ABSTRACT

In this paper, the optimal finite impulse response (OFIR) filter with embedded unbiasedness (EU) is derived by minimizing the mean square error (MSE) subject to the unbiasedness constraint for discrete time-invariant state-space models. Unlike the OFIR filter, the OFIR-EU filter does not require the initial conditions. In terms of accuracy, the OFIR-EU filter occupies an intermediate place between the UFIR and OFIR filters. With a two-state harmonic model, we show that the OFIR-EU filter has higher immunity against errors in the noise statistics and better robustness against temporary model uncertainties than the OFIR and Kalman filters.

1. INTRODUCTION

The finite impulse response (FIR) filter uses finite measurements over the most recent time horizon of \( N \) discrete points. Basically, the unbiasedness can be met in FIR filters using two different strategies: 1) one may test an estimator by the unbiasedness condition or 2) one may embed the unbiasedness constraint into the filter design. We therefore recognize two different strategies: 1) one may test an estimator by the unbiasedness condition or 2) one may embed the unbiasedness (EU) filter is derived by minimizing the variance. For the real-time state-space model, the FIR-CU filter and smoother were found in [2] for CU by employing the maximum likelihood and in [3] for EU by minimizing the variance. For the real-time state space model, the FIR-CU filter and smoother were proposed by Shmaliy in \([4,5]\) for polynomial systems. In \([6]\), a \( p \)-shift unbiased FIR filter (UFIR) was derived as a special case of the OFIR filter. Here, the unbiasedness was checked a posteriori and the solution thus belongs to CU. Soon after, the UFIR filter \([6]\) was extended to time-variant systems \([7,8]\). For nonlinear models, an extended UFIR filter was proposed in \([9]\) and unified forms for FIR filtering and smoothing were discussed in \([10]\).

It has to be remarked now that all of the aforementioned FIR estimators related to real-time state-space model belong to the CU solutions. Still no optimal FIR estimator was addressed of the EU type. In this paper, we derive a new FIR filter, called OFIR-EU filter, by minimizing the mean square error (MSE) subject to the unbiasedness constraint. We also investigate properties of the OFIR-EU filter in a comparison with the OFIR and UFIR filters and KF.

2. STATE-SPACE MODEL AND PRELIMINARIES

Consider a linear discrete time-invariant model given with the state-space equations

\[
x_k = A x_{k-1} + B w_k, \quad (1)
\]
\[
y_k = C x_k + D v_k, \quad (2)
\]

in which \( k \) is the discrete time index, \( x_k \in \mathbb{R}^n \) is the state vector, and \( y_k \in \mathbb{R}^p \) is the measurement vector. Matrices \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \), \( C \in \mathbb{R}^{p \times n} \) and \( D \in \mathbb{R}^{p \times p} \) are time-invariant and known. We suppose that the process noise \( w_k \in \mathbb{R}^m \) and the measurement noise \( v_k \in \mathbb{R}^p \) are zero mean, \( E\{w_k\} = 0 \) and \( E\{v_k\} = 0 \), mutually uncorrelated and have arbitrary distributions and known covariances \( Q(i, j) = E\{w_i w_j^T\} \) for all \( i \) and \( j \), to mean that \( w_k \) and \( v_k \) are not obligatorily white Gaussian.

The state-space model (1) and (2) can be represented in the batch form on a discrete time interval \([l, k]\) with recursively computed forward-in-time solutions as

\[
X_{k,l} = A_{k-l} x_l + B_{k-l} W_{k,l}, \quad (3)
\]
\[
Y_{k,l} = C_{k-l} x_l + H_{k-l} W_{k,l} + D_{k-l} V_{k,l}, \quad (4)
\]

where \( l = k - N + 1 \) is a start point of the averaging horizon. The time-variant state vector \( X_{k,l} \in \mathbb{R}^{n \times 1} \), observation vector \( Y_{k,l} \in \mathbb{R}^{p \times 1} \), process noise vector \( W_{k,l} \in \mathbb{R}^{m \times 1} \), and observation noise vector \( V_{k,l} \in \mathbb{R}^{p \times 1} \) are specified as, respectively,

\[
X_{k,l} = [x_l^T x_l^T \cdots x_l^T]^{T}, \quad (5)
\]
\[
Y_{k,l} = [y_l^T y_l^T \cdots y_l^T]^{T}, \quad (6)
\]
\[
W_{k,l} = [w_l^T w_l^T \cdots w_l^T]^{T}, \quad (7)
\]
\[
V_{k,l} = [v_l^T v_l^T \cdots v_l^T]^{T}. \quad (8)
\]

The expanded model matrix \( A_{k-l} \in \mathbb{R}^{n \times n} \), process noise matrix \( B_{k-l} \in \mathbb{R}^{n \times m} \), observation matrix \( C_{k-l} \in \mathbb{R}^{p \times n} \), auxiliary matrix \( H_{k-l} \in \mathbb{R}^{p \times m} \), and measurement noise matrix \( D_{k-l} \in \mathbb{R}^{p \times p} \) are all time-invariant and dependent on the horizon length of \( N \) points. Model (1) and (2) suggests that these matrices can be written as, respectively

\[
A_{i} = \left[ (A^i)^T (A^i)^{-1} \right]^T, \quad (9)
\]

\[
B = \begin{bmatrix} B & AB & \cdots & A^{i-2}B & A^{i-1}B \\ 0 & B & \cdots & A^{i-2}B & A^{i-1}B \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & B & AB \\
0 & 0 & \cdots & 0 & B \end{bmatrix}, \quad (10)
\]
In the derivation of the OFIR-EU filter, the following lemma is used:

Lemma 1 The trace optimization problem is given by

\[
\begin{align*}
\text{arg min}_{K} \, & \, \text{tr} \left[ (KF - G)(KF - G)^T \right] \\
& + (KL - M)P(KL - M)^T + KS^2K^T, \\
\end{align*}
\]

subject to \( \Sigma_{KU=Z}|\theta \),

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

\[
K = \begin{bmatrix}
Z & G \\
M & \theta \left( U^T \Xi^{-1} U \right)^{-1} H P F^T \Xi^{-1} H \Xi^{-1} U^T \\
0 & \Xi^{-1} U^T \Xi^{-1}
\end{bmatrix},
\]

where \( \Xi = I - \theta U \left( U^T \Xi^{-1} U \right)^{-1} U^T \Xi^{-1} \) and

\[
\Xi = \begin{cases}
L P L^T + S, & \text{if } F = U, \, G = Z, \, \text{and } \theta = 1 \\
F H P^T + S, & \text{if } L = U, \, M = Z, \, \text{and } \theta = 1 \\
F H P^T + L P L^T + S, & \text{if } \theta = 0
\end{cases}.
\]

Proof: The proof can be obtained by modifying the results presented in [11, 12], which is omitted here due to the strict of paper length.

3.1 OFIR-EU Filter Design

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ e_{k} e_{k}^T \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ e_{k} e_{k}^T \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} - S_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

If to take into account constraint (18), provide the averaging, and rearrange the terms, (27) can be transformed to

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]

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

\[
K_{k}^{\text{OEU}} = \arg \min_{K_{k}} \text{tr} \left\{ \left( A^{N-1} x_{l} + B_{k-l} W_{k,l} \right) \right\}
\]
where the fact is invoked that the system noise vector \( W_{k,l} \) and the measurement noise vector \( V_{k,l} \) are pairwise independent. The auxiliary matrices are

\[
\Theta_w = E\{W_{k,l}W_{k,l}^T\}, \quad (29)
\]

\[
\Delta_v = D_{k-1}E\{V_{k,l}V_{k,l}^T\}D_{k-1}^T. \quad (30)
\]

Referring to Lemma 1 with \( \theta = 1 \), the solution to the optimization problem (28) can be obtained by neglecting \( L \) and \( M \) and using the replacements: \( F \rightarrow H_{k-1}, \ G \rightarrow B_{k-1}, \ H \rightarrow \Theta_w, \ U \rightarrow C_{k-1}, \ Z \rightarrow A^{N-1}, \) and \( S \rightarrow \Delta_v \). We thus have

\[
K^\text{OEU}_k = K^\text{OEU}_a + K^\text{OEU}_b, \quad (31)
\]

where

\[
K^\text{OEU}_a = A^{-N}(C_{k-l}^{-1}C_{k-1}^{-1}C_{k-l}^{-1}C_{k-1}^{-1}C_{k-l}^{-1}), \quad (32)
\]

\[
K^\text{OEU}_b = B_{k-l}\Theta_w H_{k-l}^{-1}(I - \Omega_{k-l}), \quad (33)
\]

in which

\[
\Omega_{k-l} = C_{k-l}(C_{k-l}^{-1}C_{k-1}^{-1}C_{k-l}^{-1}C_{k-1}^{-1}C_{k-l}^{-1}), \quad (34)
\]

\[
\Delta_{w+v} = \Delta_w + \Delta_v, \quad (35)
\]

\[
\Delta_w = H_{k-l}\Theta_w H_{k-l}^{-1}. \quad (36)
\]

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

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

\[
\hat{x}_{k,l} = (K^\text{OEU}_a + K^\text{OEU}_b)Y_{k,l}, \quad (37)
\]

where \( Y_{k,l} \in R^{N_p \times 1} \) is the measurement vector given by (6), \( K^\text{OEU}_a \) and \( K^\text{OEU}_b \) are given by (32) and (33) with \( C_{k-l} \) and \( B_{k-l} \) specified by (11) and (10), respectively.

**Proof:** The proof is provided by (24)-(36).

Note that the horizon length \( N \) for (37) should be chosen such that the inverse in \( K^\text{OEU}_a \) exists. In general, \( N \) can be set as \( N \geq n \), where \( n \) is the number of the model states. Table 1 summarizes 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 \( K^\text{OEU}_a \) and \( K^\text{OEU}_b \) according to (32) and (33) respectively, then the OFIR-EU estimate can be obtained at time index \( k \) by (37).

<table>
<thead>
<tr>
<th>Stage</th>
<th>Given: ( N \geq n ), ( l = k - N + 1 )</th>
<th>Find: ( K^\text{OEU}_a ) by (32) and ( K^\text{OEU}_b ) by (33)</th>
<th>Compute: ( \hat{x}_{k,l} = (K^\text{OEU}_a + K^\text{OEU}<em>b)Y</em>{k,l} )</th>
</tr>
</thead>
</table>

In this section, we are going to test the OFIR-EU filter with a two-state harmonic time-invariant state-space models in different noise environments. The main purpose is to show the effect of the unbiasedness condition embedded into the UFIR filter. The KF, UFIR and OFIR filters are employed as benchmarks when necessary. Similar examples can also be found in [7, 8, 13].

The two-state harmonic model can be specified by \( B = \begin{bmatrix} 1 & 1 \end{bmatrix}^T, \ C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \ D = 1, \) and

\[
A = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix}
\]

with \( \phi = \pi/32 \). Traditionally, we investigate the cases of a completely known model and system uncertainties.

We generate a process at 400 subsequent points with the initial states \( x_{10} = 1 \) and \( x_{20} = 0.1 \) and noise variances \( \sigma_w^2 = 1 \) and \( \sigma_v^2 = 10 \). The RMSEs \( \sqrt{\text{tr} J} \) computed as functions of \( N \) are exhibited in Fig. 1. One can see that KF performs best among all the filters, as the model used is accurate. One the other hand, the MSE in the OFIR-EU and OFIR filters become constant when \( N \geq N_{\text{opt}} \). This is a quite useful property of the OFIR-EU filter proposed. Specifically, it is not necessary to choose an optimal horizon for the OFIR-EU filter, a relative large horizon is always satisfied.

In order to show effect of the model uncertainties on the estimation errors, we augment the system matrix \( A \) as

\[
A = \begin{bmatrix} \cos \phi & \sin \phi + \delta \\ -\sin \phi + \delta & \cos \phi \end{bmatrix}
\]

where we set \( \delta = 0.4 \) if \( 160 \leq k \leq 180 \) and \( \delta = 0 \) otherwise.

The process is generated with \( x_{10} = 1, \ x_{20} = 0.1, \ \sigma_w = 0.1 \) and \( \sigma_v = 100 \) at 400 subsequent points. The instantaneous estimation errors produced by the KF and OFIR-EU filter for \( p \leq 1 \) are shown in Fig. 2, where \( p \) is
a correction parameter with $p^2\sigma_y$ and $\sigma_y/p^2$. It is seen that both filters produce negligible errors in the interval of first 160 points. Beyond this interval, the performance of both filters is deteriorated by the excursions when $\delta_k \neq 0$. Furthermore, one watches for transients which are limited with $N$ points in the OFIR-EU filter and last much longer in KF.

5. CONCLUSIONS

In this paper, the unbiasedness condition is embedded into the OFIR filter to obtain a new FIR filter-OFIR-EU filter, which can be considered as the optimal unbiased FIR filter. Unlike the OFIR filter, the OFIR-EU filter completely ignores the initial conditions. In terms of accuracy, the OFIR-EU filter is in between the UFIR and OFIR filters. Unlike in the UFIR filter which minimizes MSE by $N_{opt}$, MSES in the OFIR-EU and OFIR filters diminish with $N$ and these filters are thus full-horizon. Accordingly, the OFIR-EU filter in general demonstrates better robustness against temporary model uncertainties than KF.

Referring to the fact that optimal FIR filters are essentially the full-horizon filters but their batch forms are computationally inefficient, we now focus our attention on the fast iterative form for OFIR-EU filter and plan to report the results in near future.

REFERENCES


