NEW RESULTS IN NONLINEAR STATE ESTIMATION USING EXTENDED UNBIASED FIR FILTERING

Moises Granados-Cruz, Yuriy S. Shmaliy, Sanowar H. Khan, Choon Ki Ahn, Shunyi Zhao

Department of Electronics Engineering, Universidad de Guanajuato, Salamanca, 36885, Mexico
Department of Electronics Engineering, City University London, London, UK
School of Electrical Engineering, Korea University, Seoul, Korea
Key Lab. Adv. Process Contr. for Light Ind., Jiangnan University, Wuxi, P.R. China

ABSTRACT
This paper discusses two algorithms of extended unbiased FIR (EFIR) filtering of nonlinear discrete-time state-space models used in tracking and state estimation. The basic algorithm employs the extended nonlinear state and observation equations. The modified algorithm utilizes the nonlinear-to-linear conversion of the observation equation which is provided using a batch EFIR filter having small memory. Unlike the extended Kalman filter (EKF), both EFIR algorithms ignore the noise statistics and demonstrate better robustness against temporary model uncertainties. These algorithms require an optimal horizon in order to minimize the mean square error. Applications are given for robot indoor self-localization utilizing radio frequency identification tags.

1. INTRODUCTION
Diverse problems in navigation, tracking, robotics, communications, etc. often require nonlinear state estimation. A traditional tool here is the extended Kalman filter (EKF) [1] having strong features such as high accuracy, fast computation, easy coding, and small memory. Among the recognized flaws of EKF are the following: 1) its estimate can be biased if noise is nonadditive, 2) it may diverge if nonlinearities and noise are large [2], and 3) its accuracy can be low if noise covariances are not well specified or ill-conditioned and noise is non white Gaussian, heavy-tailed, or Gaussian with outliers [3].

In line with the recursive Kalman approach, several other approaches have also been developed during decades in order to find a more robust solution [4–11, 14, 15]. The unscented Kalman filter (UKF) was proposed in [9] to transfer the mean and variance through nonlinearities with higher accuracy than in EKF when the model is strongly nonlinear. A grid-based method which was worked out to approximate the posterior process distribution has resulted in the hidden Markov model (HMM) filters [10]. A sequential Monte Carlo (SMC) method also known as a particle filter (PF) [11] was developed for Bayesian models associated with Markov chains [12]. A new filtering approach was developed in [13] for nonlinear pairwise models. A review of these and other nonlinear filters can be found in [14].

An efficient alternative to the recursive EKF is the iterative extended finite impulse response (EFIR) filter [15]. Unlike the EKF, UKF, and optimal FIR (OFIR) filters [16–18], the EFIR filter totally ignores the noise statistics and initial error statistics. Similarly to PFs, the EFIR filter exploits most recent past measurements which number is required to be optimal $N_{opt}$. A scalar $N_{opt}$ can be ascertained by using test reference measurements or even via regular measurements without a reference signal [19], thus in a way much simpler than for the noise statistics. The EFIR filter belongs to a regression-based family of Gauss’s least squares estimators which often give accuracy that is superior to the best available EKF [14]. In what follows, we propose, develop, and discuss two efficient iterative EFIR filtering algorithms.

2. NONLINEAR MODEL
Let us consider a process represented in state space with the nonlinear state and observation equations, respectively,

\[ x_n = f(x_{n-1}, u_n, w_n, e_n), \quad \text{(1)} \]
\[ z_n = h(x_n, v_n), \quad \text{(2)} \]

in which $x_n \in \mathbb{R}^k$ is the state vector, $u_n \in \mathbb{R}^l$ is the input vector, and $f(\cdot)$ and $h(\cdot)$ are nonlinear time-varying functions. All random components are zero mean white Gaussian and uncorrelated. Namely, the process noise $w_n \in \mathbb{R}^p$, the input noise $e_n \in \mathbb{R}^m$, and the observation noise $v_n \in \mathbb{R}^m$ have the properties: $E(w_n) = 0$, $E(e_n) = 0$, $E(v_n) = 0$, and $E(w_i e_j^T) = 0$, $E(w_i v_j^T) = 0$, and $E(v_i v_j^T) = 0$ for all $i$ and $j$. The noise covariances are depicted as $Q = E(w_n w_n^T)$, $L = E(e_n e_n^T)$, and $R = E(v_n v_n^T)$.

In order to estimate $x_n$ using methods of linear filtering, (1) and (2) need to be expanded to the 1-order Taylor series [1]:

\[ f_n = F_n x_{n-1} + u_n + W_n w_n + E_n e_n + \xi_n, \quad \text{(3)} \]
\[ h_n = H_n x_n + z_n + T_{nv} v_n + z_n, \quad \text{(4)} \]

where $F_n = \frac{\partial f_n}{\partial x} |_{x_{n-1}}$, $W_n = \frac{\partial f_n}{\partial w} |_{x_{n-1}}$, $E_n = \frac{\partial f_n}{\partial e} |_{x_{n-1}}$, $T_n = \frac{\partial h_n}{\partial x} |_{x_n}$, and $H_n = \frac{\partial h_n}{\partial z} |_{x_{n-1}}$ are Jacobian and

\[ u_n = f_n(x_{n-1}, u_n, 0, 0) - F_n x_{n-1}, \]
\[ z_n = h_n(x_n) - H_n x_n \]

are known. Here, $\hat{x}_n$ is the estimate\footnote{$\hat{x}_n$ means the estimate at $n$ via measurement from the past to $k$. Below, we use the following notations: $\hat{x}_n \triangleq \hat{x}_n^+$ and $\hat{x}_n \triangleq \hat{x}_n^-$} and $\hat{x}_n^-$ is the prior estimate of $x_n$. The residuals $\xi_n$ and $\zeta_n$ are supposed to be small if the model is sufficiently smooth.
The 1-order expanded state-space model is thus

$$x_n = F_n x_{n-1} + u_n + \hat{w}_n + \xi_n ,$$  \hspace{1cm} (5)

$$z_n = H_n x_n + z_n + \hat{v}_n + \zeta_n ,$$  \hspace{1cm} (6)

where the zero mean noise vectors $\hat{w}_n$, $\hat{v}_n$, and $\zeta_n$ have the covariances $Q_n = F_n Q F_n^T$, $L_n = E_n L E_n^T$, and $R_n = T_n R T_n^T$. Provided (5) and (6), the EKF can be coded as in Table 1, in which the initial state estimate $\hat{x}_0$ and covariances $P_0$, $R$, $Q$, and $L$ are supposed to be known. The prior estimation error $P_0$ and estimation error $P_n$ are defined by

$$P_n = E\left\{ (x_n - \hat{x}_n)(x_n - \hat{x}_n)^T \right\} ,$$  \hspace{1cm} (7)

$$P_n = E\left\{ (x_n - \hat{x}_n)(x_n - \hat{x}_n)^T \right\} .$$  \hspace{1cm} (8)

The extended EFIR filtering algorithms are discussed next.

### 2.1 Extended unbiased FIR filtering

Unlike the recursive EKF, the iterative EFIR filter [15] utilizes measurements $y_n$ available on an interval of $N$ past neighboring points from $m = n - N + 1$ to $n$. The EFIR filter totally ignores the covariances $R$, $Q$, $L$, and $P_0$. Instead, it requires an optimal horizon of $N_{\text{opt}}$ points. There are at least two ways to find a scalar $N_{\text{opt}}$. via the test measurements implying a known model $x_n$ by minimizing the trace of $P_n$,

$$N_{\text{opt}} = \arg \min_N \left\{ \text{tr} P(N) \right\} ,$$  \hspace{1cm} (9)

or utilizing measurements with no reference [19].

The EFIR filtering estimate has the Kalman form

$$\hat{x}_t = \hat{x}_t^- + K_t [y_t - H_t (\hat{x}_t^-)] ,$$  \hspace{1cm} (10)

in which $l$ ranges from $m + K$ to $n$, where $K$ is the number of the states. For each time index $n$, the output is taken when $l = n$. The bias correction gain

$$K_l = G_l H_l^T$$  \hspace{1cm} (11)

is defined and updated iteratively via the generalized noise power gain (GNPG)

$$G_l = [H_l^T H_l + (F_l G_{l-1} F_l^T)^{-1}]^{-1} .$$  \hspace{1cm} (12)

To avoid singularities, iterative calculation of (10) starts at $m + K$ and all values at $s = m + K - 1$ are computed in short batch forms as [21]

$$\hat{x}_s = F_s \cdots F_{m+1} \Lambda_{s,m} H_{s,m}^T y_{s,m} ,$$  \hspace{1cm} (13)

$$G_s = F_s \cdots F_{m+1} \Lambda_{s,m} H_{s,m}^T F_s \cdots F_{m+1} ,$$  \hspace{1cm} (14)

where $\Lambda_{s,m} = (H_{s,m}^T H_{s,m})^{-1}$ and

$$Y_{s,m} = \begin{bmatrix} y_s^T \cdots y_{m}^T \end{bmatrix}^T ,$$  \hspace{1cm} (15)

$$H_{s,m} = \begin{bmatrix} H_s \cdots H_{m+1} \end{bmatrix} .$$  \hspace{1cm} (16)

Unlike the EKF relying on $\hat{x}_0$, the EFIR filter needs $N_{\text{opt}}$ known initial estimates or linear measurements united in a vector $y_n$. Since $y_n$ may be unavailable in nonlinear modelling, the following options can be considered:

1) If $y_n$ is available, then compute $\hat{x}_n$ via (13) using (15) and (16) and set $y_n = \hat{x}_n$. Otherwise, if all of the states are observable by $x_n$, a solution to $z_n = h_n(x_n,0)$ for $x_n$ can be employed as $y_n$ [15].

2) If (2) is linear, then set $y_n = z_n$. The iterative EFIR filter can be coded as Table 2. Provided $\hat{x}_n$, $y_n$, and $u_n$, it needs only $N$ and $K$ to start computing and updating all the vectors and matrices. No noise statistics are involved. In this code, two specific can be taken into account: 1) because the GNPG is almost unity on an interval of $K$ points, $G_s$ in many cases can be substituted with an identity matrix $I$ and 2) the EFIR algorithm operates in $N_{\text{opt}} - 1$ times slower than EKF owing to iterations.
Table 3: EFIR-2 Filtering Algorithm for (19) and (20)

<table>
<thead>
<tr>
<th>Input</th>
<th>( \hat{x}_n, u_n, K, N )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>for ( n = N - 1 ) to 0 do</td>
</tr>
<tr>
<td>2:</td>
<td>( m = n - N + 1 ), ( s = m + K - 1 )</td>
</tr>
<tr>
<td>3:</td>
<td>( \hat{x}_s = { \hat{y}_s, \text{if } s &lt; N - 1 } )</td>
</tr>
<tr>
<td></td>
<td>( \hat{y}_s, \text{if } s \geq N - 1 )</td>
</tr>
<tr>
<td>4:</td>
<td>( G_s = F_s \cdots F_{m+1} F_T^{m+1} \cdots F_s^T )</td>
</tr>
<tr>
<td></td>
<td>Otherwise, set ( G_s = I )</td>
</tr>
<tr>
<td>5:</td>
<td>for ( l = m + K : n ) do</td>
</tr>
<tr>
<td>6:</td>
<td>( \hat{x}<em>l = f_l(\hat{x}</em>{l-1}, u_l, 0, 0) )</td>
</tr>
<tr>
<td>7:</td>
<td>Update: ( F_l )</td>
</tr>
<tr>
<td>8:</td>
<td>( G_l = [I + (F_l G_{l-1} F_l^T)^{-1}]^{-1} )</td>
</tr>
<tr>
<td>9:</td>
<td>( \hat{x}_l = \hat{x}_l + G_l(\hat{y}_l - \hat{x}_l) )</td>
</tr>
<tr>
<td>10:</td>
<td>and for</td>
</tr>
<tr>
<td>11:</td>
<td>( \hat{x}_n = \hat{x}_n )</td>
</tr>
<tr>
<td>12:</td>
<td>and for</td>
</tr>
<tr>
<td>Output</td>
<td>( \hat{x}_n )</td>
</tr>
</tbody>
</table>

2.2 Nonlinear-to-linear observation conversion

The batch estimate (13) offers another opportunity to provide nonlinear EFIR filtering. Since noise reduction is insignificant on a minimum allowed horizon of \( N = K \) points, one may consider the unbiased estimate

\[
\hat{y}_n = F_n \cdots F_{v+1}(H_n^T H_{n,v})^{-1} H_n^T z_{n,v},
\]

(17)

where \( v = n - K + 1 \) and \( Z_{n,v} = [(z_n - \tilde{z}_n)^T \cdots (z_n - \tilde{z}_n)^T]^T \),

as a linear measurement of \( x_n \) and transform (5) and (6) for negligible residuals to

\[
\hat{x}_n = F_n x_{n-1} + \hat{e}_n + w_n,
\]

(19)

\[
\hat{y}_n = x_n + v_n,
\]

(20)

where \( \hat{x}_n = x_n - \hat{u}_n \). As can be seen, \( \hat{y}_n \) has the same dimensions as \( u_n \), unlike the original measurement \( z_n \). The covariance of \( \hat{y}_n \) is defined by

\[
\tilde{R}_n = E\{ (\hat{y}_n - x_n)(\hat{y}_n - x_n)^T \}.
\]

It thus follows that (17) serves in (20) as a converter of the nonlinear observation (6) to the linear one (20) with restrictions peculiar to extended nonlinear estimators. Below, we test the EFIR-1 and EFIR-2 algorithms in a comparison to the EKF by a robot localization problem.

3. Applications

Consider a robot travelling in direction \( d \) with coordinates \( x_0 \) and \( y_0 \) on an indoor floorspace (white curve in Fig. 1). The robot measures distances to two radio frequency identification (RFID) tags, A and B, and its trajectory is controlled by the left and right wheels. The distance between the wheels is \( b = 1 \) m and the incremental distances vehicle travels by these wheels are \( d_l \) and \( d_r \). The pose angle \( \Phi_0 \) is measured with an imbedded fiber optic gyroscope (FOG) [20].

The robot extended state-space model is given by (5) and (6) in which \( x_n = [x_n y_n \Phi_n]^T \), \( u_n = [d_l d_r]^T \), \( w_n = [w_{x_n} w_{y_n} w_{\Phi_n}]^T \), \( e_n = [e_n e_{\Delta x} e_{\Delta y}]^T \), \( T_n = I \), \( W_n = F_n \),

\[
F_n = \begin{bmatrix}
1 & 0 & -d_l \sin(\Phi_{n-1} + \frac{b(\Phi_n - \Phi_{n-1})}{2}) \\
0 & 1 & d_l \cos(\Phi_{n-1} + \frac{b(\Phi_n - \Phi_{n-1})}{2}) \\
0 & 0 & 1
\end{bmatrix},
\]

(21)

\[
E_n = \frac{1}{2} \begin{bmatrix}
be_{c_n} + dq_{e_n} & be_{c_n} + dq_{e_n} & be_{c_n} + dq_{e_n} \\
be_{c_n} + dq_{e_n} & be_{c_n} + dq_{e_n} & be_{c_n} + dq_{e_n} \\
-2e_{c_n} & -2e_{c_n} & 2e_{c_n}
\end{bmatrix},
\]

(22)

\[
H_n = \begin{bmatrix}
\frac{\sin(y_n - y_{\Phi})}{\Phi} & \frac{\sin(y_n - y_{\Phi})}{\Phi} & 0 \\
0 & 0 & 1
\end{bmatrix},
\]

(23)

where

\[
u_{1n} = \sqrt{(y_1 - \tilde{y}_n)^2 + (x_1 - \tilde{x}_n)^2 + c_1^2},
\]

\[
u_{2n} = \sqrt{(y_2 - \tilde{y}_n)^2 + (x_2 - \tilde{x}_n)^2 + c_2^2},
\]

\[
\phi_{\Phi} \equiv \frac{1}{b}(d_{R_n} - d_{L_n}),
\]

\[
e_{c_n} = \cos \left( \frac{\Phi_{n} - \Phi_{n-1}}{2} \right) \text{ and } e_{\Delta \phi} = \sin \left( \frac{\Phi_{n} - \Phi_{n-1}}{2} \right).
\]

We allow all of the covariance matrices be diagonal and set the standard deviations \( \sigma_\Phi = \sigma_{\Delta \phi} = \sigma_{\Phi} = 1 \text{ mm} \), \( \sigma_{\phi} = 0.5^\circ \), \( \sigma_{\Delta \phi} = 5 \text{ sm} \), and \( \sigma_{\Delta \phi} = 2^\circ \). The reader range is supposed to be \( r = 6 \) m. We place a tag A at (0, 6) m and tag B at (0, 0) m and let \( d_l = 0.12 \) mm and \( d_r = 0.24 \) mm. Simulation is provided at 5000 points with time interval \( T \) for \( G_2 = 1 \). In Fig. 1, direct measurements of \( x_n \) and \( y_n \) are unavailable. We therefore solve the inverse problem in (6) for \( x_1 = x_2 = y_2 = 0 \) and \( y_1 = 6 \) m, go to “linear” measurements \( \hat{x}_n \) and \( \hat{y}_n \), united it in a measurement vector \( y_n = [\hat{x}_n \hat{y}_n]^T \), and depict in Fig. 1 as “actual measurement”. The “converted measurement” is provided by (17). The optimal horizon \( N_{\text{opt}} = 84 \) was found for a reference test measurement.

Typical instantaneous errors produced by the EKF and EFIR-1 and EFIR-2 filters are sketched in Fig. 2. To be closer to a practical situation, in this simulation we consider the worst case of not fully known noise statistics by introducing a correction coefficient \( p \) to the covariance matrices as \( p^2 R, Q/p^2 \) and \( L/p^2 \). As can be seen in Fig. 2 (\( p = 3 \)), the correction coefficient \( p \) strongly affects the EKF estimate, whereas the EFIR estimate is \( p \)-invariant. Next, we compute the MSE for each filter by the root square of the trace of the estimation error matrix (8), excluded the third state having an angular measure. The MSEs computed over 30 subsequent Monte Carlo runs are shown in Fig. 2. Observing this figure, one infers that the EKF\( (p = 1) \) and EFIR-1 and EFIR-2 filters produce similar errors. In contrast, errors in the noise covariances lead to larger errors in both EKF\( (p = 3) \) and EFIR\( (p = 5) \) and do not affect the EFIR filters. Moreover, excursions in MSEs indicate that EKF with \( p > 1 \) is addicted to divergence and that this addiction grows with \( p \). Note that even a stronger addiction to divergence of EKF was shown.
Figure 1: Robot actual trajectory $x_n$, “actual measurement” by solving the inverse problem, and “converted measurement” by (17), excluded $N_{\text{opt}}$ first points.

Figure 2: Typical time-domain errors of EKF and EFIR-1 and EFIR-2 filters corresponding to Fig. 1 for $p = 3$ and $N_{\text{opt}} = 84$: (a) coordinate $x$, (b) coordinate $y$, and (c) heading $\Phi$. 
their fast forms utilizing recursions. Other applications for the algorithms proposed and developed require much longer operation to complete iterations. A payment for this operation remains at the same error level. Of practical importance also is that the only tuning scalar value required by the EFIR filter can easily be specialized via test measurements or even using regular measurements with no reference, thus in a way much easier than for the noise statistics. Besides, the determination of \( N_{\text{opt}} \) implies much smaller cost, especially if the process is time-varying. A payment for these advantages of EFIR filtering is an \( N_{\text{opt}} - 1 \) times longer operation required to complete iterations.

Referring to such useful properties, we now consider other applications for the algorithms proposed and develop their fast forms utilizing recursions.

4. CONCLUSIONS

The EFIR filtering algorithms proposed, developed in part, and discussed in this paper have the following useful properties. Unlike the EKF filter, both the EFIR-1 and FIR-2 algorithms are insensitive to the imprecisely defined noise statistics. Practically, this means that it is only within a narrow region around the ideal conditions that the EKF has better accuracy than EFIR. Otherwise, errors in the EKF grow rapidly and result in the divergence, whereas the EFIR filter ignoring noise statistics remains at the same error level. The EFIR filtering algorithms proposed, developed in part, and discussed in this paper have the following useful properties. Unlike the EKF filter, both the EFIR-1 and FIR-2 algorithms are insensitive to the imprecisely defined noise statistics. Practically, this means that it is only within a narrow region around the ideal conditions that the EKF has better accuracy than EFIR. Otherwise, errors in the EKF grow rapidly and result in the divergence, whereas the EFIR filter ignoring noise statistics remains at the same error level. Of practical importance also is that the only tuning scalar value \( N_{\text{opt}} \) required by the EFIR filter can easily be specialized via test measurements or even using regular measurements with no reference, thus in a way much easier than for the noise statistics. Besides, the determination of \( N_{\text{opt}} \) implies much smaller cost, especially if the process is time-varying. A payment for these advantages of EFIR filtering is an \( N_{\text{opt}} - 1 \) times longer operation required to complete iterations.

Referring to such useful properties, we now consider other applications for the algorithms proposed and develop their fast forms utilizing recursions.

REFERENCES