RESEARCH   **Open Access**

Optimal and unbiased FIR filtering in discrete time state space with smoothing and predictive properties

Yuriy S Shmaliy[*] and Oscar Ibarra-Manzano

Abstract

We address *p*-shift finite impulse response optimal (OFIR) and unbiased (UFIR) algorithms for predictive filtering ($p > 0$), filtering ($p = 0$), and smoothing filtering ($p < 0$) at a discrete point *n* over *N* neighboring points. The algorithms were designed for linear time-invariant state-space signal models with white Gaussian noise. The OFIR filter self-determines the initial mean square state function by solving the discrete algebraic Riccati equation. The UFIR one represented both in the batch and iterative Kalman-like forms does not require the noise covariances and initial errors. An example of applications is given for smoothing and predictive filtering of a two-state polynomial model. Based upon this example, we show that exact optimality is redundant when $N \gg 1$ and still a nice suboptimal estimate can fairly be provided with a UFIR filter at a much lower cost.

Keywords: State estimation, Optimal FIR filter, Unbiased FIR filter, Kalman-like algorithm

Introduction

There is a class of estimation problems requiring optimal filtering at a discrete-time current point *n* employing measurement on an averaging interval (horizon) of preceding or/and succeeding neighboring but not obligatorily nearest *N* points. To solve such problems, filtering is usually organized employing the finite impulse response (FIR) structures. Because the averaging interval can be placed with an arbitrary time shift *p* with respect to *n*, there can be recognized three kinds of *p*-shift FIR filters as shown in Figure 1, namely the *p*-step predictive filter ($p > 0$), filter ($p = 0$), and |*p*|-lag smoothing filter ($p < 0$).

Predictive FIR filtering is fundamental for discrete-time feedback systems and required in signal processing when measurement is temporary unavailable in the nearest past of *p* points. The one-step predictive filter known as the receding horizon filter has been put into the concept of the receding horizon (or model predictive) control [1,2]. For polynomial signals, an unbiased predictive FIR filter was proposed by Heinonen and Neuvo in [3]. Further, this

filter was investigated by many authors [4] and developed in state-space to *p*-step predictive filtering [5].

Smoothing FIR filtering is a key solution whenever denoising of signals is required with highest efficiency. Savitzky-Golay smoothing filter [6] is one of the most popular here. In recent decades, we meet a few new substantial results. Linear FIR smoothers were developed and used by Zhou and Wang in the FIR-median hybrid filters [7]. In state space, order-recursive FIR smoothers were proposed by Yuan and Stuller in [8]. Most recently, the general receding horizon FIR smoother theory has been developed in [9,10] and, for polynomial signals, the |*p*|-lag smoothing FIR filter theory addressed in [11].

It follows from the above-given short survey that the authors prefer solving the problems of filtering, smoothing, and prediction employing different algorithms. In [12,13], a universal scheme has been proposed for the *p*-shift FIR estimators (filters, predictors, and smoothers). Still no universal solution has been addressed in state space for FIR filtering with smoothing and prediction properties.

In this article, we follow the approach developed in [12] and address universal *p*-shift optimal FIR (OFIR)

*Correspondence: shmaliy@ugto.mx
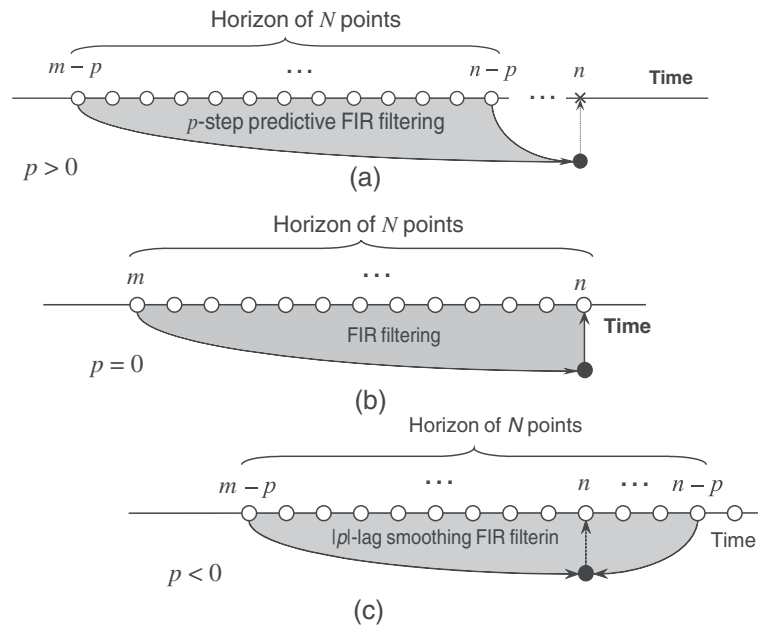Electronics Department, DICIS, Universidad de Guanajuato, Salamanca, 36855, Gto., Mexico

Figure 1 FIR filtering at a discrete point *n*. (a) *p*-step predictive filtering ($p > 0$), (**b**) filtering ($p = 0$), and (**c**) $|p|$-lag smoothing filtering ($p < 0$). Measurement is organized on an interval of *N* points from $m - p$ to $n - p$, where $m = n - N + 1$.

and unbiased FIR (UFIR) filters for predictive filtering ($p > 0$), filtering ($p = 0$), and smoothing filtering ($p < 0$) at a current point *n* of linear discrete-time-invariant state-space models with white noise. The rest of the article is organized as follows. In Section 'Signal model and problem formulation', we describe the model and formulate the problem. The *p*-shift OFIR filter is derived in Section '*p*-Shift OFIR filter with predictive and smoothing properties'. Here, we also find its gain and estimate the initial mean square state function. The UFIR filter is considered in detain in Section '*p*-shift UFIR filter with predictive and smoothing properties' along with the estimation error. An application to the two-state model is given in Section 'Applications' and concluding remarks are drawn in Section 'Conclusion'.

Signal model and problem formulation

Consider a class of discrete time-invariant linear signal models represented in state space with the state and observation equations, respectively,

$$\mathbf{x}_n = \mathbf{A}\mathbf{x}_{n-1} + \mathbf{B}\mathbf{w}_n, \tag{1}$$

$$\mathbf{y}_n = \mathbf{C}\mathbf{x}_n + \mathbf{D}\mathbf{v}_n, \tag{2}$$

where $\mathbf{x}_n \in \mathcal{R}^K$ and $\mathbf{y}_n \in \mathcal{R}^M$ are the state and observation vectors, respectively,

$$\mathbf{x}_n = [x_{1n} \, x_{2n} \, \ldots \, x_{Kn}]^T, \tag{3}$$

$$\mathbf{y}_n = [y_{1n} \, y_{2n} \, \ldots \, y_{Mn}]^T. \tag{4}$$

Here, $\mathbf{A} \in \mathcal{R}^{K \times K}$, $\mathbf{B} \in \mathcal{R}^{K \times P}$, $\mathbf{C} \in \mathcal{R}^{M \times K}$, and $\mathbf{D} \in \mathcal{R}^{M \times M}$. The system noise vector $\mathbf{w}_n \in \mathcal{R}^P$ and the measurement noise vector $\mathbf{v}_n \in \mathcal{R}^M$, respectively,

$$\mathbf{w}_n = [w_{1n} \, w_{2n} \, \ldots \, w_{Pn}]^T, \tag{5}$$

$$\mathbf{v}_n = [v_{1n} \, v_{2n} \, \ldots \, v_{Mn}]^T, \tag{6}$$

are white Gaussian with zero mean components, $E\{\mathbf{w}_n\} = \mathbf{0}$ and $E\{\mathbf{v}_n\} = \mathbf{0}$. It is implied that $\mathbf{w}_n$ and $\mathbf{v}_n$ are mutually uncorrelated, $E\{\mathbf{w}_i\mathbf{v}_j^T\} = \mathbf{0}$ for all *i* and *j*, and have the covariances, respectively,

$$\mathbf{R} = E\{\mathbf{w}_n\mathbf{w}_n^T\}, \tag{7}$$

$$\mathbf{Q} = E\{\mathbf{v}_n\mathbf{v}_n^T\}. \tag{8}$$

The problem now formulates as follows. Given the model (1) and (2), we would like to derive a *p*-shift OFIR filter covering the problems of predictive filtering ($p > 0$), filtering ($p = 0$), and smoothing filtering ($p < 0$) as shown in Figure 1. We also wish to find its unbiased version, represent it in the iterative Kalman-like form, and investigate errors based on a typical example.

p-Shift OFIR filter with predictive and smoothing properties

A *p*-shift OFIR filter can be derived following Figure 1, if to represent (1) and (2) on a horizon of *N* points, similarly

to [11], with recursively computed forward-in-time solutions [14] as follows, respectively,

$$\mathbf{X}_{n-p,m-p} = \mathbf{A}_{N-1}\mathbf{x}_{m-p} + \mathbf{B}_{N-1}\mathbf{W}_{n-p,m-p}, \qquad (9)$$

$$\mathbf{Y}_{n-p,m-p} = \mathbf{C}_{N-1}\mathbf{x}_{m-p} + \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p}$$
$$+ \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p}, \qquad (10)$$

where $\mathbf{X}_{n,m} \in \mathcal{R}^{KN}$, $\mathbf{Y}_{n,m} \in \mathcal{R}^{MN}$, $\mathbf{W}_{n,m} \in \mathcal{R}^{PN}$, and $\mathbf{V}_{n,m} \in \mathcal{R}^{MN}$ are given by, respectively,

$$\mathbf{X}_{n-p,m-p} = \left[\mathbf{x}_{n-p}^T \, \mathbf{x}_{n-1-p}^T \, \cdots \, \mathbf{x}_{m-p}^T \right]^T, \qquad (11)$$

$$\mathbf{Y}_{n-p,m-p} = \left[\mathbf{y}_{n-p}^T \, \mathbf{y}_{n-1-p}^T \, \cdots \, \mathbf{y}_{m-p}^T \right]^T, \qquad (12)$$

$$\mathbf{W}_{n-p,m-p} = \left[\mathbf{w}_{n-p}^T \, \mathbf{w}_{n-1-p}^T \, \cdots \, \mathbf{w}_{m-p}^T \right]^T, \qquad (13)$$

$$\mathbf{V}_{n-p,m-p} = \left[\mathbf{v}_{n-p}^T \, \mathbf{v}_{n-1-p}^T \, \cdots \, \mathbf{v}_{m-p}^T \right]^T. \qquad (14)$$

The matrices $\mathbf{A}_{N-1} \in \mathcal{R}^{KN \times K}$, $\mathbf{B}_{N-1} \in \mathcal{R}^{KN \times PN}$, $\mathbf{C}_{N-1} \in \mathcal{R}^{MN \times K}$, $\mathbf{G}_{N-1} \in \mathcal{R}^{MN \times PN}$, and $\mathbf{D}_{N-1} \in \mathcal{R}^{MN \times MN}$ are specified with, respectively,

$$\mathbf{A}_i = \left[(\mathbf{A}^i)^T \, (\mathbf{A}^{i-1})^T \, \cdots \, \mathbf{A}^T \, \mathbf{I} \right]^T, \qquad (15)$$

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{i-1}\mathbf{B} & \mathbf{A}^i\mathbf{B} \\ \mathbf{0} & \mathbf{B} & \cdots & \mathbf{A}^{i-2}\mathbf{B} & \mathbf{A}^{i-1}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B} & \mathbf{AB} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{B} \end{bmatrix}, \qquad (16)$$

$$\mathbf{C}_i = \bar{\mathbf{C}}_i \mathbf{A}_i, \qquad (17)$$

$$\mathbf{G}_i = \bar{\mathbf{C}}_i \mathbf{B}_i, \qquad (18)$$

$$\mathbf{D}_i = \mathrm{diag}\left(\underbrace{\mathbf{D}\ \mathbf{D}\ \cdots\ \mathbf{D}}_{i+1} \right), \qquad (19)$$

where we have assigned $\bar{\mathbf{C}}_i = \mathrm{diag}\left(\underbrace{\mathbf{C}\ \mathbf{C}\ \cdots\ \mathbf{C}}_{i+1} \right)$.

In this model, the initial state $\mathbf{x}_{m-p}$ is supposed to be known exactly and $\mathbf{w}_{m-p}$ is thus allowed to have zero components.

Optimal gain

One can now assign the gain matrix $\mathbf{H}(p) \triangleq \mathbf{H}(p, n, m) \in \mathcal{R}^{K \times MN}$ implementing the convolution principle and find the filtering estimate[a] of $\mathbf{x}_n$ as

$$\tilde{\mathbf{x}}_{n|n-p} = \mathbf{H}(p)\mathbf{Y}_{n-p,m-p} \qquad (20)$$
$$= \mathbf{H}(p)[\, \mathbf{C}_{N-1}\mathbf{x}_{m-p} + \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p}$$
$$+ \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p} \,]. \qquad (21)$$

For $\mathbf{H}(p)$ to be optimal in the minimize mean square error (MSE) sense, the following cost function needs to be minimized,

$$\mathbf{J}(p) = E\{(\mathbf{x}_n - \tilde{\mathbf{x}}_{n|n-p})(\mathbf{x}_n - \tilde{\mathbf{x}}_{n|n-p})^T\}$$
$$= E\{[\, \mathbf{x}_n - \mathbf{H}(p)(\mathbf{C}_{N-1}\mathbf{x}_{m-p} + \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p}$$
$$+ \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p})]\,[\,\mathbf{x}_n - \mathbf{H}(p)(\mathbf{C}_{N-1}\mathbf{x}_{m-p}$$
$$+ \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p} + \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p})]^T\}, \qquad (22)$$

where $E(x)$ means an average of x. The minimization can be provided employing the orthogonality condition [14] in the form of [12],

$$\mathbf{0} = E\{[\, \mathbf{x}_n - \hat{\mathbf{H}}(p)(\mathbf{C}_{N-1}\mathbf{x}_{m-p} + \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p}$$
$$+ \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p})]\,(\mathbf{C}_{N-1}\mathbf{x}_{m-p}$$
$$+ \mathbf{G}_{N-1}\mathbf{W}_{n-p,m-p} + \mathbf{D}_{N-1}\mathbf{V}_{n-p,m-p})^T\}, \qquad (23)$$

to produce the optimal gain matrix $\hat{\mathbf{H}}(p)$. In doing so, one needs substituting $\mathbf{x}_n$ with the first vector row[b] in (9); that is,

$$\mathbf{x}_n = \mathbf{A}^{N-1+p}\mathbf{x}_{m-p} + \bar{\mathbf{B}}_{N-1}\mathbf{W}_{n-p,m-p}, \qquad (24)$$

where $\bar{\mathbf{B}}_{N-1}$ is the first vector row in (16).

Substituting (24) to (23) and supposing that the initial state and measurement noise are mutually uncorrelated for all p, we provide the averaging in (23) and arrive at the optimal gain matrix

$$\hat{\mathbf{H}}(p) = (\mathbf{A}^{N-1+p}\mathbf{R}_{m-p}\mathbf{C}_{N-1}^T + \bar{\mathbf{Z}}_w)$$
$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}, \qquad (25)$$

in which $\bar{\mathbf{Z}}_w = \bar{\mathbf{B}}_{N-1}\Psi\mathbf{G}_{N-1}^T$,

$$\mathbf{Z}_{m-p} = \mathbf{C}_{N-1}\mathbf{R}_{m-p}\mathbf{C}_{N-1}^T, \qquad (26)$$

$$\tilde{\mathbf{Z}}_w = \mathbf{G}_{N-1}\Psi\mathbf{G}_{N-1}^T, \qquad (27)$$

$$\tilde{\mathbf{Z}}_v = \mathbf{D}_{N-1}\Phi\mathbf{D}_{N-1}^T, \qquad (28)$$

the mean square initial state is specified by

$$\mathbf{R}_{m-p} = E\left\{ \mathbf{x}_{m-p}\mathbf{x}_{m-p}^T \right\}, \qquad (29)$$

and the signal and measurement white noise covariance functions are formed as, respectively,

$$\Psi = E\left\{\mathbf{W}_{n-p,m-p}\mathbf{W}_{n-p,m-p}^T\right\}$$

$$= \mathrm{diag}\big(\underbrace{\mathbf{R}\ \mathbf{R}\ \dots\ \mathbf{R}}_{N}\big), \tag{30}$$

$$\Phi = E\left\{\mathbf{V}_{n-p,m-p}\mathbf{V}_{n-p,m-p}^T\right\}$$

$$= \mathrm{diag}\big(\underbrace{\mathbf{Q}\ \mathbf{Q}\ \dots\ \mathbf{Q}}_{M}\big). \tag{31}$$

By multiplying $\mathbf{R}_{m-p}$ in (25) from the left-hand side with the identity matrix $(\mathbf{C}_{n-m}^T\mathbf{C}_{n-m})^{-1}\mathbf{C}_{n-m}^T\mathbf{C}_{n-m}$, we have finally

$$\hat{\mathbf{H}}(p) = [\bar{\mathbf{H}}(p)\mathbf{Z}_{m-p} + \bar{\mathbf{Z}}_w]\,(\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}, \tag{32}$$

where, by $n - m = N - 1$, the unbiased gain attains two equivalent forms of

$$\bar{\mathbf{H}}(p) = \mathbf{A}^{N-1+p}(\mathbf{C}_{n-m}^T\mathbf{C}_{n-m})^{-1}\mathbf{C}_{n-m}^T \tag{33}$$

$$= \mathbf{A}^{N-1+p}(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T. \tag{34}$$

Note that $\bar{\mathbf{H}}(p)$ satisfies the unbiasedness condition

$$E\{\hat{\mathbf{x}}_{n|n-p}\} = E\{\mathbf{x}_n\} \tag{35}$$

and has an important applied property: it does not depend on noise and initial errors, although it is p- and N-dependent.

As shown in Appendix, matrix $\mathbf{Z}_{m-p}$ representing in (32) the mean square initial state $\mathbf{R}_{m-p}$ on an averaging interval of N points can optimally be estimated by solving the discrete algebraic Riccati equation (DARE)

$$\mathbf{0} = \mathbf{Z}_{m-p}(\tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Z}_{m-p} + 2\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v$$

$$- \mathbf{Y}_{n-p,m-p}\mathbf{Y}_{n-p,m-p}^T(\tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Z}_{m-p}, \tag{36}$$

whose analytic solution can be found following [15]. We notice that this equation also serves for filtering out all of the noise components [12].

Optimal filtering estimate

Determined $\mathbf{Z}_{m-p}$, by (36), the p-shift OFIR filtering estimate $\hat{\mathbf{x}}_{n|n-p}$ can now be generalized as follows. Given (1) and (2) with uncorrelated zero-mean white noise vectors $\mathbf{w}_n$ and $\mathbf{v}_n$. Then p-step OFIR predictive filtering ($p > 0$), filtering ($p = 0$), and $|p|$-lag smoothing filtering ($p < 0$) can be provided at n employing data taken from $m - p$ to $n - p$ by

$$\hat{\mathbf{x}}_{n|n-p} = \hat{\mathbf{H}}(p)\mathbf{Y}_{n-p,m-p} \tag{37}$$

$$= [\bar{\mathbf{H}}(p)\mathbf{Z}_{m-p} + \bar{\mathbf{Z}}_w]\,(\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}$$

$$\times \mathbf{Y}_{n-p,m-p} \tag{38}$$

$$= [\mathbf{A}^{N-1+p}(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T\mathbf{Z}_{m-p} + \bar{\mathbf{Z}}_w]$$

$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Y}_{n-p,m-p}, \tag{39}$$

where $\mathbf{Y}_{n-p,m-p}$ is the data vector (12), $\mathbf{C}_i$ is given by (17), and $\bar{\mathbf{Z}}_w$, $\tilde{\mathbf{Z}}_w$, and $\tilde{\mathbf{Z}}_v$ are specified for (25). The algorithm should be applied to any $N \geqslant K$, in order to avoid problems with singularities. Note that K is typically not larger in state space modeling.

p-shift UFIR filter with predictive and smoothing properties

There are at least two cases when exact optimality is redundant and OFIR can fairly be substituted with UFIR at much lower price to produce still a nice near optimal estimate [12]. In fact, if $\mathbf{Z}_{m-n}$ substantially dominates $\bar{\mathbf{Z}}_w$, $\tilde{\mathbf{Z}}_w$, and $\tilde{\mathbf{Z}}_v$ in the order of magnitudes for all p, we have $\hat{\mathbf{H}}(p) \cong \bar{\mathbf{H}}(p)$. The same effect is achieved with $N \gg 1$.

Thus, the UFIR filter should also be generalized. Given (1) and (2) with uncorrelated zero-mean white noise components $\mathbf{w}_n$ and $\mathbf{v}_n$. Then p-step UFIR predictive filtering ($p > 0$), filtering ($p = 0$), and $|p|$-lag smoothing filtering ($p < 0$) can be provided at n employing data taken from $m - p$ to $n - p$ by

$$\bar{\mathbf{x}}_{n|n-p} = \bar{\mathbf{H}}(p)\mathbf{Y}_{n-p,m-p} \tag{40}$$

$$= \mathbf{A}^{N-1+p}(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T\mathbf{Y}_{n-p,m-p}. \tag{41}$$

Note that both (40) and (41) follow from (38) and (39) straightforwardly if to refer to the unbiasedness condition (35) and neglect $\bar{\mathbf{Z}}_w$, $\tilde{\mathbf{Z}}_w$, and $\tilde{\mathbf{Z}}_v$.

Kalman-like UFIR filtering algorithm

Noticing that the UFIR filter (41) ignoring the noise covariances and initial error is highly attractive for engineering applications, one also notes that the computational problem may arise in its batch form when $N \gg 1$. To circumvent this problem, a fast iterative Kalman-like form has been addressed in [13] for filters, predictors, and smoothers. If to introduce a time shift p to $\bar{\mathbf{x}}_{n+p|n}$ stated by Theorem 2 in [13] for time-invariant models and take into consideration that initial $\mathbf{F}_l$ is time-invariant, then the iterative Kalman-like form of (41) appears as follows:

$$\Upsilon = (\mathbf{C}_{s-m}^T \mathbf{C}_{s-m})^{-1}, \qquad (42)$$

$$\mathbf{F}_s = \mathbf{A}^{s-m+p} \Upsilon \mathbf{A}^{s-m+p\,T}, \qquad (43)$$

$$\bar{\mathbf{x}}_{s|s-p} = \mathbf{A}^{s-m+p} \Upsilon \mathbf{C}_{s-m}^T \mathbf{Y}_{s-p,m-p}, \qquad (44)$$

$$\mathbf{F}_l = [\mathbf{C}^T \mathbf{C} + (\mathbf{A} \mathbf{F}_{l-1} \mathbf{A}^T)^{-1}]^{-1}, \qquad (45)$$

$$\bar{\mathbf{x}}_{l|l-p} = \mathbf{A} \bar{\mathbf{x}}_{l-1|l-p-1} + \mathbf{K}_l(\mathbf{y}_{l-p}$$
$$- \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1|l-p-1}), \qquad (46)$$

where $\mathbf{K}_l \triangleq \mathbf{K}_l(p) = \mathbf{A}^p \mathbf{F}_l \mathbf{C}^T$ is the bias correction gain, $m = n - N + 1$, $s = \alpha - 1$, and an iterative variable l ranges from $\alpha = \max(m + K, m + 2, m + 2 - p)$ to n. The true estimate corresponds to $l = n$.

As well as in the case of UFIR estimator [13], the gain $\mathbf{K}_l$ in (46) also does not depend on noise and initial errors. In this algorithm, we have two batch forms, (43) and (44), which can be computed fast for typically small K. To avoid singularities, the computation starts with $l = \alpha$ and finishes at $l = n$. This last value is used as true and the procedure repeated iteratively for each new measurement. The iterative p-shift Kalman-like algorithm (42)–(46) is listed in Table 1 in a convenient computation form.

Estimation error
Although the estimation error is not involved to the algorithm (42)–(46) that is its extremely remarkable property, the MSE may be required to characterize the filter performance.

The MSE in the p-shift FIR filtering estimate can be evaluated by the matrix

$$\mathbf{P}_l = E\{(\mathbf{x}_l - \bar{\mathbf{x}}_{l|l-p})(\mathbf{x}_l - \bar{\mathbf{x}}_{l|l-p})^T\}. \qquad (47)$$

Substituting $\bar{\mathbf{x}}_{l|l-p}$ with (46), assigning $\bar{\mathbf{x}}_l \triangleq \bar{\mathbf{x}}_{l|l-p}$ and $\varepsilon_l = \mathbf{x}_l - \bar{\mathbf{x}}_l$, and employing (1) and (2), we first write

Table 1 Iterative p-shift Kalman-like UFIR filtering algorithm

Stage	
Given:	$K, N, p, l = \alpha, \ldots, n, s = \alpha - 1$
	$\alpha = \max(m + K, m + 2, m + 2 - p),$
Set:	$\Upsilon = (\mathbf{C}_{s-m}^T \mathbf{C}_{s-m})^{-1}$
	$\mathbf{F}_s = \mathbf{A}^{s-m+p} \Upsilon \mathbf{A}^{s-m+p\,T}$
	$\bar{\mathbf{x}}_{s|s-p} = \mathbf{A}^{s-m+p} \Upsilon \mathbf{C}_{s-m}^T \mathbf{Y}_{s-p,m-p}$
Update:	$\mathbf{F}_l = [\mathbf{C}^T \mathbf{C} + (\mathbf{A} \mathbf{F}_{l-1} \mathbf{A}^T)^{-1}]^{-1}$
	$\bar{\mathbf{x}}_{l|l-p} = \mathbf{A} \bar{\mathbf{x}}_{l-1|l-p-1}$
	$+ \mathbf{A}^p \mathbf{F}_l \mathbf{C}^T(\mathbf{y}_{l-p} - \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1|l-p-1})$
Remark:	Use $\bar{\mathbf{x}}_{l|l-p}$ as the output when $l = n$

$$\mathbf{P}_l = E\{(\mathbf{x}_n - \mathbf{A}\bar{\mathbf{x}}_{l-1} - \mathbf{K}_l \mathbf{y}_{l-p} + \mathbf{K}_l \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1})$$
$$\times (\mathbf{x}_n - \mathbf{A}\bar{\mathbf{x}}_{l-1} - \mathbf{K}_l \mathbf{y}_{l-p} + \mathbf{K}_l \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1})^T\}$$
$$= E\{(\mathbf{A}\varepsilon_{l-1} + \mathbf{B}\mathbf{w}_l - \mathbf{K}_l \mathbf{C}\mathbf{x}_{l-p} - \mathbf{K}_l \mathbf{D}\mathbf{v}_{l-p}$$
$$+ \mathbf{K}_l \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1})(\mathbf{A}\varepsilon_{l-1} + \mathbf{B}\mathbf{w}_l - \mathbf{K}_l \mathbf{C}\mathbf{x}_{l-p}$$
$$- \mathbf{K}_l \mathbf{D}\mathbf{v}_{l-p} + \mathbf{K}_l \mathbf{C} \mathbf{A}^{1-p} \bar{\mathbf{x}}_{l-1})^T\}. \qquad (48)$$

As a next step, it needs to express $\mathbf{x}_{l-p}$ via $\mathbf{x}_{l-1}$. That can be done if we write (1) forward and backward in time for different p and provide the transformations in order to have finally

$$\mathbf{x}_{l-p} = \mathbf{A}^{1-p} \mathbf{x}_{l-1} + \beta_l, \qquad (49)$$

where

$$\beta_l = \begin{cases} \sum_{i=0}^{|p|} \mathbf{A}^{|p|-i} \mathbf{B}\mathbf{w}_{l+i}, & p \leqslant 0 \\ \mathbf{0}, & p = 1 \\ -\sum_{i=1}^{p-1} \mathbf{A}^{i-p} \mathbf{B}\mathbf{w}_{l-i}, & p > 1 \end{cases}. \qquad (50)$$

By substituting (49) with (50) to (48), taking into consideration that $E\{\varepsilon_{l-1}\beta_l^T\}$ and $E\{\beta_l \varepsilon_{l-1}^T\}$ have zero components, and providing the averaging, we finally come up with

$$\mathbf{P}_l = E\{[\,(\mathbf{I} - \mathbf{K}_l \mathbf{C} \mathbf{A}^{-p})\mathbf{A}\varepsilon_{l-1} + \mathbf{B}\mathbf{w}_l - \mathbf{K}_l \mathbf{C}\beta_l$$
$$- \mathbf{K}_l \mathbf{D}\mathbf{v}_{l-p}\,][\,(\mathbf{I} - \mathbf{K}_l \mathbf{C} \mathbf{A}^{-p})\mathbf{A}\varepsilon_{l-1} + \mathbf{B}\mathbf{w}_l$$
$$- \mathbf{K}_l \mathbf{C}\beta_l - \mathbf{K}_l \mathbf{D}\mathbf{v}_{l-p}]^T\}$$
$$= (\mathbf{I} - \mathbf{K}_l \mathbf{C} \mathbf{A}^{-p})\mathbf{A} \mathbf{P}_{l-1} \mathbf{A}^T (\mathbf{I} - \mathbf{K}_l \mathbf{C} \mathbf{A}^{-p})^T$$
$$+ \mathbf{B}\mathbf{R}\mathbf{B}^T - \mathbf{B}\bar{\mathbf{R}}\mathbf{C}^T \mathbf{K}_l^T - \mathbf{K}_l \mathbf{C}\hat{\mathbf{R}}\mathbf{B}^T$$
$$+ \mathbf{K}_l \mathbf{C}\tilde{\mathbf{R}}\mathbf{C}^T \mathbf{K}_l^T + \mathbf{K}_l \mathbf{D}\tilde{\mathbf{Q}}\mathbf{D}^T \mathbf{K}_l^T, \qquad (51)$$

where $\bar{\mathbf{R}} = E\{\mathbf{w}_l \beta_l^T\}$, $\hat{\mathbf{R}} = E\{\beta_l \mathbf{w}_l^T\}$, and $\tilde{\mathbf{R}} = E\{\beta_l \beta_l^T\}$ are given with, respectively,

$$\bar{\mathbf{R}} = \begin{cases} \mathbf{R}\mathbf{B}^T \mathbf{A}^{|p|\,T}, & p \leqslant 0 \\ \mathbf{0}, & p > 0 \end{cases}, \qquad (52)$$

$$\hat{\mathbf{R}} = \begin{cases} \mathbf{A}^{|p|} \mathbf{B}\mathbf{R}, & p \leqslant 0 \\ \mathbf{0}, & p > 0 \end{cases}, \qquad (53)$$

$$\tilde{\mathbf{R}} = \begin{cases} \sum_{i=0}^{|p|} \mathbf{A}^{|p|-i} \mathbf{B}\mathbf{R}\mathbf{B}^T \mathbf{A}^{|p|-i\,T}, & p \leqslant 0 \\ \mathbf{0}, & p = 1 \\ \sum_{i=1}^{p-1} \mathbf{A}^{i-p} \mathbf{B}\mathbf{R}\mathbf{B}^T \mathbf{A}^{i-p\,T}, & p > 1 \end{cases}. \qquad (54)$$

By (51), the estimation error can thus be computed iteratively along with the estimate (46). As can be seen, $\mathbf{P}_n$ inherently diminishes in smoothing filtering ($p < 0$), by $\bar{\mathbf{R}}$ and $\hat{\mathbf{R}}$. It rises with higher rate in predictive filtering

($p > 0$) owing to the effect of $\tilde{\mathbf{P}}$. In the case of filtering ($p = 0$), $\mathbf{P}_n$ is computed by

$$\mathbf{P}_l = (\mathbf{I} - \mathbf{K}_l\mathbf{C})(\mathbf{A}\mathbf{P}_{l-1}\mathbf{A}^T + \mathbf{B}\mathbf{R}\mathbf{B}^T)(\mathbf{I} - \mathbf{K}_l\mathbf{C})^T$$

$$+ \mathbf{K}_l\mathbf{D}\tilde{\mathbf{Q}}\mathbf{D}^T\mathbf{K}_l^T . \tag{55}$$

In all of the cases, $\mathbf{P}_n$ becomes zero if the model is deterministic and the filter order is exactly that of a system.

Note that $\mathbf{P}_n$ computed in such a way ranges upper the true value due to the accumulating effect caused by the noise covariances. Alternatively, $\mathbf{P}_n$ can be well bounded with the error bound (EB) specified in [16] in the three-sigma sense via the noise power gain (NPG) as

$$\beta_{k(vg)}(N,p) = 3\sigma_k\sqrt{g_{k(vg)}(N,p)} \tag{56}$$

to characterize the noise standard deviation in the v-to$-g$ filter channel via measurement of the kth state in the presence of white noise having the variance σ_k^2. The NPG coefficient $g_{k(vg)} \triangleq g_{k(vg)}(N,p)$ is defined here as a components of the NPG matrix $\mathcal{K}_k \triangleq \mathcal{K}_k(N,p)$

$$\mathcal{K}_k = \mathcal{H}_k\mathcal{H}_k^T , \tag{57}$$

$$= \begin{bmatrix} g_{k(11)} & \cdots & g_{k(1k)} & \cdots & g_{k(1K)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g_{k(k1)} & \cdots & g_{k(kk)} & \cdots & g_{k(kK)} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g_{k(K1)} & \cdots & g_{k(Kk)} & \cdots & g_{k(KK)} \end{bmatrix}$$

$$= \mathbf{A}^{N-1+p}(\tilde{\mathbf{C}}_{N-1}^T\tilde{\mathbf{C}}_{N-1})^{-1}\mathbf{A}^{N-1+p^T} . \tag{58}$$

where the thinned gain $\mathcal{H}_k \triangleq \bar{\mathbf{H}}(p)_k \in \mathcal{R}^{K\times N}$ is composed with the Kth columns of $\bar{\mathbf{H}}(p)$ given by (42) starting with the kth one as

$$\mathcal{H}_k = \mathbf{A}^{N-1+p}(\tilde{\mathbf{C}}_{N-1}^T\tilde{\mathbf{C}}_{N-1})^{-1}\tilde{\mathbf{C}}_{N-1}^T \tag{59}$$

and $\tilde{\mathbf{C}}_i \triangleq (\mathbf{C}_i)_k$ is the kth row of $\mathbf{C}_i$.

To avoid the computational problem with $N \gg 1$, the NPG $\mathcal{K}_k$ can be computed iteratively [12] as

$$\mathcal{K}_{kj} = \mathbf{A}[\mathbf{A}^{1-p^T}\mathbf{C}^T\mathbf{C}\mathbf{A}^{1-p} + \mathcal{K}_{k(j-1)}^{-1}]^{-1}\mathbf{A}^T , \tag{60}$$

by changing an iterative variable j from $j = \gamma \geqslant K$, to $N-1$. The initial value $\mathcal{K}_{k(\gamma-1)}$ is provided by (58), if to substitute N with γ, and the true $\mathcal{K}_k$ is taken when $l = N-1$.

Applications

A comparison of errors in the FIR and Kalman filters has been provided in many articles [2,9,10,12,13]. Much lesser attention has been paid to the trade-off between the OFIR and UFIR filter outputs. To investigate errors in the proposed p-shift OFIR and UFIR filters and thereby learn

their facilities, below we exploit a two-state model represented with (1) and (2) having $\mathbf{A} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}$, $\mathbf{C} =[\ 1\ 0\ ]$, $\mathbf{B}$, and $\mathbf{D}$ identity, $x_0 = 1$, $y_0 = 0.01$ s^{-1}, $\tau = 1$, $\sigma_x = 0.1$, and $\sigma_y = 10^{-3}/s$. The covariances (7) and (8) of zero mean noise components, $\mathbf{w}_n$ and v_n, are allowed to be $\mathbf{R} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$ and $\mathbf{Q} =[\ \sigma_v^2\ ]$, respectively.

Measurement has been provided in the presence of the zero-mean noise v_n uniformly distributed from -2 to 2 with the variance $\sigma_v = 2/\sqrt{3}$.

Both the OFIR algorithm and UFIR one (Table 1) have been tested and the filtering errors evaluated in the first state at a current point n for different p and fixed N. Errors were bounded with EB $\beta(p) \triangleq \beta_{1(11)}(N,p)$ calculated by (56).

Errors in predictive FIR filtering

Supposing that the estimate is required at $n = 50$ and assuming that measurement may not be available in the nearest past points (as it sometimes occurs in wireless systems), we let $0 \leqslant p \leqslant 30$ and find the predictive filtering estimate for $N = 10$ and $N = 20$. Figure 2 sketches errors in OFIR and UFIR estimates as functions of p.

Here, we also show the bounds $\pm\beta(p)$ for each N as functions of p. Note that for the model in question, EB can also be calculated via NPG $g_1(N,p)$ found in [5] as
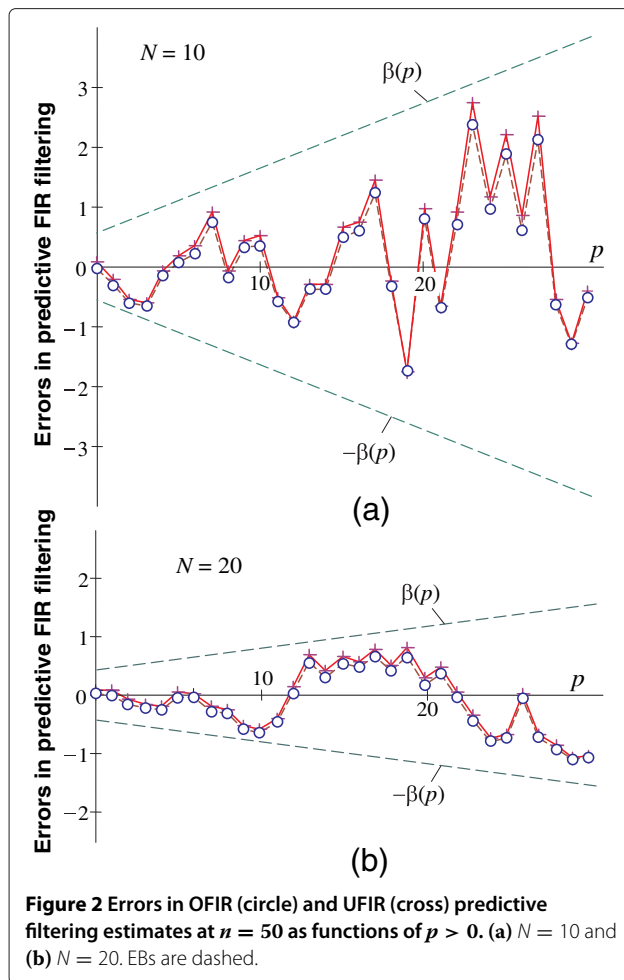
$$\beta_{1(11)}(N,p) = 3\sigma_v\sqrt{g_1(N,p)} \tag{61}$$

$$= 3\sigma_v \left\{ \frac{1}{N(N^2-1)} [2(2N-1)(N-1) \right.$$

$$\left. + 12p(N-1+p)] \right\}^{0.5} . \tag{62}$$

Observing Figure 2, one infers that the estimates are closely related and that the errors range well within EBs stretched by growing p. Inherently, the prediction error is reduced by increasing N that can be seen by comparing Figure 2a,b.

Errors in smoothing FIR filtering

In the second experiment, we change p from zero to $-N+1$ and evaluate errors in the smoothing filters at $n = 30$. Figure 3 illustrates the results for $N = 10$ (Figure 3a) and $N = 20$ (Figure 3b).

The first conclusion that can be made is that errors reach a minimum at a center of the averaging horizon with $p = -N/2$. This should not be surprising, since the ramp impulse response associated with linear models reduces at this point to the uniform one associated with simple averaging [11] producing noise minimum among all other filters [6]. It can also be seen that the difference between the optimal and unbiased estimates exists but it is not

Figure 2 Errors in OFIR (circle) and UFIR (cross) predictive filtering estimates at $n = 50$ as functions of $p > 0$. (a) $N = 10$ and **(b)** $N = 20$. EBs are dashed.

Figure 3 Errors in OFIR (circle) and UFIR (cross) smoothing filtering estimates at $n = 50$ as functions of $p < 0$. (a) $N = 10$ and **(b)** $N = 20$. EBs are dashed.

large, in view of the scale in Figures 2 and 3. And we notice again that errors in the smoothing filter range well within a gap between the EBs.

Conclusion

In this article, the p-shift OFIR and UFIR algorithms with the properties of predictive filtering ($p > 0$), filtering ($p = 0$), and smoothing filtering ($p < 0$) have been addressed for linear discrete time-invariant state-space models. The OFIR filter is shown to self-determine the mean square initial state function by solving the DARE. The UFIR filter represented both in the batch and iterative Kalman-like forms ignores covariances and initial errors, unlike the Kalman filter. As an example of applications, we have exploited the two-state polynomial model and investigated errors in the OFIR and UFIR filters. Based upon, we have confirmed the statement made earlier for FIR filters, predictors and smoothers: the UFIR estimate converges to the OFIR one by increasing N and the estimation errors are well bounded with EB. That means that exact optimality may be redundant with $N \gg 1$ and still a nice
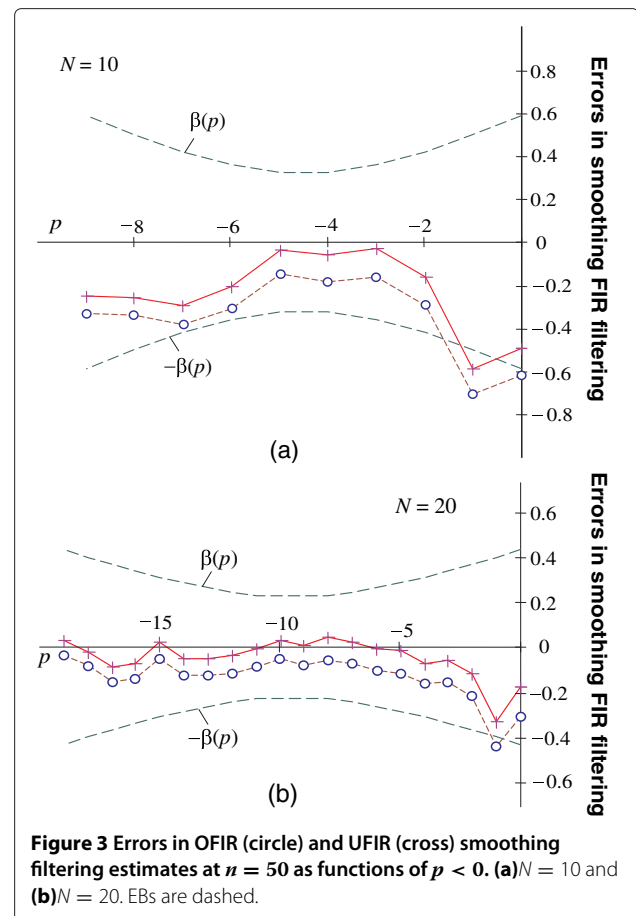
suboptimal estimate can be provided with UFIR filters at much lower cost.

An importance of the OFIR and UFIR filtering algorithms proposed resides in the fact that they are both general for linear discrete time-invariant state-space models. The algorithms virtually generalize the well-known Savitzky-Golay solution for smoothing and predictive filtering in state-space. However, unlike the latter, both OFIR and UFIR filters have the convolution-based forms more familiar for electronics engineers. Moreover, the convolution computation can efficiently be provided in the frequency domain that is its another benefit. Finally, engineers should certainly appreciate the iterative Kalman-like algorithm. Paying attention to these advantages, our current investigations are focused on several applied problems associated with signal and image processing.

Endnotes

[a] $\tilde{x}_{n|k}$ is the filtering estimate at n via measurement from the past to k; $\hat{x}_{n|k}$ is optimal and $\bar{x}_{n|k}$ unbiased.
[b] The case of filtering out all of the noise components is considered in [12].

Appendix

Mean square initial state function

Consider the estimate provided by (20) with gain (32),

$$\hat{\mathbf{x}}_{n|n-p} = [\mathbf{A}^{N-1+p}(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T\mathbf{Z}_{m-p} + \bar{\mathbf{Z}}_w]$$
$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Y}_{n-p,m-p}. \quad (A.1)$$

Following [12], find the smoothing estimate at the initial point $n - N + 1$ of the averaging interval. By letting $p = -(n - m) = -(N - 1)$, go to

$$\hat{\mathbf{x}}_{n|n+N-1} = [(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T\mathbf{Z}_n + \bar{\mathbf{Z}}_w]$$
$$\times (\mathbf{Z}_n + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Y}_{n+N-1,n}. \quad (A.2)$$

Substitute n with $m - p$ and find the estimate at $m - p$,

$$\hat{\mathbf{x}}_{m-p|m+N-1-p} = [(\mathbf{C}_{N-1}^T\mathbf{C}_{N-1})^{-1}\mathbf{C}_{N-1}^T\mathbf{Z}_{m-p} + \bar{\mathbf{Z}}_w]$$
$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Y}_{m+N-1-p,m-p}$$
$$= (\mathbf{R}_{m-p}\mathbf{C}_{N-1}^T + \bar{\mathbf{Z}}_w)(\mathbf{Z}_{m-p}+\tilde{\mathbf{Z}}_w+\tilde{\mathbf{Z}}_v)^{-1}$$
$$\times \mathbf{Y}_{m+N-1-p,m-p}. \quad (A.3)$$

Now, substitute the unknown $\mathbf{x}_{m-p}$ with its optimal estimate $\hat{\mathbf{x}}_{m-p|m+N-1-p}$ and recall that $\mathbf{x}_{m-p}$ is supposed to be known exactly. This allows providing the following transformations:

$$\mathbf{R}_{m-n} = E\{\mathbf{x}_{m-p}\mathbf{x}_{m-p}^T\} = \mathbf{x}_{m-p}\mathbf{x}_{m-p}^T \quad (A.4)$$

$$\cong E\{\hat{\mathbf{x}}_{m-p|m+N-1-p}\hat{\mathbf{x}}_{m-p|m+N-1-p}^T\} \quad (A.5)$$

$$\cong \hat{\mathbf{x}}_{m-p|m+N-1-p}\hat{\mathbf{x}}_{m-p|m+N-1-p}^T. \quad (A.6)$$

By employing (A.6) and taking into account that $\mathbf{R}_{m-p}$, $\bar{\mathbf{Z}}_w$, $\mathbf{Z}_{m-p}$, $\tilde{\mathbf{Z}}_w$, and $\tilde{\mathbf{Z}}_v$ are all symmetric, transform (A.6) to

$$\mathbf{R}_{m-p} \cong \tilde{\mathbf{x}}_{m-p|m+N-1-p}\tilde{\mathbf{x}}_{m-p|m+N-1-p}^T \quad (A.7)$$
$$= (\mathbf{R}_{m-p}\mathbf{C}_{N-1}^T + \bar{\mathbf{Z}}_w)(\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}$$
$$\times \mathbf{Y}_{m+N-1-p,m-p}\mathbf{Y}_{m+N-1-p,m-p}^T$$
$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}(\mathbf{C}_{N-1}\mathbf{R}_{m-p} + \bar{\mathbf{Z}}_w). \quad (A.8)$$

A supposition that $\mathbf{x}_{m-p}$ is deterministic makes both $\mathbf{R}_{m-p}$ and $\mathbf{Z}_{m-p}$ singular. However, if we allow an equality in (A.7) and solve (A.8) for $\mathbf{Z}_{m-p}$, these matrices will be found approximately in the minimum MSE sense and thus no longer be singular. Next, observe that the second term in the first parenthesis on (A.8) represents the noise variance on the averaging interval and is commonly much smaller than the first term representing the initial state gained. Then neglect $\bar{\mathbf{Z}}_w$, accept an equality in (A.7), multiply (A.8) with $\mathbf{C}_{n-m}$ and $\mathbf{C}_{n-m}^T$ from the left-hand and

right-hand sides, respectively, invoke (26), remove nonsingular $\mathbf{Z}_{m-p}$ from both sides, and substitute $m+N-1$ with n. That leads to

$$\mathbf{I} = (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Y}_{n-p,m-p}\mathbf{Y}_{n-p,m-p}^T$$
$$\times (\mathbf{Z}_{m-p} + \tilde{\mathbf{Z}}_w + \tilde{\mathbf{Z}}_v)^{-1}\mathbf{Z}_{m-p}. \quad (A.9)$$

By rearranging the terms, (A.9) becomes the DARE (36), whose solution with respect to $\mathbf{Z}_{m-p}$ can be found following [15].

References
1. EF Camacho, C Bordons, *Model Predictive Control* (Springer-Verlag, Berlin, 2004)
2. WH Kwon, S Han, *Receding Horizon Control: Model Predictive Control for State Models* (Springer, Berlin, 2005)
3. P Heinonen, Y Neuvo, FIR-median hybrid filters with predictive FIR structures. IEEE Trans. Acoust. Speech Signal Process. **36**(6), 892–899 (1988)
4. TG Campbell, Y Neuvo, Predictive FIR filters with low computational complexity. IEEE Trans. Circ. Syst. **38**(9), 1067–1071 (1991)
5. YS Shmaliy, An unbiased *p*-step predictive FIR filter for a class of noise-free discrete-time models with independently observed states. Signal Image Video Process. **3**(2), 127–135 (2009)
6. A Savitzky, MJE Golay, Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. **36**(8), 1627–1639 (1964)
7. X Zhou, X Wang, FIR-median hybrid filters with polynomial fitting. Digital Signal Process. **39**(2), 112–124 (2004)
8. J-T Yuan, JA Stuller, Order-recursive FIR smoothers. IEEE Trans. Signal Process. **42**(5), 1242–1246 (1994)
9. BK Kwon, S Han, OK Kwon, WH Kwon, Minimum variance FIR smoothers for discrete-time state space models. IEEE Signal Process. Lett. **14**(8), 557–560 (2007)
10. S Han, WH Kwon, $L_2–E$ FIR smoothers for deterministic discrete-time state-space signal models. IEEE Trans. Autom. Control. **52**(5), 927–932 (2007)
11. YS Shmaliy, L Morales-Mendoza, FIR smoothing of discrete-time polynomial signals in state space. IEEE Trans. Signal Process. **58**(5), 2544–2555 (2010)
12. YS Shmaliy, Linear optimal FIR estimation of discrete time-invariant state-space models. IEEE Trans. Signal Process. **58**(6), 3086–3096 (2010)
13. YS Shmaliy, An iterative Kalman-like algorithm ignoring noise and initial conditions. IEEE Trans. Signal Process. **59**(6), 2465–2473 (2011)
14. H Stark, JW Woods, *Probability, Random Processes, and Estimation Theory for Engineers,* 2nd edn. (Prentice Hall, Upper Saddle River, 1994)
15. P Lancaster, L Rodman, *Algebraic Riccati Equations* (Oxford University Press, New York, 1995)
16. YS Shmaliy, O Ibarra-Manzano, Noise power gain for discrete-time FIR estimators. IEEE Signal Process. Lett. **18**(4), 207–210 (2011)