol{B}_{k-l}\boldsymbol{W}_{k,l}, \tag{3}$$

$$\boldsymbol{Y}_{k,l} = \boldsymbol{C}_{k-l}\boldsymbol{x}_l + \boldsymbol{H}_{k-l}\boldsymbol{W}_{k,l} + \boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l}, \tag{4}$$

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

$$\boldsymbol{X}_{k,l} = \left[\boldsymbol{x}_k^{\mathrm{T}}\ \boldsymbol{x}_{k-1}^{\mathrm{T}} \cdots \boldsymbol{x}_l^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{5}$$

$$\boldsymbol{Y}_{k,l} = \left[\boldsymbol{y}_k^{\mathrm{T}}\ \boldsymbol{y}_{k-1}^{\mathrm{T}} \cdots \boldsymbol{y}_l^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{6}$$

$$\boldsymbol{W}_{k,l} = \left[\boldsymbol{w}_k^{\mathrm{T}}\ \boldsymbol{w}_{k-1}^{\mathrm{T}} \cdots \boldsymbol{w}_l^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{7}$$

$$\boldsymbol{V}_{k,l} = \left[\boldsymbol{v}_k^{\mathrm{T}}\ \boldsymbol{v}_{k-1}^{\mathrm{T}} \cdots \boldsymbol{v}_l^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{8}$$

The expanded model matrix $\boldsymbol{A}_{k-l} \in \mathbb{R}^{Nn\times n}$, process noise matrix $\boldsymbol{B}_{k-l} \in \mathbb{R}^{Nn\times Nu}$, observation matrix $\boldsymbol{C}_{k-l} \in \mathbb{R}^{Np\times n}$, auxiliary matrix $\boldsymbol{H}_{k-l} \in \mathbb{R}^{Np\times Nu}$, and measurement noise matrix $\boldsymbol{D}_{k-l} \in \mathbb{R}^{Np\times Nv}$ are all time invariant and dependent on the horizon length of N points. Models (1) and (2) suggest that these matrices can be written as, respectively

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

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

$$\boldsymbol{C}_i = \bar{\boldsymbol{C}}_i\boldsymbol{A}_i, \tag{11}$$

$$\boldsymbol{H}_i = \bar{\boldsymbol{C}}_i\boldsymbol{B}_i, \tag{12}$$

$$\boldsymbol{D}_i = \mathrm{diag}(\underbrace{\boldsymbol{D}\,\boldsymbol{D}\cdots\boldsymbol{D}}_{i+1}), \tag{13}$$

$$\bar{\boldsymbol{C}}_i = \mathrm{diag}(\underbrace{\boldsymbol{C}\,\boldsymbol{C}\cdots\boldsymbol{C}}_{i+1}). \tag{14}$$

At the initial point, we have $\boldsymbol{x}_l = \boldsymbol{x}_l + \boldsymbol{B}\boldsymbol{w}_l$, which means that the initial state $\boldsymbol{x}_l$ is known in advance.

On a horizon $[l, k]$, the FIR filter using N past neighbouring measurement points can be described by

$$\hat{\boldsymbol{x}}_{k|k} = \boldsymbol{K}_k\boldsymbol{Y}_{k,l}, \tag{15}$$

where $\hat{\boldsymbol{x}}_{k|k}$ is the estimate ($\hat{\boldsymbol{x}}_{k|k}$ means the estimate at k via measurements from the past to k), and $\boldsymbol{K}_k$ is the FIR filter gain which is determined using a cost function. The operation principles of the FIR and IIR (Kalman) filters are illustrated in Fig. 1. One can easily notice that the recursive IIR (Kalman) filter uses only one nearest past measurement to compute the estimate, while the convolution-based batch FIR filter requires N most recent measurements. Four basic FIR filters such as UFIR, OFIR, and in-betweens (OFIR-EU and OUFIR) are considered in the following sections.

3 UFIR filter

Intuitively, one wants to provide tracking, state estimation, localisation, and so on very accurately – without bias in the estimate. This leads to a concept of the *unbiased estimate*. In optimal filtering, the estimate (15) is said to be unbiased if it obeys the following unbiasedness condition [2]

$$E\{\boldsymbol{x}_k\} = E\{\hat{\boldsymbol{x}}_{k|k}\}, \tag{16}$$

which means that the average of the estimate $\hat{\boldsymbol{x}}_{k|k}$ is required to be equal to that of the model $\boldsymbol{x}_k$.

To arrive at the UFIR estimate, one can combine (3) and (4) and specify the model $\boldsymbol{x}_k$ as

$$\boldsymbol{x}_k = \boldsymbol{A}^{N-1}\boldsymbol{x}_l + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l}, \tag{17}$$

where $\bar{\boldsymbol{B}}_{k-l}$ is the first vector row in $\boldsymbol{B}_{k-l}$. Then substituting (15) and (17) into (16), replacing the term $\boldsymbol{Y}_{k,l}$ with (4), and providing the averaging, lead to the *unbiasedness constraint* [23]

$$\boldsymbol{A}^{N-1} = \bar{\boldsymbol{K}}_k\boldsymbol{C}_{k-l} \tag{18}$$

which is also known as the *deadbeat constraint* [15]. Here, $\bar{\boldsymbol{K}}_k$ represents the UFIR filter gain which can be derived similarly to the OLS estimator. In fact, if we multiply $\boldsymbol{A}^{N-1}$ in (18) from the right-hand side with the identity matrix $(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}$ and neglect $\boldsymbol{C}_{k-l}$ in both sides, we get

$$\bar{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}, \tag{19}$$

where N should be chosen as $N \geqslant n$ to guarantee the invertibility of $\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}$.

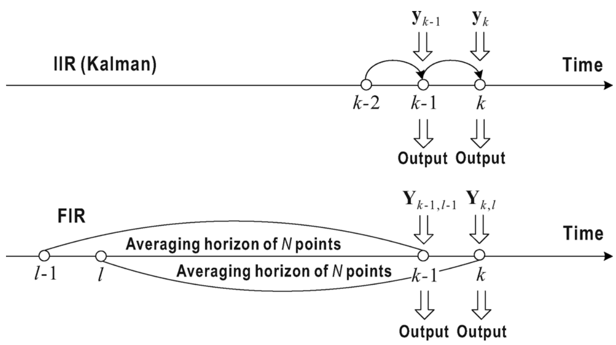

Fig. 1 *Operation time diagrams of the IIR (Kalman) and FIR filters*

The gain (19) set to (15) guarantees that $\hat{\boldsymbol{x}}_{k|k}$ obeys (16) and the estimate is thus unbiased. The main advantage of (19) is that no information about noise is required and the UFIR can thus be used in any noise environment. Design and developments of UFIR filters for diverse models can be found in [6, 7, 22, 42, 44, 45].

4 OFIR filter

By obeying only (16), the UFIR filter quarantines the best estimation accuracy on a horizon of N points. However, the precision of its estimate may not be sufficiently high. If so, then an optimisation problem should be solved in order to increase the precision at some expense in accuracy. The MSE sense is most common in linear optimal filtering as providing the best trade-off between the accuracy and precision. The minimisation of MSE leads to the OFIR filter (Strictly speaking, optimality in filtering refers to linear models with Gaussian noise. The OFIR filter considered in this paper minimises the MSE assuming zero mean noise having arbitrary distribution.).

To minimise MSE in the estimate $\hat{\boldsymbol{x}}_{k|k}$, the instantaneous estimation error can be defined at k as

$$\boldsymbol{e}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}. \tag{20}$$

Then the optimal gain $\tilde{\boldsymbol{K}}_k$ for the OFIR filter can be found by solving the optimisation problem

$$\tilde{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} E\{\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}\} \tag{21}$$

which, if to substitute $\boldsymbol{x}_k$ with (17) and $\hat{\boldsymbol{x}}_k$ with (15) into (20), becomes

$$\tilde{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} E\big\{\big(\boldsymbol{A}^{N-1}\boldsymbol{x}_l + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l} - \boldsymbol{K}_k\boldsymbol{Y}_{k,l}\big)(\cdots)^{\mathrm{T}}\big\}, \tag{22}$$

where $(\cdots)$ denotes the term that is equal to the relevant preceding term. There are many ways of solving (22). One of the most straightforward approaches implies invoking the orthogonality condition [2]. It has been shown in [41] that (22) in this case transforms to

$$E\{(\boldsymbol{A}^{N-1}\boldsymbol{x}_l + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l} - \tilde{\boldsymbol{K}}_k\boldsymbol{Y}_{k,l})\boldsymbol{Y}_{k,l}^{\mathrm{T}}\} = \boldsymbol{0} \tag{23}$$

and, by incorporating (4), it can be rewritten as

$$\begin{aligned}\boldsymbol{0} = E\big\{&[(\boldsymbol{A}^{N-1} - \tilde{\boldsymbol{K}}_k\boldsymbol{C}_{k-l})\boldsymbol{x}_l + (\bar{\boldsymbol{B}}_{k-l} \\ &- \tilde{\boldsymbol{K}}_k\boldsymbol{H}_{k-l})\boldsymbol{W}_{k,l} - \tilde{\boldsymbol{K}}_k\boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l}] \\ &\times (\boldsymbol{C}_{k-l}\boldsymbol{x}_l + \boldsymbol{H}_{k-l}\boldsymbol{W}_{k,l} + \boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l})^{\mathrm{T}}\big\}.\end{aligned} \tag{24}$$

Under the assumption that the initial state and the system and measurement noise sources are mutually uncorrelated and independent process, the averaging and rearranging the terms bring the solution to (24) into the form of [41]

$$\tilde{\boldsymbol{K}}_k = \big(\boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}} + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\big)\boldsymbol{\Delta}_{x+w+v}^{-1}, \tag{25}$$

where $\boldsymbol{\Theta}_x = E\{\boldsymbol{x}_l\boldsymbol{x}_l^{\mathrm{T}}\}$, $\boldsymbol{\Theta}_w = E\{\boldsymbol{W}_{k,l}\boldsymbol{W}_{k,l}^{\mathrm{T}}\}$, $\boldsymbol{\Theta}_v = E\{\boldsymbol{V}_{k,l}\boldsymbol{V}_{k,l}^{\mathrm{T}}\}$, $\boldsymbol{\Delta}_x = \boldsymbol{C}_{k-l}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}$, $\boldsymbol{\Delta}_w = \boldsymbol{H}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}$, $\boldsymbol{\Delta}_v = \boldsymbol{D}_{k-l}\boldsymbol{\Theta}_v\boldsymbol{D}_{k-l}^{\mathrm{T}}$, and $\boldsymbol{\Delta}_{x+w+v} = \boldsymbol{\Delta}_x + \boldsymbol{\Delta}_w + \boldsymbol{\Delta}_v$.

Further multiplying $\boldsymbol{\Theta}_x$ in (25) from the left-hand side with the identity matrix $(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}^{-1}$, the OFIR filter gain becomes

$$\tilde{\boldsymbol{K}}_k = \big(\bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_x + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\big)\boldsymbol{\Delta}_{x+w+v}^{-1}, \tag{26}$$

where $\bar{\boldsymbol{K}}_k$ is the UFIR filter gain given by (19). Note that matrix $\boldsymbol{\Delta}_x$ associated with the mean square initial state $\boldsymbol{\Theta}_x$ can be computed by solving the discrete algebraic Riccati equation (DARE) [41]

$$\boldsymbol{\Delta}_x\boldsymbol{\Delta}_{w+v}^{-1}\boldsymbol{\Delta}_x + 2\boldsymbol{\Delta}_x + \boldsymbol{\Delta}_{w+v} - \boldsymbol{Y}_{k,l}\boldsymbol{Y}_{k,l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}\boldsymbol{\Delta}_x = \boldsymbol{0}.$$

Detailed theory, developments, and some application of OFIR filters can be found in [15, 16, 18, 19, 23, 30, 41, 46].

5 OFIR-EU filter

A significant advantage of the OFIR filter is that it minimises the MSE. However, it requires the noise statistics and the initial error statistics. Although the latter can be found by solving the DARE, the solution which is not always available in real time implies extra computational burden. An intermediate solution between the OFIR filter and the UFIR filter which ignores the noise statistics and the initial error statistics can be found if we subject (21) to the unbiasedness constraint (18). The relevant OFIR-EU filter gain can then be found by solving the following optimisation problem

$$\begin{aligned}\hat{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} E\{\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}\}&, \\ \text{subject to (18).}&\end{aligned} \tag{27}$$

A solution to (27) still has not been addressed in the literature. We provide it below and derive the OFIR-EU filter gain. To this end, the following lemma will be used.

Lemma 1: The trace optimisation problem is given by

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

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

$$\boldsymbol{K} = \begin{bmatrix}\boldsymbol{Z} \\ \boldsymbol{G} \\ \boldsymbol{M}\end{bmatrix}^{\mathrm{T}} \begin{bmatrix}\theta\Big(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Xi}^{-1}\boldsymbol{U}\Big)^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Xi}^{-1} \\ \boldsymbol{HF}^{\mathrm{T}}\boldsymbol{\Xi}^{-1}\boldsymbol{\Pi} \\ \boldsymbol{PL}^{\mathrm{T}}\boldsymbol{\Xi}^{-1}\boldsymbol{\Pi}\end{bmatrix},$$

where $\boldsymbol{\Pi} = \boldsymbol{I} - \theta\boldsymbol{U}\Big(\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Xi}^{-1}\boldsymbol{U}\Big)^{-1}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Xi}^{-1}$ and

$$\boldsymbol{\Xi} = \begin{cases}\boldsymbol{LPL}^{\mathrm{T}} + \boldsymbol{S}, & \text{if } \boldsymbol{F} = \boldsymbol{U},\ \boldsymbol{G} = \boldsymbol{Z},\ \text{and } \theta = 1 \\ \boldsymbol{FHF}^{\mathrm{T}} + \boldsymbol{S}, & \text{if } \boldsymbol{L} = \boldsymbol{U},\ \boldsymbol{M} = \boldsymbol{Z},\ \text{and } \theta = 1 \\ \boldsymbol{FHF}^{\mathrm{T}} + \boldsymbol{LPL}^{\mathrm{T}} + \boldsymbol{S}, & \text{if } \theta = 0\end{cases}.$$

Proof: The proof can be obtained directly by following the line presented in [15, 47]. Therefore, we omit it. □

5.1 OFIR-EU filter design

Using the trace operation, the optimisation problem (27) subject to (18) can be reformulated via the cost function

$$\begin{aligned}\hat{\boldsymbol{K}}_k &= \arg\min_{\boldsymbol{K}_k} E\big\{\mathrm{tr}\big[\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}\big]\big\} \\ &= \arg\min_{\boldsymbol{K}_k} E\big\{\mathrm{tr}\big[\big(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k}\big)(\cdots)^{\mathrm{T}}\big]\big\}.\end{aligned}$$

By substituting $\boldsymbol{x}_k$ with (17) and $\hat{\boldsymbol{x}}_{k|k}$ with (15), and using the extended measurement (4), the above cost function becomes

$$\begin{aligned}\hat{\boldsymbol{K}}_k &= \arg\min_{\boldsymbol{K}_k} E\{\mathrm{tr}[(\boldsymbol{A}^{N-1}\boldsymbol{x}_l + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l} - \boldsymbol{K}_k\boldsymbol{Y}_{k,l})(\cdots)^{\mathrm{T}}]\} \\ &= \arg\min_{\boldsymbol{K}_k} E\{\mathrm{tr}[((\boldsymbol{K}_k\boldsymbol{C}_{k-l} - \boldsymbol{A}^{N-1})\boldsymbol{x}_l \\ &\quad + (\boldsymbol{K}_k\boldsymbol{H}_{k-l} - \bar{\boldsymbol{B}}_{k-l})\boldsymbol{W}_{k,l} + \boldsymbol{K}_k\boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l})(\cdots)^{\mathrm{T}}]\}. \end{aligned} \quad (28)$$

If we subject (28) to the constraint (18), provide the averaging, and rearrange the terms, (28) can be transformed to

$$\hat{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} \mathrm{tr}[(\boldsymbol{K}_k\boldsymbol{H}_{k-l} - \bar{\boldsymbol{B}}_{k-l})\boldsymbol{\Theta}_w(\cdots)^{\mathrm{T}} + \boldsymbol{K}_k\boldsymbol{\Delta}_v(\cdots)^{\mathrm{T}}] \quad (29)$$

referring to the fact that the system noise vector $\boldsymbol{W}_{k,l}$ and the measurement noise vector $\boldsymbol{V}_{k,l}$ are pairwise independent.

By Lemma 1, case $\theta = 1$, the solution to the optimisation problem (29) can be found by neglecting $\boldsymbol{L}$, $\boldsymbol{M}$, and $\boldsymbol{P}$ and using the replacements: $\boldsymbol{F} \leftarrow \boldsymbol{H}_{k-l}$, $\boldsymbol{G} \leftarrow \bar{\boldsymbol{B}}_{k-l}$, $\boldsymbol{H} \leftarrow \boldsymbol{\Theta}_w$, $\boldsymbol{U} \leftarrow \boldsymbol{C}_{k-l}$, $\boldsymbol{Z} \leftarrow \boldsymbol{A}^{N-1}$, and $\boldsymbol{S} \leftarrow \boldsymbol{\Delta}_v$. We thus have

$$\hat{\boldsymbol{K}}_k = \boldsymbol{K}_k(a) + \boldsymbol{K}_k(b), \quad (30)$$

where

$$\boldsymbol{K}_k(a) = \boldsymbol{A}^{N-1}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}, \quad (31)$$

$$\boldsymbol{K}_k(b) = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}(\boldsymbol{I} - \boldsymbol{\Omega}_{k-l}), \quad (32)$$

in which

$$\boldsymbol{\Omega}_{k-l} = \boldsymbol{C}_{k-l}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}. \quad (33)$$

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

Theorem 1: Given the state-space models (1) and (2) with zero mean mutually independent and uncorrelated noise vectors $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$, then OFIR-EU filter utilising measurements from l to k is stated by

$$\hat{\boldsymbol{x}}_{k|k} = [\boldsymbol{K}_k(a) + \boldsymbol{K}_k(b)]\boldsymbol{Y}_{k,l}, \quad (34)$$

where $\boldsymbol{Y}_{k,l} \in \mathbb{R}^{Np\times 1}$ is the measurement vector given by (6), and $\boldsymbol{K}_k(a)$ and $\boldsymbol{K}_k(b)$ are given by (30) and (31) with $\boldsymbol{C}_{k-l}$ and $\bar{\boldsymbol{B}}_{k-l}$ specified by (11) and (10), respectively.

Proof: The proof is provided by (27)–(33). □

Note that the horizon length N for (30) should be chosen such that the inverse in $\hat{\boldsymbol{K}}_k$ exists. Generally, one can set $N \geqslant n$, where n is the number of the model states. Table 1 summarises the steps in the OFIR-EU estimation algorithm, in which the noise statistics are assumed to be known for measurements available from l to k. Given N, compute $\boldsymbol{K}_k(a)$ and $\boldsymbol{K}_k(b)$ according to (31) and (32), respectively, then the OFIR-EU estimate can be obtained at time index k by (34).

5.2 Unified form for OFIR and OFIR-EU filters

To ascertain a correspondence between the OFIR filter and its modifications associated with the unbiasedness constraint (18), the optimisation problem (28) can be rewritten with respect to the unified gain $\breve{\boldsymbol{K}}_k$ as

$$\begin{aligned}\breve{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} \mathrm{tr}[&(\boldsymbol{K}_k\boldsymbol{H}_{k-l} - \bar{\boldsymbol{B}}_{k-l})\boldsymbol{\Theta}_w(\cdots)^{\mathrm{T}} \\ &+ (\boldsymbol{K}_k\boldsymbol{C}_{k-l} - \boldsymbol{A}^{N-1})\boldsymbol{\Theta}_x(\cdots)^{\mathrm{T}} + \boldsymbol{K}_k\boldsymbol{\Delta}_v(\cdots)^{\mathrm{T}}]\end{aligned} \quad (35)$$

subject to the constraint $\mathfrak{L}_{\{\boldsymbol{K}_k\boldsymbol{C}_{k-l}=\boldsymbol{A}^{N-1}\}|\theta}$. By Lemma 1, a solution to (35) attains the form of

$$\begin{aligned}\breve{\boldsymbol{K}}_k = \theta\boldsymbol{A}^{N-1}\boldsymbol{\Lambda}_{k-l} &+ \boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}\bar{\boldsymbol{\Delta}}_{x+w+v}^{-1}(\boldsymbol{I} - \theta\boldsymbol{C}_{k-l}\boldsymbol{\Lambda}_{k-l}) \\ &+ \boldsymbol{B}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\bar{\boldsymbol{\Delta}}_{x+w+v}^{-1}(\boldsymbol{I} - \theta\boldsymbol{C}_{k-l}\boldsymbol{\Lambda}_{k-l}),\end{aligned} \quad (36)$$

where

$$\boldsymbol{\Lambda}_{k-l} = \left(\boldsymbol{C}_{k-l}^{\mathrm{T}}\bar{\boldsymbol{\Delta}}_{x+w+v}^{-1}\boldsymbol{C}_{k-l}\right)^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\bar{\boldsymbol{\Delta}}_{x+w+v}^{-1}, \quad (37)$$

in which

$$\bar{\boldsymbol{\Delta}}_{x+w+v} = \begin{cases}\boldsymbol{\Delta}_{w+v}, & \text{if } \theta = 1 \\ \boldsymbol{\Delta}_{x+w+v}, & \text{if } \theta = 0\end{cases}. \quad (38)$$

In a special case of $\theta = 1$, (36) reduces to

$$\begin{aligned}\breve{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}\boldsymbol{\Lambda}_{k-l} &+ \boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}(\boldsymbol{I} - \boldsymbol{C}_{k-l}\boldsymbol{\Lambda}_{k-l}) \\ &+ \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}(\boldsymbol{I} - \boldsymbol{C}_{k-l}\boldsymbol{\Lambda}_{k-l}),\end{aligned} \quad (39)$$

where $\boldsymbol{\Lambda}_{k-l}$ is given by (37), in which $\bar{\boldsymbol{\Delta}}_{x+w+v}$ is specified by (38) with $\theta = 1$. Referring to (31) and (32) and taking into consideration that $\boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}(\boldsymbol{I} - \boldsymbol{C}_{k-l}\boldsymbol{\Lambda}_{k-l}) = \boldsymbol{0}$, we come up with a conclusion that

$$\breve{\boldsymbol{K}}_k = \hat{\boldsymbol{K}}_k. \quad (40)$$

In the unconstrained case of $\theta = 0$, (36) transforms to

$$\breve{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{x+w+v}^{-1} + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{x+w+v}^{-1}. \quad (41)$$

By multiplying $\boldsymbol{\Theta}_x$ with identity $(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}$ from the left-hand side, (41) turns out to be

$$\begin{aligned}\breve{\boldsymbol{K}}_k &= \left(\boldsymbol{A}^{N-1}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_x \right. \\ &\quad \left. + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\right)\boldsymbol{\Delta}_{x+w+v}^{-1} \\ &= (\bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_x + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}})\boldsymbol{\Delta}_{x+w+v}^{-1} \\ &= \tilde{\boldsymbol{K}}_k,\end{aligned} \quad (42)$$

and we infer that this case corresponds to the OFIR filter which gain was found in (26). At this point, we notice that (36) is a unified generalised form for the OFIR filter gain which minimises the MSE in the estimate of discrete time-invariant state-space model. In this regard, the OFIR filter gain derived in [41] and OFIR-EU filter gain specified by Theorem 1 can be considered as special cases of (36).

Table 1 OFIR-EU filtering algorithm

Stage	
Given	$N \geqslant n$, $l = k-N+1$
Find	$\boldsymbol{K}_k(a)$ by (31) and $\boldsymbol{K}_k(b)$ by (32)
Compute	$\hat{\boldsymbol{x}}_{k\vert k} = [\boldsymbol{K}_k(a) + \boldsymbol{K}_k(b)]\boldsymbol{Y}_{k,l}$

6 OUFIR filter

Another in-between solution can be found to minimise the MSE in the UFIR filter. The relevant filter is called the OUFIR filter, because it evolves as UFIR-to-OFIR filter by increasing N. The OUFIR filter can be obtained by representing its gain $\boldsymbol{K}_{k,\mathrm{ou}}$ as a linear combination of $\bar{\boldsymbol{K}}_k$ given by (19) and an auxiliary unknown term of the same class $\boldsymbol{K}'_k$, and then defining $\boldsymbol{K}'_k$ by solving the optimisation problem

$$\boldsymbol{K}'_k = \arg\min_{\boldsymbol{K}_k} E\{\boldsymbol{e}_k\boldsymbol{e}_k^{\mathrm{T}}\}. \qquad (43)$$

Similar reasonings as for (27) bring (43) to

$$\begin{aligned}\boldsymbol{K}'_k &= \arg\min_{\boldsymbol{K}_k} E\big\{\mathrm{tr}\big[\big((\bar{\boldsymbol{K}}_k + \boldsymbol{K}'_k)\boldsymbol{Y}_{k,l} - \boldsymbol{A}^{N-1}\boldsymbol{x}_l \\ &\quad - \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l}\big)(\cdots)^{\mathrm{T}}\big]\big\}. \\ &= \arg\min_{\boldsymbol{K}_k} E\big\{\mathrm{tr}\big[\big((\bar{\boldsymbol{K}}_k + \boldsymbol{K}'_k)(\boldsymbol{H}_{k-l}\boldsymbol{W}_{k,l} + \boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l}) \\ &\quad + \boldsymbol{K}'_k\boldsymbol{C}_{k-l}\boldsymbol{x}_l - \bar{\boldsymbol{B}}_{k-l}\boldsymbol{W}_{k,l}\big)(\cdots)^{\mathrm{T}}\big]\big\}\end{aligned}$$

that, for noise sources which are mutually uncorrelated and independent on the initial state $\boldsymbol{x}_l$, becomes

$$\begin{aligned}\boldsymbol{K}'_k = \arg\min_{\boldsymbol{K}_k} \mathrm{tr}\big\{&\big((\bar{\boldsymbol{K}}_k + \boldsymbol{K}'_k)\boldsymbol{H}_{k-l} - \bar{\boldsymbol{B}}_{k-l}\big)\boldsymbol{\Theta}_w(\cdots)^{\mathrm{T}} \\ &+ (\bar{\boldsymbol{K}}_k + \boldsymbol{K}'_k)\boldsymbol{\Delta}_v(\cdots)^{\mathrm{T}} + \boldsymbol{K}'_k\boldsymbol{\Delta}_x(\cdots)^{\mathrm{T}}\big\}.\end{aligned} \qquad (44)$$

A solution to (44) provided by using Lemma 1 is given by

$$\boldsymbol{K}'_k = \bar{\boldsymbol{\Upsilon}}_{k-l}(\boldsymbol{I} - \bar{\boldsymbol{\Omega}}_{k-l}), \qquad (45)$$

where

$$\bar{\boldsymbol{\Upsilon}}_{k-l} = (\bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}} - \bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v})\boldsymbol{\Delta}_{x+w+v}^{-1}, \qquad (46)$$

$$\bar{\boldsymbol{\Omega}}_{k-l} = \boldsymbol{C}_{k-l}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{x+w+v}^{-1}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{x+w+v}^{-1}. \qquad (47)$$

At this point, the OUFIR filter gain $\boldsymbol{K}_{k,\mathrm{ou}}$ appears to be

$$\boldsymbol{K}_{k,\mathrm{ou}} = \bar{\boldsymbol{K}}_k + \boldsymbol{K}'_k. \qquad (48)$$

Lemma 1 suggests that $\boldsymbol{K}_{k,\mathrm{ou}}$ does not depend on the initial state matrix $\boldsymbol{\Delta}_x$, and $\boldsymbol{\Delta}_x$ can thus be arbitrarily defined in (45), provided that the inverse in (47) exists. Thus, the last term $\boldsymbol{K}'_k\boldsymbol{\Delta}_x(\cdots)^{\mathrm{T}}$ in (44) can be omitted, by $\boldsymbol{\Delta}_x = \boldsymbol{0}$. This fundamental property was postulated in many papers [15, 21, 27, 42] and, based upon, $\boldsymbol{K}_{k,\mathrm{ou}}$ can be rewritten equivalently as

$$\boldsymbol{K}_{k,\mathrm{ou}} = \bar{\boldsymbol{K}}_k + \boldsymbol{K}''_k, \qquad (49)$$

where

$$\boldsymbol{K}''_k = \boldsymbol{\Upsilon}_{k-l}\big(\boldsymbol{I} - \boldsymbol{\Omega}_{k-l}\big),$$

$$\boldsymbol{\Upsilon}_{k-l} = (\bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}} - \boldsymbol{K}_k^{\mathrm{U}}\boldsymbol{\Delta}_{w+v})\boldsymbol{\Delta}_{w+v}^{-1},$$

and $\boldsymbol{\Omega}_{k-l}$ is given by (33). Referring to (33) and making some rearrangements, one infers that

$$\begin{aligned}\boldsymbol{K}_{k,\mathrm{ou}} &= \bar{\boldsymbol{K}}_k - \bar{\boldsymbol{K}}_k(\boldsymbol{I} - \boldsymbol{\Omega}_{k-l}) + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}(\boldsymbol{I} - \boldsymbol{\Omega}_{k-l}) \\ &= \bar{\boldsymbol{K}}_k\boldsymbol{C}_{k-l}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_{w+v}^{-1} + \boldsymbol{K}_k(b) \\ &= \hat{\boldsymbol{K}}_k\end{aligned}$$

which is formalised below with a theorem.

Theorem 2: The OUFIR filter gain specified by (48) is identical to the OFIR-EU filter gain specified by Theorem 1

$$K_{k,\mathrm{ou}} = \hat{\boldsymbol{K}}_k.$$

Proof: The proof has been given above. □

It follows from Theorem 2 that the gain $\boldsymbol{K}_{k,\mathrm{ou}}$ is not unique. One may assume any initial state matrix $\boldsymbol{\Delta}_x$, compute it by solving the DARE as in [16], or even neglect $\boldsymbol{\Delta}_x$ as we have done above. Although each of these cases require particular algorithms, Lemma 1 suggests that the estimation accuracy will not be affected by $\boldsymbol{\Delta}_x$. We use this property below while comparing different kinds of FIR filters.

6.1 Deterministic state-space model

Having no noise in (1) and (2), the cost function in (28) becomes

$$\hat{\boldsymbol{K}}_k = \arg\min_{\boldsymbol{K}_k} E\big\{\mathrm{tr}\big[(\boldsymbol{A}^{N-1}\boldsymbol{x}_l - \boldsymbol{K}_k\boldsymbol{C}_{k-l}\boldsymbol{x}_l)(\cdots)^{\mathrm{T}}\big]\big\}. \qquad (50)$$

According to (18), the above equation becomes identically zero. Hence, the solution to (50) is not unique and any gain $\boldsymbol{K}_k$ can be considered as the solution. Then, we have a conclusion that *the UFIR filter is a deadbeat filter for deterministic systems*.

If (18) is not used, the solution to (50) is found to be

$$\tilde{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}\boldsymbol{\Theta}_x\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{\Delta}_x^{-1}. \qquad (51)$$

Multiplying $\boldsymbol{\Theta}_x$ with $(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}$ from the left-hand side of (51) gives us

$$\tilde{\boldsymbol{K}}_k = \bar{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}\big(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l}\big)^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}$$

which can also be obtained by setting the terms $\boldsymbol{\Delta}_w$ and $\boldsymbol{\Delta}_v$ in (26) to zero. Therefore, *OFIR filter is also a deadbeat filter*. Table 2 summarises the gains for the UFIR, OFIR-EU (OUFIR), and OFIR filters. Developments and applications of different kinds of OUFIR filters can be found in [15, 28, 48, 49].

Provided the developments and comparative analysis of the UFIR and OFIR filters, and two possible in-betweens such as the OFIR-EU and OUFIR filters, we analyse the estimation errors in the following section.

7 Estimation errors

As long as the MSE is recognised to be the main characteristic of any linear filter, we compare the MSEs of the above-observed filters and outline their common properties and special features referring to Table 2.

Most generally, the MSE $\boldsymbol{J}_k$ at time k is specified as

$$\boldsymbol{J}_k = E\big\{(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|k})(\cdots)^{\mathrm{T}}\big\} \qquad (52)$$

and further transformed to

$$\boldsymbol{J}_k = E\{(E\{\boldsymbol{x}_k\} - E\{\hat{\boldsymbol{x}}_{k|k}\})(\cdots)^{\mathrm{T}}\} + E\{(\boldsymbol{x}_k - E\{\boldsymbol{x}_k\})(\cdots)^{\mathrm{T}}\} + E\{(E\{\hat{\boldsymbol{x}}_{k|k}\} - \hat{\boldsymbol{x}}_{k|k})(\cdots)^{\mathrm{T}}\} - 2E\left\{\boldsymbol{x}_k\hat{\boldsymbol{x}}_{k|k}^{\mathrm{T}} - \boldsymbol{x}_k E\{\hat{\boldsymbol{x}}_{k|k}\}^{\mathrm{T}}\right\} \quad (53a)$$

$$= \mathrm{Bias}^2(\hat{\boldsymbol{x}}_{k|k}) + \mathrm{Var}(\boldsymbol{x}_k) + \mathrm{Var}(\hat{\boldsymbol{x}}_{k|k}) - 2\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}), \quad (53b)$$

where the terms in (53b) correspond to the relevant terms in (53a). Here, $\mathrm{Bias}^2(\hat{\boldsymbol{x}}_{k|k})$ represents the squared bias in $\hat{\boldsymbol{x}}_{k|k}$, $\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k})$ represents the covariance between $\boldsymbol{x}_k$ and $\hat{\boldsymbol{x}}_{k|k}$, and $\mathrm{Var}(\boldsymbol{x}_k)$ and $\mathrm{Var}(\hat{\boldsymbol{x}}_{k|k})$ represents the variances of state $\boldsymbol{x}_k$ and estimate $\hat{\boldsymbol{x}}_{k|k}$, respectively. The state variance $\mathrm{Var}(\boldsymbol{x}_k)$ is specified as

$$\mathrm{Var}(\boldsymbol{x}_k) = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\bar{\boldsymbol{B}}_{k-l}^{\mathrm{T}}. \quad (54)$$

For unbiased estimates, we get

$$\mathrm{Bias}(\hat{\boldsymbol{x}}_{k|k}) = \boldsymbol{0}. \quad (55)$$

An analysis of MSEs for diverse FIR filters is given below.

7.1 MSE of the UFIR filter

In this case, $\mathrm{Var}(\hat{\boldsymbol{x}}_{k|k})$ on the right-hand side of (53b) has the following equalities

$$\mathrm{Var}(\hat{\boldsymbol{x}}_{k|k}) = E\{(\bar{\boldsymbol{K}}_k\boldsymbol{H}_{k-l}\boldsymbol{W}_{k,l} + \bar{\boldsymbol{K}}_k\boldsymbol{D}_{k-l}\boldsymbol{V}_{k,l})(\cdots)^{\mathrm{T}}\} = \bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}}. \quad (56)$$

Since $\boldsymbol{W}_{k,l}$ and $\boldsymbol{V}_{k,l}$ are mutually independent, the covariance $\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k})$ can be written as

$$\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}) = E\left\{(\boldsymbol{x}_k - E\{\boldsymbol{x}_k\})(\hat{\boldsymbol{x}}_{k|k} - E\{\hat{\boldsymbol{x}}_{k|k}\})^{\mathrm{T}}\right\} = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\bar{\boldsymbol{K}}_k)^{\mathrm{T}}, \quad (57)$$

Finally, the MSE in the UFIR filter can be computed as

$$\bar{\boldsymbol{J}}_k = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w(\cdots)^{\mathrm{T}} + \bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}} - 2\bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\bar{\boldsymbol{K}}_k)^{\mathrm{T}}, \quad (58)$$

where $\bar{\boldsymbol{K}}_k$ is given by (19).

7.2 MSE of the OFIR-EU filter

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

$$\mathrm{Var}(\hat{\boldsymbol{x}}_{k|k}) = \hat{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}}, \quad (59)$$

$$\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}) = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\hat{\boldsymbol{K}}_k)^{\mathrm{T}}. \quad (60)$$

Table 2 Gains of UFIR, OFIR-EU (or OUFIR), and OFIR filters

Filter	Gain
UFIR	$\bar{\boldsymbol{K}}_k = \boldsymbol{A}^{N-1}(\boldsymbol{C}_{k-l}^{\mathrm{T}}\boldsymbol{C}_{k-l})^{-1}\boldsymbol{C}_{k-l}^{\mathrm{T}}$
OFIR-EU	$\hat{\boldsymbol{K}}_k = \boldsymbol{K}_k(a) + \boldsymbol{K}_k(b)$
OFIR	$\tilde{\boldsymbol{K}}_k = (\bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_x + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}})\boldsymbol{\Delta}_{x+w+v}^{-1}$

With (49), we arrive at

$$\mathrm{Var}(\hat{\boldsymbol{x}}_{k|k}) = \bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}} + 2\bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\boldsymbol{K}_k'')^{\mathrm{T}} + \boldsymbol{K}_k''\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}}, \quad (61)$$

$$\mathrm{Cov}(\boldsymbol{x}_k, \hat{\boldsymbol{x}}_{k|k}) = \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\bar{\boldsymbol{K}}_k)^{\mathrm{T}} + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\boldsymbol{K}_k'')^{\mathrm{T}}. \quad (62)$$

Substituting (55), (61), and (62) into (53b) and rearranging the terms yields

$$\hat{\boldsymbol{J}}_k = \bar{\boldsymbol{J}}_k + \boldsymbol{K}_k''\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}} - 2(\bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}} - \bar{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v})(\boldsymbol{K}_k'')^{\mathrm{T}}. \quad (63)$$

Using $\boldsymbol{\Upsilon}_{k-l}$, (63) can be rewritten as

$$\hat{\boldsymbol{J}}_k = \bar{\boldsymbol{J}}_k + \boldsymbol{K}_k''\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}} - 2\boldsymbol{\Upsilon}_{k-l}\boldsymbol{\Delta}_{w+v}(\boldsymbol{K}_k'')^{\mathrm{T}}. \quad (64)$$

7.3 MSE of the OFIR filter

Before calculating estimation errors, the OFIR filter gain $\tilde{\boldsymbol{K}}_k$ represented by (26) can be equivalently given by

$$\tilde{\boldsymbol{K}}_k = \bar{\boldsymbol{K}}_k + \bar{\boldsymbol{\Upsilon}}_{k-l}. \quad (65)$$

Then, the bias-dependent term becomes $\mathrm{Bias}^2(\hat{\boldsymbol{x}}_{k|k}) = \bar{\boldsymbol{\Upsilon}}_{k-l}\boldsymbol{\Delta}_x(\cdots)^{\mathrm{T}}$. By combining (53b), (54), and (58), the MSE of the OFIR filter can now be computed as

$$\tilde{\boldsymbol{J}}_k = \bar{\boldsymbol{\Upsilon}}_{k-l}\boldsymbol{\Delta}_x(\cdots)^{\mathrm{T}} + \bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w(\cdots)^{\mathrm{T}} + \tilde{\boldsymbol{K}}_k\boldsymbol{\Delta}_{w+v}(\cdots)^{\mathrm{T}} - 2\bar{\boldsymbol{B}}_{k-l}\boldsymbol{\Theta}_w\boldsymbol{H}_{k-l}^{\mathrm{T}}(\tilde{\boldsymbol{K}}_k)^{\mathrm{T}}. \quad (66)$$

If we further substitute $\tilde{\boldsymbol{K}}_k$ with (65), refer to (58), and rearrange the terms, we arrive at the final form of

$$\tilde{\boldsymbol{J}}_k = \bar{\boldsymbol{J}}_k - \bar{\boldsymbol{\Upsilon}}_{k-l}\boldsymbol{\Delta}_{x+w+v}(\cdots)^{\mathrm{T}}. \quad (67)$$

The above-provided relations allow us to analyse a correspondence between the MSEs in different kinds of FIR filters that we do next.

7.4 Correspondence between the MSEs

Theorem 3: Given the MSEs $\bar{\boldsymbol{J}}_k$ by (58), $\hat{\boldsymbol{J}}_k$ by (64), and $\tilde{\boldsymbol{J}}_k$ by (67). We have

$$\tilde{\boldsymbol{J}}_k \leqslant \hat{\boldsymbol{J}}_k \leqslant \bar{\boldsymbol{J}}_k, \quad (68)$$

where the equalities hold when the state-space model is deterministic.

Proof: The proof of this theorem given in [48]. □

Fig. 2 sketches an algorithmic link between the FIR filtering estimates. In turn, Table 3 summarises the properties of the UFIR, OFIR-EU (or OUFIR), and OFIR filters and we notice again that all these filters are deadbeat filters. In the UFIR filter, bias is guaranteed to be zero, but the noise variance is not minimised and reduced by averaging as $\propto(1/N)$. Therefore, the MSE in UFIR filter generally exceeds those in two other filters. On the contrary, the OFIR filter minimises MSE to obtain an optimal balance between the bias and variance. The OFIR-EU filter occupies an intermediate position between the UFIR and OFIR filters, as it minimises MSE subject to the unbiasedness constraint.

8 Applications to basic systems and special features

We have already pointed out several basic properties of UFIR, OFIR, OFIR-EU, and OUFIR filters. Many others can be learned in particular situations. To investigate the trade-off between these filters in more detail, in this section we test them in a comparison with the KF by the two-state polynomial and harmonic models in different noise environments. The reader can also find some comparisons of KF and FIR filters in [20, 22, 41, 50].

8.1 Two-state polynomial model

The two-state polynomial model serves well in many practical situations associated with moving objects, timekeeping, and navigation. We specify such a model given by (1) and (2) with $\boldsymbol{B} = \boldsymbol{I}$, $\boldsymbol{D} = \boldsymbol{I}$, $\boldsymbol{C} = [1, 0]$, and

$$\boldsymbol{A} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix},$$

where τ is a time step in second. Practical applications dictate that this model may approximate a process under the errors in the noise covariances and/or in the presence of uncertainties. Below, we test the FIR filters in such environments. To evaluate limiting capabilities of FIR filters, we start with an ideal case of exactly known model and noise which fit the process exactly.

8.1.1 Accurate model – ideal case: Suppose that the model represents a process exactly and the noise statistics are completely known. As such a supposition is commonly not supported by practice and rather reflects a desire to avoid mismodelling, we call this case *ideal*. Specifically, we set $\tau = 0.1$ s, the noise variances for the first and second states as $\sigma_{w1}^2 = 1$ and $\sigma_{w2}^2 = 1/s^2$, the measurement noise variance as $\sigma_v^2 = 10$, and the initial states as $x_{10} = 1$ and $x_{20} = 0.01/s$. The process is simulated at 400 subsequent points and, by minimising the MSE in the UFIR filter as suggested in [45], we find $N_{\text{opt}} = 20$.

The purpose of the first simulation is to learn the effect of N on the estimates. With this aim, we first show in Fig. 3 the estimation errors produced by the UFIR, OFIR, and OFIR-EU filters with $N = 20$ along with the KF errors. What can instantly be deduced by observing this figure is that the OFIR and OFIR-EU estimates are almost identical to the KF one. However, the UFIR filter stays a bit away with its not fully consistent estimates and larger variance.

Admitting that the relation between the estimates found in the ideal case (Fig. 3) may be violated if to vary N, we further compute the root MSE (RMSE) by $\text{tr}(\boldsymbol{J}_k)$ as function of N. Typical errors are sketched in Fig. 4*a* for $\sigma_v^2 = 10$ and in Fig. 4*b*

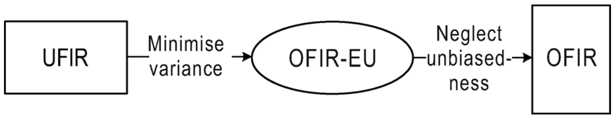

Fig. 2 *Link between the UFIR, OFIR-EU (or OUFIR), and OFIR filters*

Table 3 Basic properties of UFIR, OFIR, and OFIR-EU (or OUFIR) filters

Performance	UFIR	OFIR-EU	OFIR
noise statistics	ignored	required	required
optimality	unbiased	unbiased/optimal	optimal
initial conditions	ignored	ignored	required
complexity	middle	high	high

for $\sigma_v^2 = 100$. One notices that the MSE function of UFIR filter is traditionally concave on N with a minimum at N_{opt} [45]: $N < N_{\text{opt}}$ makes noise reduction inefficient and $N > N_{\text{opt}}$ the bias error dominating. Noticing that the KF is N-invariant, we arrive at the following generalisations:

- The embedded unbiasedness puts OFIR-EU filter in between the UFIR and OFIR filters: the *OFIR-EU filter* becomes essentially the *UFIR filter when* $N < N_{\text{opt}}$ and it becomes the *OFIR filter if* $N > N_{\text{opt}}$.
- The OFIR and OFIR-EU estimates converge to the KF one by increasing N. These estimates become practically indistinguishable when $N \gg N_{\text{opt}}$.
- Since the MSEs in the *OFIR* and *OFIR-EU filters* diminish with N, these filters *are full-horizon* [22, 46]. As the full-horizon batch FIR filters are computationally inefficient in view of large dimensions of all the vectors and matrices, they can be exploited on N_{opt} with no essential loss in accuracy.
- To implement all kinds of batch FIR filters, fast iterative algorithms are required [46, 51].

8.1.2 Filtering with errors in the noise statistics: The noise statistics required by the KF are commonly not completely known to the engineer. To investigate the effect of the imprecisely defined noise covariances in the worst case, we introduce a correction coefficient p as $p^2\boldsymbol{Q}$ and $\boldsymbol{R}/p^2$, vary p from 0.1 to 10, and plot the RMSE $\sqrt{\text{tr}\boldsymbol{J}_k}$ as shown in Fig. 5. Here, p is used to introduce uncertainties into the noise variances. With $p \gg 1$ or $p \ll 1$, large uncertainties exist in the noise variances used in algorithms. When $p = 1$, the noise variances are known accurately. Therefore, the MSE functions of optimal filters are inherently concave on p with a minimum at $p = 1$. On the contrary, the MSE of the UFIR filter is p-invariant.

As expected, the KF is a bit more accurate than the UFIR filter when $p = 1$. However, that is only within a narrow range of $0.7 < p < 1.5$ that KF slightly outperforms UFIR filter if $p \neq 1$. Otherwise, the UFIR filter demonstrates smaller errors. Referring to difficulties in practical determination of the noise statistics [8], the latter can be considered as an important engineering advantage against KF. Some other generalisations also emerge from Fig. 5:

- The embedded unbiasedness makes the OFIR-EU and OUFIR filters p-invariant with $p < 1$ to possess the properties of the UFIR filter.
- With $p < 1$, the KF is more sensitive to errors in the noise statistics than the FIR filters.
- By $p > 1$, the KF, OFIR, OFIR-EU, and OUFIR estimates converge and their MSEs increase.

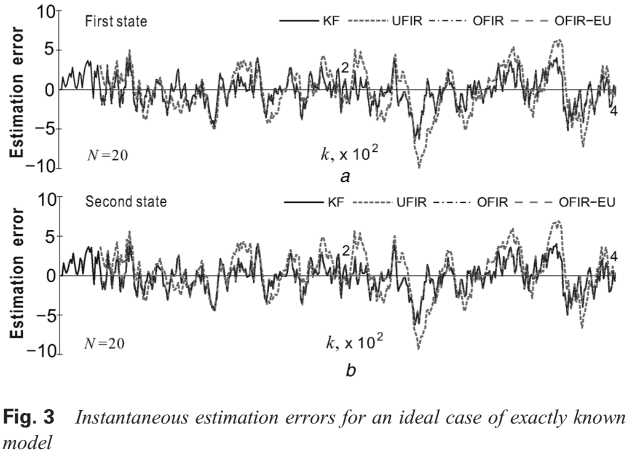

Fig. 3 *Instantaneous estimation errors for an ideal case of exactly known model*

a First state
b Second state

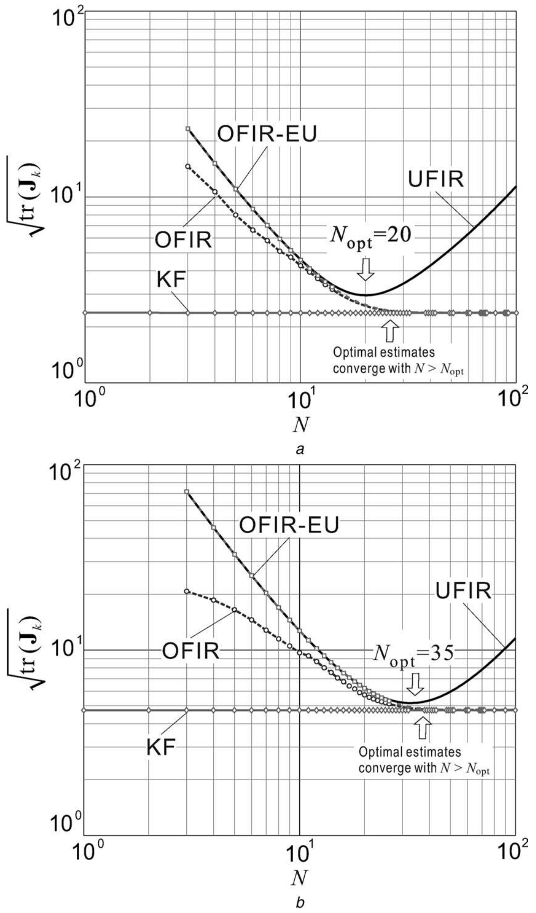

Fig. 4 *RMSEs as functions of N for KF and FIR filters with* $\sigma_w^2 = 1$
a $\sigma_v^2 = 10$
b $\sigma_v^2 = 100$

Overall, we conclude that the OFIR-EU and OUFIR filters have better performance than the OFIR filter and KF in practical situations when $p \neq 1$.

8.1.3 Filtering under the model temporary uncertainties: System behaviours often imply sudden unpredictable jumps and temporary departures from the normal mode. This causes uncertainties in (1) and (2) which in some situations should be

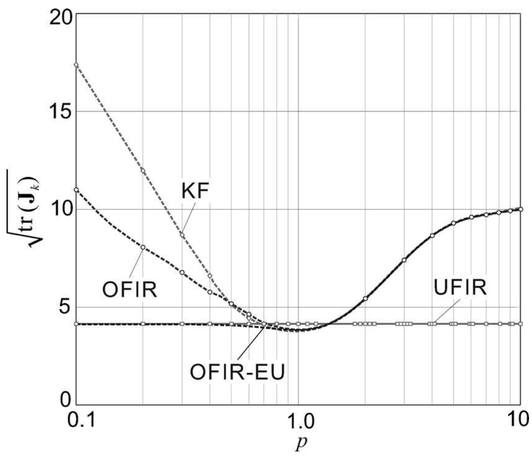

Fig. 5 *RMSEs* $\sqrt{trJ_k}$ *as functions of p for KF and FIR filters*

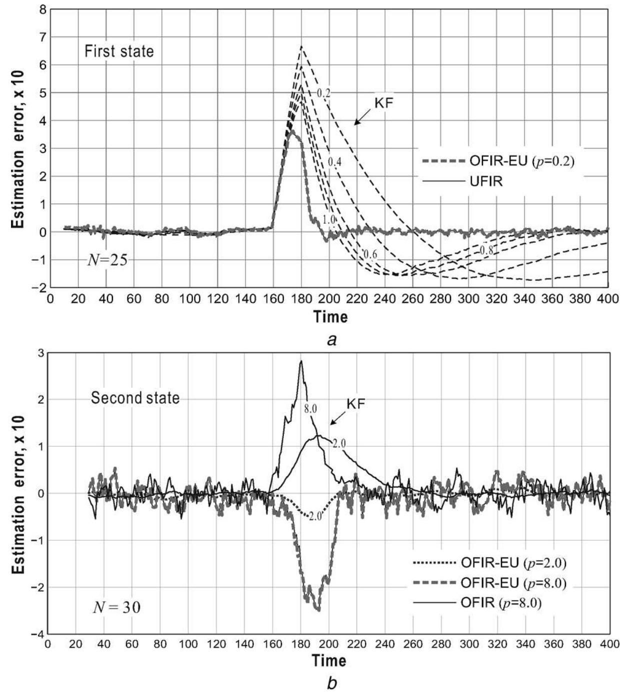

Fig. 6 *Effect of the model temporary uncertainty on the estimation errors*
a First state for $p \leqslant 1$
b Second state for $p > 1$

tracked and in some others ignored by an estimator. To learn the effect of the model temporary uncertainty on the estimates, in this section we set $\tau = 5$ s when $160 \leqslant k \leqslant 180$ and $\tau = 0.1$ s otherwise. The noise variances are allowed to be $\sigma_{w1}^2 = 10^{-2}$, $\sigma_{w2}^2 = 10^{-2}/s^2$, and $\sigma_v^2 = 25$. The process was simulated at 400 subsequent points and we found $N_{opt} = 25$ for the first state and $N_{opt} = 30$ for the second state.

Typical estimates are sketched in Fig. 6. As can be seen in Fig. 6*a*, the OFIR-EU (or OUFIR) filter (case $p = 0.2$) and the UFIR filter produce similar errors. However, the KF demonstrates much worse robustness for any $p \leqslant 1$. Otherwise, with $p > 1$, no preference can be granted to any of the filters (Fig. 6*b*). Note that the same deduction flows from Fig. 5.

8.2 Harmonic model

Oscillating systems and processes corrupted by noise are well modelled by a harmonic model in diverse environments. Referring to the properties of FIR filters discussed in Section 8.1 based on a polynomial model, one may wonder if they are saved in the harmonic model. To ascertain this, we test the FIR filters by a two-state harmonic model allowing $\boldsymbol{B} = [11]^{\mathrm{T}}$, $\boldsymbol{C} = [10]$, $\boldsymbol{D} = 1$, and

$$\boldsymbol{A} = \begin{bmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{bmatrix}$$

with $\varphi = \pi/32$. Traditionally, we investigate an ideal case of completely known model which fits the process exactly and the cases of imprecisely defined noise statistics and systems with temporary uncertaainties.

8.2.1 Accurate model – ideal case: To learn the estimation errors in the ideal case, we draw 400 samples for the initial states $x_{10} = 1$ and $x_{20} = 0.1$ and noise variances $\sigma_w^2 = 1$ and $\sigma_v^2 = 10$. Supposing that all these values are known exactly, we provide state estimation and compute the RMSEs $\sqrt{\mathrm{tr}\boldsymbol{J}_k}$ as functions of N (Fig. 7*a*). Referring to [22], we notice that the properties of FIR filters learned based on the polynomial model have similar appearances in the harmonic models. In fact, the OFIR-EU (or OUFIR) estimate still evolves as UFIR-to-OFIR by changing N around N_{opt}. The only specific can be noticed when $N > N_{\text{opt}}$: the UFIR estimate grows with oscillations in the harmonic model and it grows monotonously in the polynomial model.

8.2.2 Filtering with error in the noise statistics: The case of imprecisely defined noise statistics does not reveal new features. In fact, the RMSEs $\sqrt{\text{tr}\boldsymbol{J}_k}$ computed as functions of the correction coefficient p behave in Fig. 7*b* very similarly to those shown in Fig. 5 for the polynomial model. Functions shown in this figure basically confirm our early statements: with $p < 1$, the OFIR-EU and OUFIR filters behave as the UFIR filter and, if $p > 1$, the KF, OFIR, OFIR-EU, and OUFIR estimates converge and their MSEs grow.

8.2.3 Filtering under the model temporary uncertainties: We finally would like to learn the effect of a model temporary uncertainty on the estimates. With this purpose in mind, we augment the system matrix $\boldsymbol{A}$ as

$$\boldsymbol{A} = \begin{bmatrix} \cos\varphi & \sin\varphi + \delta \\ -\sin\varphi + \delta & \cos\varphi \end{bmatrix},$$

where we set $\delta = 0.4$ if $160 \leqslant k \leqslant 180$ and $\delta = 0$ otherwise. The process is generated with $x_{10} = 1$, $x_{20} = 0.1$, $\sigma_w^2 = 0.1$, and $\sigma_v^2 = 100$ at 400 subsequent points.

The instantaneous estimation errors produced by the KF and OFIR-EU (or OUFIR) filter for $p \leqslant 1$ are shown in Fig. 8. It is seen that all the filtering estimates have negligible errors in the interval of first 160 points where the model is supposed to be known exactly. Beyond this interval, the performance of all filters is deteriorated by $\delta_k \neq 0$. One also notices that after the model returned back to normal mode of $\delta = 0$, the filters demonstrate transients with highly perceptible differences. Inherently, the transients in OFIR-EU and OUFIR filters are limited by the horizon length, which in our case was found to be $N_{\text{opt}} = 19$. Just on the contrary, the transient in the KF which has IIR lasts much longer as shown in Figs. 8 and 9. The case of $p > 1$ in Fig. 9 slightly corrects the above-made observations. Here, we see larger excursions in the OFIR-EU (or OUFIR) and OFIR filters. Nevertheless, the KF transient still lasts much longer.

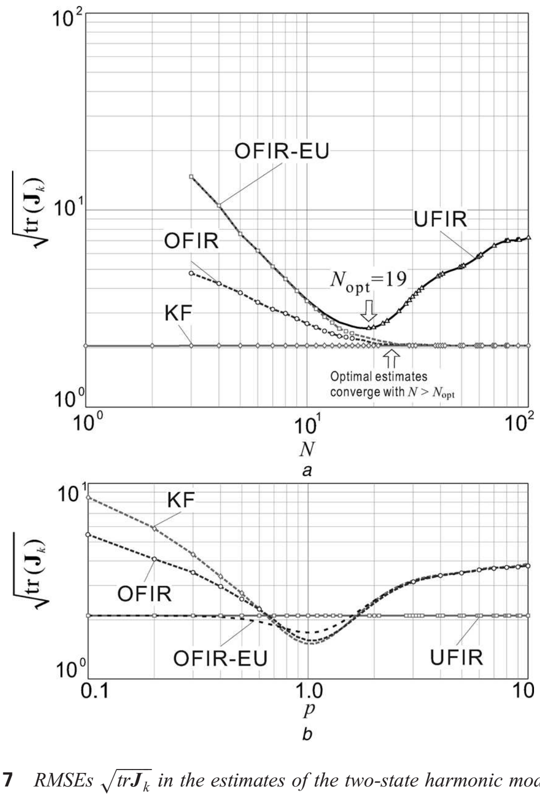

Fig. 7 *RMSEs $\sqrt{tr\boldsymbol{J}_k}$ in the estimates of the two-state harmonic model*

a As function of N
b As function of p for $N_{\text{opt}} = 19$

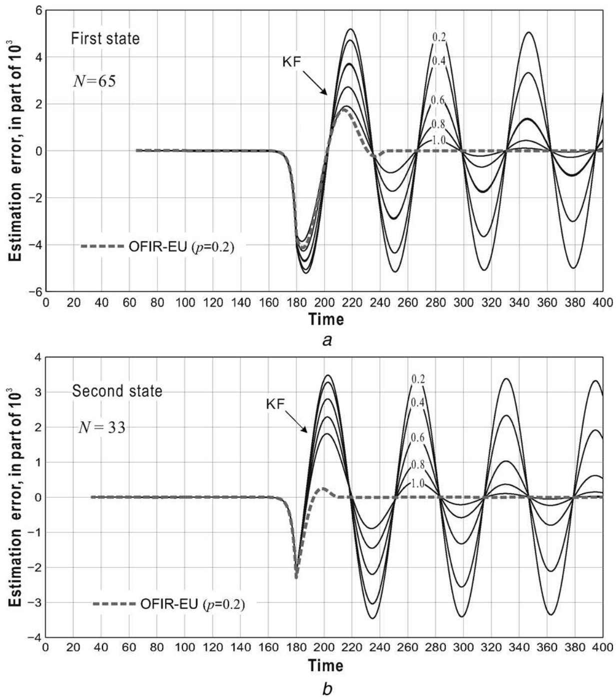

Fig. 8 *Instantaneous estimation errors caused by the temporary model uncertainties with $p \leq 1$*

a First state
b Second state

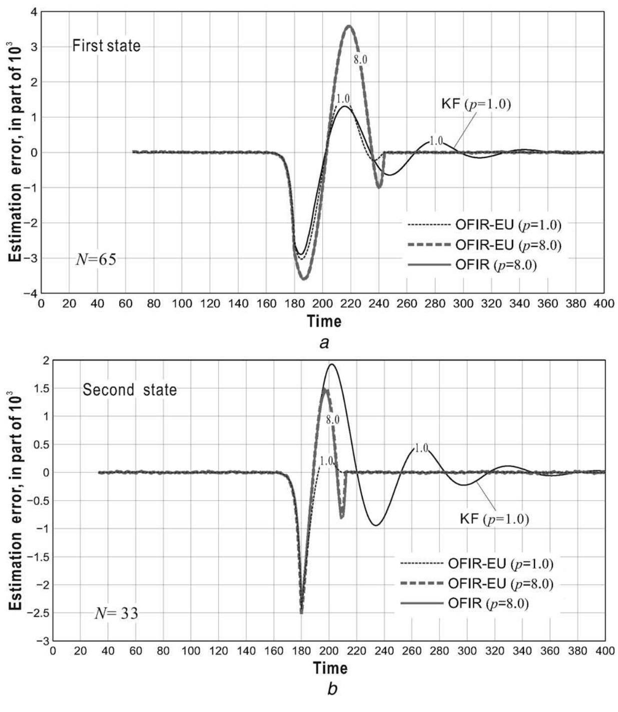

Fig. 9 *Instantaneous estimation errors caused by the temporary model uncertainties with $p > 1$*

a First state
b Second state

9 Conclusions

A survey provided in this paper overviews the trade-off between the UFIR, OFIR, OFIR-EU, and OUFIR filters. It shows that the filters have diverse levels of computational complexities and produce different estimation errors as functions of the horizon N, mismodelling, and environmental conditions. Most appealing from the engineering perspective is the UFIR filter that does not require the noise statistics. it also has a simplest algorithm with lowest computational burden and its MSE is minimised by N_{opt}. However, the UFIR filter produces larger errors than in KF, because it does not guarantee optimality. The output of the OFIR filter is consistent to the KF output. The OFIR and KF estimates become virtually equal when $N > N_{opt}$ and exactly equal when $N \to \infty$. The latter is due to the fact that $N = \infty$ makes the OFIR filter has optimal IIR and the KF has IIR as well.

The intermediate solutions such as the OFIR-EU and OUFIR filters are equivalent to each other. In terms of accuracy, they stay between the UFIR and OFIR filters and evolve from one to another by changing the conditions. Of practical importance is that the OFIR-EU and OUFIR filters adopt properties of both the UFIR and OFIR filters. In fact, like the UFIR filter, the OFIR-EU (or OUFIR) filter completely ignores the initial conditions. Unlike in the UFIR filter which MSE is minimised by N_{opt}, the MSEs in the OFIR-EU filter decrease with N and this filter is thus full-horizon, as well as the OFIR filter. The performance of OFIR-EU filter is developed by varying the horizon length 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.

10 Acknowledgment

This investigation was jointly supported by the Royal Academy of Engineering under the Newton Research Collaboration Programme NRCP/1415/140, National Natural Science Foundation of China (Grants No.: 61273087), and the 111 Project (B12018). The authors would like to thank the reviewers for their valuable comments and suggestions to improve the quality of this paper.

11 References

1 Gauss, C.F.: 'Theory of the combination of observations least subject to errors' (SIAM Publication, Philadelphia, 1995), Translated by G. W. Stewart
2 Stark, H., Woods, J.W.: 'Probability, random processes, and estimation theory for engineers' (Prentice Hall, Upper Saddle River, NJ, 1994, 2nd edn.)
3 Stapleton, J.H.: 'Linear statistical models' (Wiley, New York, 2009, 2nd edn.)
4 Aitken, A.C.: 'On least squares and linear combinations of observations'. Proc. of the Royal Society of Edinburgh, 1935, vol. 55, pp. 42–48
5 Kay, S.M.: 'Fundamentals of statistical signal processing' (Prentice Hall, New York, 2001)
6 Kim, P.S., Lee, M.E.: 'A new FIR filter for state estimation and its applications', *J. Comput. Sci. Technol.*, 2007, **22**, (5), pp. 779–784
7 Shmaliy, Y.S.: 'An unbiased FIR filter for TIE model of a local clock in applications to GPS-based timekeeping', *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 2006, **53**, (5), pp. 862–870
8 Gibbs, B.P.: 'Advanced Kalman filtering, least-squares and modeling' (John Wiley & Sons, Hoboken, NJ, 2011)
9 Hardy, M.: 'An illuminating counterexample', *Am. Math. Mon.*, 2003, **110**, (3), pp. 232–238
10 Simon, D.: 'Optimal state estimation: Kalman, Hinf, and nonlinear approaches' (John Wiley & Sons, Honboken, NJ, 2006)
11 Liu, Q., Wang, Z., He, X., *et al.*: 'Event-based recursive distributed filtering over wireless sensor networks', *IEEE Trans. Autom. Control*, 2015, **60**, (9), pp. 2470–2475
12 Dong, H., Wang, Z., Ding, S.X., *et al.*: 'Event-based H-infinity filter design for a class of nonlinear time-varying systems with fading channels and multiplicative noises', *IEEE Trans. Signal Process.*, 2015, **63**, (13), pp. 3387–3395
13 Zhang, S., Wang, Z., Ding, D., *et al.*: 'On design of fault detection filter in finite frequency domain with regional pole assignment', *IEEE Trans. Circuits Syst. II*, 2015, **62**, (4), pp. 382–386
14 Jazwinski, A.H.: 'Stochastic processes and filtering theory' (Academic, New York, 1970)
15 Kwon, W.H., Han, S.: 'Receding horizon control: model predictive control for state models' (Springer, London, UK, 2005)
16 Shmaliy, Y.S.: 'Linear optimal FIR estimation of discrete time-invariant state-space models', *IEEE Trans. Signal Process.*, 2010, **58**, (6), pp. 3086–3096
17 Johnson, K.R.: 'Optimum, linear, discrete filtering of signals containing a nonrandom component', *IRE Trans. Inf. Theory*, 1956, **2**, (2), pp. 49–55
18 Jazwinski, A.H.: 'Limited memory optimal filtering', *IEEE Trans. Autom. Control*, 1968, **13**, (10), pp. 558–563
19 Kwon, O.K., Kwon, W.H., Lee, K.S.: 'FIR filters and recursive forms for discrete-time state-space models', *Automatica*, 1989, **25**, (5), pp. 715–728
20 Kwon, W.H., Kim, P.S., Park, P.: 'A receding horizon Kalman FIR filter for discrete time-invariant systems', *IEEE Trans. Autom. Control*, 1999, **99**, (9), pp. 1787–1791
21 Kwon, W.H., Kim, P.S., Han, S.: 'A receding horizon unbiased FIR filter for discrete-time state space models', *Automatica*, 2002, **38**, (3), pp. 545–551
22 Shmaliy, Y.S.: 'An iterative Kalman-like algorithm ignoring noise and initial conditions', *IEEE Trans. Signal Process.*, 2011, **59**, (6), pp. 2465–2473
23 Shmaliy, Y.S.: 'Optimal gains of FIR estimations for a class of discrete-time state-space models', *IEEE Trans. Signal Process. Lett.*, 2008, **15**, pp. 517–520
24 Ahn, C.K.: 'Strictly passive FIR filtering for state-space models with external disturbance', *Int. J. Electron. Commun.*, 2012, **66**, (11), pp. 944–948
25 Ahn, C.K., Kim, P.S.: 'Fixed-lag maximum likelihood FIR smoother for state-space models', *IEICE Electron. Express*, 2008, **5**, (1), pp. 11–16
26 Mirkin, L.: 'On the H1 fixed-lag smoothing: how to exploit the information preview', *Automatica*, 2003, **39**, (8), pp. 1495–1504
27 Shmaliy, Y.S., Mendoza, L.J.M.: 'FIR smoothing of discrete-time polynomial signals in state space', *IEEE Trans. Signal Process.*, 2010, **58**, (5), pp. 2544–2555
28 Kwon, B.K., Han, S., Kim, O.K., *et al.*: 'Minimum variance FIR smoothers for discrete-time state space models', *IEEE Trans. Signal Process. Lett.*, 2007, **14**, (8), pp. 557–560
29 Liu, D., Liu, X.: 'Optimal state estimation without the requirement of a priori statistics information of the initial state', *IEEE Trans. Autom. Control*, 1994, **39**, (10), pp. 2087–2091
30 Ling, K.V., Lim, K.W.: 'Receding horizon recursive state estimation', *IEEE Trans. Autom. Control*, 1999, **44**, (9), pp. 1750–1753
31 Makhoul, J.: 'Linear prediction: a tutorial review', *Proc. IEEE*, 1975, **63**, pp. 561–580
32 Iwata, T., Imae, M., Suzuyama, T., *et al.*: 'Simulation and ground experiments of remote synchronization system for on-board crystal oscillator of quazi-zenith satellite system', *J. Inst. Navig.*, 2006, **53**, pp. 231–235
33 Levine, J.: 'Invited review article: the statistical modeling of atomic clocks and the design of time scales', *Rev. Sci. Instrum.*, 2012, **83**, (2), p. 021101
34 Kou, Y., Jiao, Y., Xu, D., *et al.*: 'Low-cost precise measurement of oscillator frequency instability based on GNSS carrier observation', *Adv. Space Res.*, 2013, **51**, (6), pp. 969–977
35 Choi, J.-W., Han, S., Cioffi, J.M.: 'An FIR channel estimation filter with robustness to channel mismatch condition', *IEEE Trans. Broadcast.*, 2008, **54**, (1), pp. 127–130
36 Ding, D., Wang, Z., Shen, B., *et al.*: 'Envelope-constrained H-infinity filtering with fading measurements and randomly occurring nonlinearities: the finite horizon case', *Automatica*, 2015, **55**, pp. 37–45
37 Dong, H., Wang, Z., Ding, S.X., *et al.*: 'Finite-horizon estimation of randomly occurring faults for a class of nonlinear time-varying systems', *Automatica*, 2014, **50**, (12), pp. 3182–3189
38 Ahn, C.K., Han, S., Kwon, W.H.: 'H_∞ FIR filters for linear continuous-time state-space systems', *IEEE Signal Process. Lett.*, 2006, **13**, (9), pp. 557–560
39 Salmi, J., Richter, A., Koivunen, V.: 'Detection and tracking of MIMO propagation path parameters using state-space approach', *IEEE Trans. Signal Process.*, 2009, **57**, (4), pp. 1538–1550
40 Nevat, I., Yuan, J.: 'Joint channel tracking and decoding for BICM-OFDM systems using consistency test and adaptive detection selection', *IEEE Trans. Veh. Technol.*, 2009, **58**, (8), pp. 4316–4328
41 Shmaliy, Y.S., Manzano, O.I.: 'Time-variant linear optimal finite impulse response estimator for discrete state-space models', *Int. J. Adapt. Control Signal Process.*, 2012, **26**, (2), pp. 95–104
42 Shmaliy, Y.S.: 'Unbiased FIR filtering of discrete-time polynomial state-space models', *IEEE Trans. Signal Process.*, 2009, **57**, (4), pp. 1241–1249
43 Shmaliy, Y.S.: 'Suboptimal FIR filtering of nonlinear models in additive white Gaussian noise', *IEEE Trans. Signal Process.*, 2012, **60**, (10), pp. 5519–5527
44 Simon, D., Shmaliy, Y.S.: 'Unified forms for Kalman and finite impulse response filtering and smoothing', *Automatica*, 2013, **49**, (6), pp. 1892–1899
45 Echeverria, F.R., Sarr, A., Shmaliy, Y.S.: 'Optimal memory for discrete-time FIR filters in state-space', *IEEE Trans. Signal Process.*, 2014, **62**, pp. 557–561
46 Zhao, S., Shmaliy, Y.S., Liu, F.: 'Fast computation of discrete optimal FIR estimates in white Gaussian noise', *IEEE Signal Process. Lett.*, 2015, **22**, (6), pp. 718–722
47 Zhao, S., Shmaliy, Y.S., Liu, F., *et al.*: 'Effect of embedded unbiasedness on discrete-time optimal FIR filtering estimates', *EURASIP J. Adv. Signal Process.*, 2015, **2015**, (1), pp. 1–13
48 Zhao, S., Shmaliy, Y.S., Huang, B., *et al.*: 'Minimum variance unbiased FIR filter for discrete time-variant models', *Automatica*, 2015, **53**, pp. 355–361
49 Lee, Y.S., Han, S.: 'A minimum variance FIR filter with an N_∞ error bound and its application to the current measuring circuitry', *Measurement*, 2014, **50**, pp. 115–120
50 Kim, P.S.: 'An alternative FIR filter for state estimation in discrete-time systems', *Digit. Signal Process.*, 2010, **20**, (3), pp. 935–943
51 Shmaliy, Y.S., Simon, D.: 'Iterative unbiased FIR state estimation: a review of algorithms', *EURASIP J. Adv. Signal Process.*, 2013, **113**, (1), pp. 1–16