Studies of Discrete-Time Unbiased FIR Filters of Polynomial State-Space Models

Yuriy S. Shmaliy 1, Oscar Ibarra-Manzano 2
Electronics Department, Guanajuato University, Salamanca, 36885, Mexico
1shmaliy@salamanca.ugto.mx
2ibarrao@salamanca.ugto.mx

Abstract—We study an unbiased finite impulse response (FIR) filter in applications to discrete-time state space models with polynomial representation of the states. The unique l-degree polynomial FIR filter gain and the estimate variance are found for a general case. The noise power gain (NG) is derived for white Gaussian noises in the model and in the measurement. The filter does not involve any knowledge about noise in the algorithm. It is unstable at short horizons, 2 ≤ N ≤ l, and inefficient (NG exceeds unity) in the narrow range l < N ≤ N_o, where N_o is ascertained by the cross-components in the measurement matrix C. With N ≫ N_o, the filter NG poorly depends on C and fits the asymptotic function (l+1)^2/N. With very large N ≫ 1, the estimate noise becomes negligible and the filter thus optimal in the sense of zero bias and zero noise. An example is given for a two-state system.

I. INTRODUCTION

There is a wide class of systems, which states change with time slowly1 and relevant signals can therefore be represented with degree polynomials. Examples can be found in time synchronization of digital communication networks, target tracking provided by radars, positioning obtained with the Global Positioning System (GPS), etc. Inherently slowly changing are narrowband channels in wireless communications, remote wireless control, remote sensing, etc. For such systems, finite impulse response (FIR) filtering, providing weighted averaging on a horizon of N points, typically obtains a nice signal restoration owing to some important properties. In contrast to the infinite impulse response (IIR) structures, including the Kalman-Bucy filter [1], [2], FIR structures are inherently bounded input/bounded output stable and more robust against temporary model uncertainties and round-off errors [3]. Moreover, they allow noises to have arbitrary distributions and covariances.

For polynomial models, the optimal FIR filter gain can be found using the Johnson discrete-time analog [4] of Zadeh-Ragazzini’s filter [5]. Following [4], the gain coefficients can be defined via the unbiasedness constraint using the method of Lagrange multipliers. The approach was efficiently developed in [6] by Heinonen and Neuvo, who proposed a one-parameter family of the 1-step toward predictive FIR filters with short polynomial gain functions, thus having strong engineering features. Polynomial FIR estimators can also be obtained employing regression analysis [7] and in state space [8].

In state space, Jazwinski addressed a finite memory filter for systems without noise employing a maximum likelihood criterion [9]. Soon after, Roberts and Tooley in [10] and Mullis with Roberts in [11] made some generalizations for the approach. The same problem was later solved by Ling and Lim in [12] with the proposed moving horizon least-square filter. In [13], Kwon, Kim, and Park used weighted averaging to derive the Kalman-FIR filter. Most recently, Kwon, Kim, and Han proposed in [14] an unbiased optimal FIR filter for a state space model with white Gaussian noise both in the system and in the measurement and Han, Kwon, and Kim found in [15] a solution to deterministic systems assuming the system matrix nonsingular. Shmaliy in [16] proposed an unbiased FIR filter involving no noise and initial state conditions to the algorithm for real-time signal processing and Kim and Lee found a similar solution in [17] for receding horizon control. The polynomial gain of the unbiased FIR filter was found in [16] for deterministic systems with independently observed states. The gain was derived fitting the unbiasedness condition

\[ E[\hat{x}_n|n] = E[x_n], \]

where x_n and \( \hat{x}_n|n \) are the signal vector and its a posteriori estimate, respectively, at a current discrete time point n. It was then shown in [18] that, by large averaging horizons, N ≫ 1, a relatively simple ramp unbiased FIR filter and the optimal one, existing in the sense of the minimum mean square error (MSE), produce virtually indistinguishable estimates.

Below, the approach [16] is extended to a general discrete-time polynomial state-space model. In more detail, solutions of filtering, smoothing, and prediction problems are elucidated in [19]–[21].

II. POLYNOMIAL STATE SPACE MODEL AND PROBLEM FORMULATION

Consider a discrete time-invariant linear model represented in state space with

\[ x_n = Ax_{n-1} + Bw_n, \]  
\[ y_n = Cx_n + Dw_n, \]
where the $K \times 1$ state and observation vectors are given by, respectively,
\[
X_n = [x_{1n} \ x_{2n} \ \ldots \ x_{Kn}]^T,
\]
\[
Y_n = [y_{1n} \ y_{2n} \ \ldots \ y_{Kn}]^T,
\]

in which $x_{kn}, k \in [1,K]$, is the $k$th state associated with the derivative of the $(k-1)$th state, starting with $k = 2$, and $y_{kn}$ is the measurement of the $k$th state. Suppose that the $k$th state is expanded to the Taylor series such that the $K$-degree series represents the first state and the 1-degree the $K$th state [16]. The $K \times K$ transition matrix, projecting the nearest past state $x_{n-1}$ to the present state $x_n$, is thus triangular,
\[
A = \begin{bmatrix}
1 & \tau & \tau^2 & \ldots & \tau^{K-1} \\
0 & 1 & \tau & \ldots & \tau^{K-2} \\
0 & 0 & 1 & \ldots & \tau^{K-3} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1
\end{bmatrix},
\] (6)

where $\tau = t_n - t_{n-1}$ is the sampling time and $t_n$ is a current discrete time at $n$. By (6), the state model (2) becomes polynomial and the $K \times K$ measurement matrix $C$ triangular,
\[
C = \begin{bmatrix}
c_{11} & c_{12} & \ldots & c_{1K} \\
0 & c_{22} & \ldots & c_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & c_{KK}
\end{bmatrix}.
\] (7)

Both matrices, $B$ and $D$, have dimensions dependent on applications. The input and measurement noise vectors, both of $K \times 1$ dimensions, respectively,
\[
w_n = [w_{1n} \ w_{2n} \ \ldots \ w_{Kn}]^T,
\]
\[
v_n = [v_{1n} \ v_{2n} \ \ldots \ v_{Kn}]^T,
\]

have zero-mean components, $E\{w_n\} = 0$ and $E\{v_n\} = 0$. It is implied that $w_n$ and $v_n$ are mutually uncorrelated and independent, $E\{w_nv_n^T\} = 0$, having arbitrary covariances, respectively,
\[
R = E\{w_nw_n^T\},
\] (10)
\[
Q = E\{v_nv_n^T\}
\] (11)

for all $i$ and $j$. The latter means that $w_n$ and $v_n$ are not obligatorily Gaussian and delta-correlated.

III. AN UNBIASED FIR FILTER

An unbiased FIR filter can be derived on a horizon of $N$ points by the discrete-time convolution (1), if to represent (2) and (3) with recursively computed forward-in-time solutions as follows [22]:

\[
X_N = A_NX_{n-N+1} + B_NN_N,
\]

\[
Y_N = C_NX_{n-N+1} + G_NN_N + D_NV_N,
\]

where
\[
X_N = [x_{1n}^T \ x_{2n}^T \ \ldots \ x_{n-N+1}^T]^T,
\] (14)

\[
Y_N = [y_{1n}^T \ y_{n-N+1}^T \ \ldots \ y_{n-N-N+1}^T]^T,
\]

\[
N_N = [w_{1n}^T \ w_{n-N+1}^T \ \ldots \ w_{n-N-N+1}^T]^T,
\]

\[
V_N = [v_{1n}^T \ v_{n-N+1}^T \ \ldots \ v_{n-N-N+1}^T]^T,
\]

\[
A_N = [(A^{N-1})^T \ (A^{N-2})^T \ \ldots \ A^T \ I]^T,
\]

\[
B_N = \begin{bmatrix}
B & AB & \ldots & A^{N-2}B & A^{N-1}B \\
0 & B & \ldots & A^{N-3}B & A^{N-2}B \\
0 & 0 & \ldots & B & AB \\
0 & 0 & \ldots & 0 & B
\end{bmatrix},
\] (20)

\[
C_N = \begin{bmatrix}
CA^{N-1} \\
CA^{N-2} \\
\vdots \\
CA \\
C
\end{bmatrix},
\] (21)

\[
D_N = \text{diag}(D \ D \ \ldots \ D\ N),
\] (22)

\[
G_N = \begin{bmatrix}
CB & CAB & \ldots & CA^{N-2}B & CA^{N-1}B \\
0 & CB & \ldots & CA^{N-3}B & CA^{N-2}B \\
0 & 0 & \ldots & CB & CAB \\
0 & 0 & \ldots & 0 & CB
\end{bmatrix},
\] (23)

Note that the initial state $x_{n-N+1}$ is supposed to be known exactly, although it is randomly valued. Therefore, $w_{n-N+1}$ in (16) is always zero valued.

It has been shown in [14] that optimal FIR filtering can be applied to each of the states separately. Utilizing $N$ measurements from $n-N+1$ to $n$, the FIR filtering estimate $\hat{x}_{n|n} = [\hat{x}_{1n|n} \ \hat{x}_{2n|n} \ \ldots \ \hat{x}_{Kn|n}]^T$ of $x_n$ is hence obtained at $n$ as follows:

\[
\hat{x}_{n|n} = \sum_{i=0}^{N-1} H_iy_{n-i},
\] (24a)

\[
= W^TY_N
\] (24b)

\[
= W^T[C_Nx_{n-N+1} + G_NN_N + D_NV_N],
\] (24c)

where the filter gain is represented with the matrices $H_i = \text{diag}(h_{k(K-1)}, h_{k(K-2)}, \ldots, h_{k})$, $W = [H_0^T \ H_1^T \ \ldots \ H_{N-1}^T]^T$, (25)
By (1), the estimate \( x_{n|n} \) becomes unbiased. Applying (1) to (24c), we thus arrive at the strong unbiasedness (or deadbeat) constraint

\[
A^{N-1} = W_T C_N ,
\]

where \( A^{N-1} \) is specified by (19). As can be seen, the constraint (27) has an inherent property: it does not depend on the noise and the initial state.

A. Estimate for the \( k \)th State

Regarding the \( k \)th state, \( k \in [1, K] \), the estimate (24c) can be rewritten as

\[
\hat{x}_{kn|n} = W^T_i C_{Nk} x_{n-N+1} + W^T_i (G_{Nk} N_N + D_{Nk} V_N) ,
\]

where the \( l \)-degree filter gain matrix is [16]

\[
W_l = [h_{l0} h_{l1} \ldots h_{l(N-1)}]^T ,
\]

in which, for the sake of unbiasedness, one must set\(^2 l = K - k \). The other matrices are defined as

\[
C_{Nk} = \begin{bmatrix}
(CA^{N-1})_k \\
(CA^{N-2})_k \\
\vdots \\
(CA)_k \\
C_k
\end{bmatrix} ,
\]

\[
G_{Nk} = \begin{bmatrix}
(CB)_k \\
(CAB)_k \\
\vdots \\
0 \\
0 \ldots (CB)_k
\end{bmatrix} ,
\]

\[
D_{Nk} = \text{diag} \left( D_k \ D_k \ldots D_k \right) ,
\]

in which \((Z)_k \) denotes the \( k \)th row of a matrix \( Z \).

For the \( k \)th row, (1) means \( E\{\hat{x}_{kn|n}\} = E\{x_{kn}\} \) and the constraint (27) accordingly becomes

\[
(A^{N-1})_k = W^T_i C_{Nk} .
\]

A matrix equation (33) can now be solved for the coefficients of a degree polynomial \( h_{kn} \) [16] if we allow \( l = K - k \).

B. Unique Polynomial Gain

Following [16], the filter gain can be represented with the degree polynomial

\[
h_{li} = \frac{1}{c_{kk}} \sum_{j=0}^{l} a_{lj} i^j , \quad l \in [1, K] ,
\]

where the coefficient \( a_{lj} \) is given as

\[
a_{lj} = (-1)^j \frac{M_{(j+1)1}}{|A|} ,
\]

(35)

\(^2\)Unbiasedness is also achieved when \( l > K - k \), although with larger noise [25].

Determined \( h_{li} \), the unbiased FIR filtering estimate of the model \( k \)th state, \( k \in [1, K] \), is provided by

\[
\hat{x}_{kn|n} = \sum_{i=0}^{N-1} h_{(K-k)i} y_{k(n-i)} ,
\]

where \( Y_N = [y_{n-k} y_{n-k+1} \ldots y_{n-N+1}]^T \).

Most generally, combining all of the estimates in a vector, the unbiased FIR estimate of the model state is ascertained by

\[
\hat{x}_{kn|n} = \sum_{i=0}^{N-1} H_i y_{n-i} ,
\]

(38a)

\[
= W^T_i Y_{N_k} .
\]

(38b)

C. Properties of the Unbiased FIR Filter Gain

It follows that the polynomial gain (35) existing from 0 to \( N - 1 \) satisfies the property

\[
\sum_{i=0}^{N-1} h_{li} i^u = 0 , \quad u \in [1, l] ,
\]

(40)

\[
\sum_{i=0}^{N-1} h_{li}^2 = \frac{a_{kl}}{c_{kk}^2} .
\]

(41)

For identity \( C \), the filter gain satisfies (39)–(41) if to set \( c_{kk} = 1 \).

IV. ESTIMATE VARIANCE

As can be seen, \( h_{li} \) does not involve noise to the coefficients. Thus, \( w_n \) and \( v_n \) are allowed to have arbitrary distributions and covariances. The latter affect the estimate variance.

The MSE of the unbiased FIR estimate of the \( k \)th state can be written as

\[
J_k = E\{(x_{kn} - \hat{x}_{kn|n})^2\}
\]

\[
= E\{[(A^{N-1})_k x_{n-N+1} - W^T_i C_{Nk} x_{n-N+1} - W^T_i G_{Nk} N_N - W^T_i D_{Nk} V_N]^2\} .
\]

(42)

Embedded (33) and accounted for the commutativity of \( W^T_i G_{Nk} N_N = (G_{Nk} N_N)^T W_i \) and \( W^T_i D_{Nk} V_N = (D_{Nk} V_N)^T W_i \), the MSE (42) represents the variance

\[
\sigma_k^2 = E\{(W^T_i G_{Nk} N_N + W^T_i D_{Nk} V_N)^2\}
\]

\[
= W^T_i (G_{Nk} N N_{Nk}^T C_{Nk} + D_{Nk} V V_N^T D_{Nk}^T W_i ) .
\]

(43)

V. LOW-DEGREE GAINS

Typically, slowly changing signals are represented at a horizon of some points with low-degree polynomials.
where the coefficients are

\[ a_i = \frac{2(2N-1) + 6\beta_i}{N(N+1)} \]

where the coefficients are

\[ a_i = \frac{6}{N(N+1)} - \frac{12\beta_i}{N(N^2-1)} \]

in which the coefficient \( \beta_{kk} \) is

\[ \beta_{kk} = (-1)^{v-k} \frac{c_{kk}}{c_{kk}} \frac{(v-k)!}{\tau^{v-k}} , \quad v \geq k \]

and such that \( \beta_{kk} = 1 \). Note that \( \beta_1 \) means \( \beta_{12} \) specified by (47). Here and in the following, we assign \( \beta_s \triangleq \beta_{k(k+s)} \).

The second state, \( k = 2 \), is processed with simple averaging, \( l = K-k = 0 \).

Because C has the only cross-component \( c_{12} \), the unique coefficient exists that is \( \beta_1 = -c_{12} \). Fig. 1 sketches (44) for different \( \beta_1 \) and one realizes that \( c_{12} \) rotates \( h_{11} \) about a center of the averaging horizon.

The noise power gain (NPG) (41) is given for this gain as

\[ g_{v1} = \frac{2(2N-1)}{N(N+1)} + \frac{12(N-1 + \beta_1)}{N(N^2-1)} \]

A. Two-State Model

Consider the two-state space system, \( K = 2 \). The first state, \( k = 1 \), is unbiasedly filtered with a ramp gain, \( l = K-k = 1 \). By (34) and (35), this gain is represented with

\[ h_{11} = a_{01} + a_{11}i \]

where the coefficients are

\[ a_{01} = \frac{2(2N-1) + 6\beta_1}{N(N+1)} \]

and such that \( \beta_{kk} = 1 \). Note that \( \beta_1 \) means \( \beta_{12} \) specified by (47). Here and in the following, we assign \( \beta_s \triangleq \beta_{k(k+s)} \).

The second state, \( k = 2 \), is processed with simple averaging, \( l = K-k = 0 \).

Because C has the only cross-component \( c_{12} \), the unique coefficient exists that is \( \beta_1 = -c_{12} \). Fig. 1 sketches (44) for different \( \beta_1 \) and one realizes that \( c_{12} \) rotates \( h_{11} \) about a center of the averaging horizon.

The noise power gain (NPG) (41) is given for this gain as

\[ g_{v1} = \frac{2(2N-1)}{N(N+1)} + \frac{12(N-1 + \beta_1)}{N(N^2-1)} \]

B. Three-State Model

For systems having three states, \( K = 3 \), the first state, \( k = 1 \), is filtered, letting \( l = K-k = 2 \), with a quadratic gain

\[ h_{21} = a_{02} + a_{12}i + a_{22}i^2 \]

for which the coefficients can be defined, by (42), to be

\[ a_{02} = 3(3N^2-3N+2) \]

and such that \( \beta_{kk} = 1 \). Note that measurements with \( \beta_1 < 0 \), contrary to the case of \( \beta_1 > 0 \), are more preferable, because of lower noise in the estimate and better stability of the filter (Fig. 2).

We thus deduce that rotating \( h_{11} \) (Fig. 1) compensates, if \( c_{12} \neq 0 \), an influence of the second state upon the first one, thereby guarantying unbiasedness. Note that measurements with \( \beta_1 < 0 \), contrary to the case of \( \beta_1 > 0 \), are more preferable, because of lower noise in the estimate and better stability of the filter (Fig. 2).
...and the third one, \( g_1 \), and the fourth \( g_2 \) by simple averaging. We thus have

\[
\bar{x}_{1n|n} = \sum_{i=0}^{N-1} \frac{1}{N} h_{1n}(u_{n-i}) y_{n-i}\]  

VI. EXAMPLE: TWO-STATE SYSTEM

As an example of applications, we chose a polynomial two-state model, represented in state space with

\[
\begin{bmatrix}
    x_{1n} \\
    x_{2n}
\end{bmatrix} = \begin{bmatrix}
    1 & \tau \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    x_{1(n-1)} \\
    x_{2(n-1)}
\end{bmatrix} + \begin{bmatrix}
    w_{1n} \\
    0
\end{bmatrix},
\]

\[
\begin{bmatrix}
    y_{1n} \\
    y_{2n}
\end{bmatrix} = \begin{bmatrix}
    1 & c_{12} \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    x_{1n} \\
    x_{2n}
\end{bmatrix} + \begin{bmatrix}
    v_{1n} \\
    0
\end{bmatrix},
\]

where \( w_{1n} \) is delta-correlated Gaussian noise with the variance \( \sigma_{w1}^2 \). The measurement noise \( v_{1n} \) is uniformly distributed with the variance \( \sigma_{v1}^2 \). The noises are mutually uncorrelated; that is \( E\{w_{1i}v_{1j}\} = 0 \) for all \( i \) and \( j \).

The FIR filtering estimate of the first state \( x_{1n} \) is provided by the ramp gain \( h_{1s} \) and the second state \( x_{2n} \) by simple averaging. We thus have

\[
\bar{x}_{1n|n} = \sum_{i=0}^{N-1} \frac{1}{N} h_{1i}(u_{n-i}) y_{n-i}
\]

It has been shown experimentally in [16] and [22] that, for the uniformly distributed measurement noise and \( C \) identity, the unbiased FIR filter produces lower errors than the standard Kalman filter. Referring to, in this example we observe solely an influence of the cross-component \( c_{12} \) of \( C \) upon the unbiased FIR filtering estimate.

Defining \( \beta_1 = -c_{12} \) and allowing for \( x_{10} = 0 \), \( x_{20} = 0.01/c \), \( \sqrt{\sigma_{w1}^2} = 0.05 \), \( \sqrt{\sigma_{v1}^2} = 2.89/c \), and \( \tau = 1 \), we simulate the noisy functions and ascertain \( \bar{x}_{1n|n} \) for \( N = 800 \). Fig. 4a shows typical behaviors of the functions for \( c_{12} = 300/c \). Here, we also show a deterministic behavior of the first state \( \bar{x}_{1n|n} \), neglecting \( w_{1n} \), and the estimate \( \tilde{x}_{1n|n} \), if to neglect \( c_{12} \), by setting \( \beta_1 = 0 \). As can be seen, the gain \( h_{1s} \) compensates the bias caused by \( c_{12} \), attenuates substantially both noise components, \( w_{1n} \) and \( v_{1n} \), and produces the estimate closely related to the mean value of \( x_{1n} \).

For the instantaneous errors,

\[
\epsilon_n = x_{1n} - \bar{x}_{1n|n},
\]

\[
\epsilon_n(\beta_1 = 0) = x_{1n} - \bar{x}_{1n|n}(\beta_1 = 0),
\]

we evaluate the estimate mean errors (biases), \( E\{\epsilon_n\} \) and \( E\{\epsilon_n(\beta_1 = 0)\} \), and standard deviations, \( \sqrt{\sigma_{\epsilon1}^2} \) and \( \sqrt{\sigma_{\epsilon0}^2} \). The values calculated over \( 5 \times 10^5 \) measurements for \( c_{12} \) changed from \(-300\) to \( 300 \) are shown in Fig. 4b. It is seen that, by \( \beta_1 = 0 \), the bias in the estimate rises dramatically. Exactly as stated by Fig. 2, negative values of \( c_{12} \) increase and positive reduce a random amount of errors in the estimate. There is a near optimum value \( c_{12} \approx 100 \) making \( \sqrt{\sigma_{\epsilon1}^2} \) minimum with a still near zero bias. Finally, the standard deviation in both cases inherently lies above the lower bound \( g_1 \sqrt{\sigma_{\epsilon1}^2} \) provided by the NPG \( g_1 \).
VII. CONCLUDING REMARKS

We discussed an unbiased FIR filter for discrete-time state space models, which states are represented with degree polynomials. The filter can be used for any averaging horizon \( N \geq 2 \), although it is most efficient when \( N \) is large; that is the system state changes with time slowly. The filter does not involve any knowledge about noise in the algorithm. It is unstable only at short horizons, \( 2 \leq N \leq l \), and inefficient in the narrow range \( l < N < N_h \), where \( N_h \) is zero, by C diagonal, and increases by the cross-components in C. With \( N \geq N_h \), the filter NG poorly depends on \( c_{ij}, j \neq i \), and fits the asymptotic function \((l + 1)^2/N\). With very large \( N \gg 1 \), the estimate noise becomes negligible and the filter thus optimal in the sense of both zero bias and zero noise.

REFERENCES