AN APPROACH TO NONLINEAR STATE ESTIMATION USING EXTENDED FIR FILTERING

Shunya Zhao, *Juan Pomárico-Franquiz, *Yuriy S. Shmaliy

Key Laboratory of Advanced Process Control for Light Industry, Jiangnan University, Wuxi, P.R. China
Department of Electronics Engineering, Universidad de Guanajuato, Salamanca, 36885, Mexico

ABSTRACT
A new technique called extended finite impulse response (EFIR) filtering is developed to nonlinear state estimation in discrete time state space. The EFIR filter belongs to a family of unbiased FIR filters which completely ignore the noise statistics. An optimal averaging horizon of \( N_{\text{opt}} \) points required by the EFIR filter can be determined via measurements with much smaller efforts and cost than for the noise statistics. These properties of EFIR filtering are distinctive advantages against the extended Kalman filter (EKF). A pay-ment for this is an \( N_{\text{opt}} - 1 \) times longer operation which, however, can be reduced to that of the EKF by using parallel computing. Based on extensive simulations of diverse nonlinear models, we show that EFIR filtering is more successful in accuracy and more robust than EKF under the unknown noise statistics and model uncertainties.

1. INTRODUCTION
Nonlinear estimation problems arise in diverse fields of applications such as navigation, tracking, robotics, communications, control, etc. A traditional tool here is the extended Kalman filter (EKF) [1,2] having strong engineering features such as high accuracy, fast computation, easy coding, and small memory. However, EKF has several widely recognized flaws: 1) its estimate can be biased if noise is nonadditive, 2) it may diverge if nonlinearities and noise are large [3], and 3) its accuracy can be low if noise covariances are not well specified or ill-conditioned and noise is nonwhite Gaussian, heavy-tailed, or Gaussian with outliers [4].

Because it is desirable to have an estimator that is more robust than EKF, several other approaches were developed during decades [5–17]. The technique called the unscented transformed was introduced in [12] to transfer the mean and variance through nonlinearities. A relevant filter called the unscented Kalman filter (UKF) has demonstrated better performance than EKF when the model is highly nonlinear. For continuous-time state-space models decomposed into “cells,” a grid-based method was worked out to approximate the posterior probability density function (pdf) of the process. The approach has resulted in the hidden Markov model (HMM) filters [13,14]. A sequential Monte Carlo (SMC) method also known as a particle filter (PF) [15] was developed to estimate Bayesian models associated with Markov chains in discrete-time domain. The reader can find a comprehensive review of these and other nonlinear filters in [16].

A novel alternative to the recursive EKF is the iterative extended finite impulse response (EFIR) filter [17,18]. Unlike the EKF, UKF, and optimal FIR (OFIR) filters [19,20], 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 \( N_{\text{opt}} \). A scalar \( N_{\text{opt}} \) can be determined by using test reference measurements or even via regular measurements without a reference signal [21], thus in a way much easier than that used to determine the noise statistics required by the Kalman filter. Finally, the EFIR filter belongs to a regression-based family of Gauss’s least squares estimators which are known to often give accuracy that is superior to the best available EKF [16]. Referring to such properties of EFIR filtering, one may expect new solutions to nonlinear estimation problems in different areas of applications. Thus, efficient EFIR algorithms are required to meet practical needs. Below, we consider a general nonlinear discrete-time state-space model, develop an iterative EFIR filtering algorithm, and learn its properties in a comparison with EKF based on two examples.

2. EXTENDED UNBIASED FIR FILTERING
In order to provide state estimation on a finite interval of \( N \) points, in this section we consider a general nonlinear discrete-time state-space model and develop an iterative EFIR algorithm.

2.1 Nonlinear State-Space Model
Consider a nonlinear process represented in state space with the state and observation equations,

\[
x_n = f_n(x_{n-1}, u_n, w_n, e_n), \quad (1)
\]

\[
z_n = h_n(x_n, v_n), \quad (2)
\]

in which \( x_n \in \mathbb{R}^L \) is the state vector, \( u_n \in \mathbb{R}^L \) is the input vector, \( z_n \in \mathbb{R}^M \) is the measurement vector, and \( f_n(\cdot) \) and \( h_n(\cdot) \) are nonlinear time-varying functions. We suppose that all random components are zero mean white Gaussian and uncorrelated. Namely, the process noise \( w_n \in \mathbb{R}^L \), the input noise \( e_n \in \mathbb{R}^L \), 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 \( E\{w_n e_n^T\} = 0, E\{w_n v_n^T\} = 0, \) and \( E\{v_n e_n^T\} = 0 \) 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 \( R = E\{v_n v_n^T\} \).

To apply a technique such as Kalman filtering, both (1) and (2) need to be expanded to the 1-order or 2-order Taylor series [1,2]. Aimed at demonstrating advantages of the FIR approach and referring to the fact that the 2-order expansion gives no definitive advantages [17, 22], below we employ only the 1-order expansions of \( f_n(\cdot) \) at \( n - 1 \) and \( h_n(\cdot) \) at \( n \) under the following suppositions. We think that \( u_n \) is slow enough and such that the difference \( u_n - u_{n-1} \) is insignificant. We also allow the initial values to be known, to
mean that the noise components at the start point are zeros. Accordingly, the expanded nonlinear functions become

\[ f_n = F_n x_{n-1} + u_n + W_n w_n + E_n e_n + \xi_n, \]

\[ h_n = H_n x_n + \bar{z}_n + V_n v_n + \xi_n, \]

where \( F_n = \frac{\partial f}{\partial x} \bigg|_{x_{n-1}}, \) \( W_n = \frac{\partial f}{\partial w} \bigg|_{x_{n-1}}, \) \( E_n = \frac{\partial f}{\partial e} \bigg|_{x_{n-1}}, \) \( T_n = \frac{\partial f}{\partial \xi} \bigg|_{x_{n-1}}, \) \( H_n = \frac{\partial h}{\partial x} \bigg|_{x_{n-1}}, \) \( \bar{z}_n = \text{Jacobian and both} \ u_n = f'(\hat{x}_{n-1}, u_n, 0, 0) \) \( + F_n \hat{x}_{n-1} \) \( \) and \( \bar{z}_n = h'(\hat{x}_n) - H_n \hat{x}_n \) are known. Here, \( \hat{x}_n \) is the estimate and \( \hat{x}_n \) is the prior estimate of \( x_n. \) The residuals \( \xi_n \) and \( \xi_n \) are supposed to be small if the model is sufficiently smooth.

The 1-order expanded state-space model is thus

\[ x_n = F_n x_{n-1} + u_n + \bar{w}_n + \bar{v}_n + \xi_n, \]

\[ z_n = H_n x_n + \bar{z}_n + V_n v_n + \xi_n, \]

where the zero mean noise vectors \( \bar{w}_n = W_n w_n, \) \( \bar{e}_n = E_n e_n, \) \( \bar{v}_n = V_n v_n \) have the covariances \( Q_n = F_n Q F_n^T, \) \( L_n = E_n L E_n^T, \) and \( R_n = T_n R T_n^T, \) respectively.

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 covariance matrix \( P_n \) and estimation error covariance matrix \( P_n \) are defined as

\[ P_n = E\{(x_n - \hat{x}_n)(x_n - \hat{x}_n)^T\}, \]

\[ P_n = E\{(x_n - \hat{x}_n)(x_n - \hat{x}_n)^T\}. \]

and we notice again that the required noise statistics \( P_0, R, \) \( Q, \) and \( L \) are typically not well known to the engineer [4] that may cause unacceptable errors in EKF.

### 2.2 EFIR Filtering Algorithm

Unlike the recursive EKF, the iterative EFIR filter [17] utilizes measurements \( z_n \) 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 averaging interval of \( N_{\text{opt}} \) points in order to minimize the mean square error. There are two simple ways to find \( N_{\text{opt}}: \)

- Via test measurements implying a known model \( x_n \) by minimizing the trace of \( P_n, \)

\[ N_{\text{opt}} = \arg \min \{ \text{tr} P(N) \}. \]

- Utilizing measurements with no reference signal as shown in [21].

The EFIR filtering estimate has the Kalman form

\[ \hat{x}_n = \hat{x}_n + K_n[z_n - h'(\hat{x}_n)], \]

in which an iterative variable \( l \) 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

\[ K_n = G_n H_n^T \]

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

\[ G_n = [H_n^T H_n + (F_n G_n H_n^T)]^{-1}. \]

To avoid singularities, iterative computation of (10) starts at \( m + K \) and all values at \( s = m + K - 1 \) are computed in short batch forms as [24]

\[ \hat{x}_n = F_s \ldots F_{n+1} \Lambda_{s=n} H_{s,m}^T Y_{s,m}, \]

\[ G_n = F_s \ldots F_{n+1} \Lambda_{s=n} F_{n+1} \ldots F_s, \]

where \( \Lambda_{s=n} = (H_{s,m} H_{m,n})^{-1} \)

\[ Y_{s,m} = \begin{bmatrix} Y_{s,1} & \ldots & Y_{s,m-1} & Y_{s,m}^T \end{bmatrix}, \]

\[ H_{s,m} = \begin{bmatrix} H_{1} & F_{1} & \ldots & F_{m+1} \\ H_{m+1} F_{m+1} \end{bmatrix}. \]


<table>
<thead>
<tr>
<th>Table 1: EKF Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> ( z_n, x_0, P_0, R, Q, L )</td>
</tr>
<tr>
<td><strong>Output:</strong> ( \hat{x}_n )</td>
</tr>
<tr>
<td>1. for ( n = 1 ) to ( M ) do</td>
</tr>
<tr>
<td>2. ( \hat{x}<em>n = \hat{x}</em>{n-1} + u_n + \bar{w}_n + \bar{v}_n + \xi_n )</td>
</tr>
<tr>
<td>3. ( P_n = F_n P_{n-1} F_n^T + W_n Q W_n^T + E_n L E_n^T )</td>
</tr>
<tr>
<td>4. ( K_n = P_n H_n^T (H_n P_n H_n^T + T_n R_n T_n^T)^{-1} )</td>
</tr>
<tr>
<td>5. ( \hat{x}_n = \hat{x}_n + K_n [z_n - h'(\hat{x}_n)] )</td>
</tr>
<tr>
<td>6. ( P_n = (I - K_n H_n) P_n )</td>
</tr>
<tr>
<td>and for</td>
</tr>
</tbody>
</table>

Note: \( \hat{x}_n \) means the estimate at \( n \) via measurement from the past to \( k. \) Below, we use the following notations: \( \hat{x}_n \triangleq \hat{x}_{n|n} \) and \( \hat{x}_n \triangleq \hat{x}_{n|n-1}. \)
As examples of applications, we first consider indoor robot localization provided using radio frequency identification (RFID) tags. We then consider tracking of a moving object with temporary model uncertainties.

### 3. APPLICATIONS

#### 3.1 Robot Localization

A robot travels in direction $d$ with coordinates $x_n$ and $y_n$ on an indoor floorspace. It measures distances to two 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 $\Phi_n$ is measured with an imbedded fiber optic gyroscope (FOG) [23].

The robot expanded state-space model is (5) and (6) in which $x_n = [x_n \ y_n \ \Phi_n]^T$, $u_n = [d_L \ d_R]^T$, $w_n = [w_{x_n} \ w_{y_n} \ w_{\Phi_n}]^T$, $n_n = [\epsilon_{L_n} \ \epsilon_{R_n} \ \epsilon_{\Phi_n}]^T$, $T_n = I$, $W_n = F_n$.

$$F_n = \begin{pmatrix} 1 & 0 & -d_n \sin(\Phi_n + \frac{1}{2} \phi_n) \\ 0 & 1 & d_n \cos(\Phi_n + \frac{1}{2} \phi_n) \\ 0 & 0 & 1 \end{pmatrix},$$  \hspace{1cm} (17)$$

$$E_n = \frac{1}{2b} \begin{pmatrix} b e_n - d_n e_n & b e_n - d_n e_n & -2 \\ b e_n - d_n e_n & b e_n - d_n e_n & 2 \\ 0 & 0 & 1 \end{pmatrix},$$  \hspace{1cm} (18)$$

$$H_n = \begin{pmatrix} 0 & 0 & 0 \\ \frac{\tilde{s}_n - s_n}{u_n} & \frac{\tilde{s}_n - s_n}{u_n} & 0 \\ \frac{\tilde{y}_n - y_n}{u_n} & \frac{\tilde{y}_n - y_n}{u_n} & 0 \end{pmatrix},$$  \hspace{1cm} (19)$$

where $u_{1n} = \sqrt{(y_1 - \tilde{y}_n)^2 + (x_1 - \tilde{x}_n)^2 + c_1^2}$, $u_{2n} = \sqrt{(y_2 - \tilde{y}_n)^2 + (x_2 - \tilde{x}_n)^2 + c_2^2}$, $d_n = \frac{1}{2}(d_{Rn} + d_{Ln})$, $\phi_n \approx \frac{1}{b}(d_{Ln} - d_{Rn})$, $\epsilon_n = \cos(\Phi_n + \frac{\phi_n}{2})$, and $e_{sn} = \sin(\Phi_n + \frac{\phi_n}{2})$. We allow all the covariance matrices to be diagonal and set the standard deviations $\sigma_2 = \sigma_3 = \sigma_4 = \sigma_5 = 1$ mm, $\sigma_6 = 0.5^\circ$, $\sigma_7 = \sigma_8 = 5$ sm, and $\sigma_9 = 2^\circ$. The reader range is supposed to be $r = 6$ m. We place a tag $A$ at $(0.6, 0)$ 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 $T$ allowing $G_n = I$. By test measurements, we obtain the model $x_n$ and find $N_{opt} = 74$ by minimizing the trace of $P_n$ via (9) as shown in Fig. 2a. Under such conditions, the estimates sketched in Fig. 1 can be considered as the best available by the EFIR filter and EKF. It is seen that the estimates are consistent, although both filters produce larger errors close to the boundary linking the tags.

Because an ideal situation is unfeasible, we next learn effect of errors in the noise covariances on EKF estimates. In doing so, we introduce a correction coefficient $p$ to the covariance matrices as $P^e = P + P^p$, and compute the trace of $P^e(p)$ using (8) for EKF as shown in Fig. 2b. Here, we also depict the $p$-invariant trace of $P$ for the EFIR filter ($N_{opt} = 74$). As expected, the EKF is a bit more accurate than EFIR filter in the ideal case of $p = 1$. However, that is only when $0.6 < p < 2$ that EKF outperforms the EFIR filter with an insignificant difference of about 0.5 mm. Otherwise, the EFIR filter is more accurate. Since a scalar $N_{opt}$ can be found in a way much easier than that required for $R$, $Q$, and $L$, we consider it as a distinctive advantage of the EFIR filter.

Note that, for the sake of correctness, the covariance matrices must first be specified in continuous time and then converted to discrete time that requires additional mathematical efforts.
Moreover, cost measurements are commonly required to determine $R$, $Q$, and $L$.

Practical experience shows that errors in the determination of $R$, $Q$, and $L$ can be large [4]. On the other hand, $N_{\text{opt}}$ can be found accurately even without a reference signal [21]. We therefore admit $p = 5$ and $N_{\text{opt}} = 74$ and take a more precise look at possible estimation errors $e_n = x_n - \hat{x}_n$ in the time domain. Typical results sketched in Fig. 3 reveal larger “slow” noise in all EKF estimates. In an opposite case of $p < 0.5$, all EKF estimates are accompanied with larger “fast” noise. Moreover, the EKF has appeared to be strongly addicted to divergence when $p < 0.5$. This fact still unknown in nonlinear filtering may help viewing the Kalman divergence from the other side. Returning back to Fig. 1, we finally notice that in the ideal case of $p = 1$ the EKF and EFIR estimates do not get away essentially from each other and are almost indistinguishable.

### 3.2 Tracking with Temporary Model Uncertainties

We now consider a typical problem when two distance measurement stations (DMSs) are located at $(0,0)$ and $(0,a = 50 \text{ m})$ and that a moving object and DMSs are in a horizontal plane. Each DMS transmits a pulse that is reflected from the object and returns back to DMS. The transit time is interpreted in terms of distance $d_1$ or $d_2$.

We suppose that an object moves in the presence of noise along each of the axes and has four states ($K = 4$): $x_{1n}$ is the coordinate $x$; $x_{2n}$ velocity along $x$; $x_{3n} > 0$ coordinate $y$; and $x_{4n}$ velocity along $y$. The behavior is modeled with equations

\[
x_n = Ax_{n-1} + w_n, \tag{20}
\]

\[
y_n = h_n(x_n) + v_n, \tag{21}
\]

in which $x_n = [x_{1n} x_{2n} x_{3n} x_{4n}]^T$ and

\[
A = \begin{bmatrix} 1 & \tau + \delta_n & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tau + \delta_n \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{22}
\]

where $\delta_n \neq 0$ if $400 \leq n \leq 440$ and is zero otherwise. To gain the effect, we set $\delta_n = 10 \text{ s} \gg \tau = 0.1 \text{ s}$. It is supposed that white Gaussian noise $w_n = [0 w_{2n} 0 w_{4n}]^T$ is zero mean with the variances $\sigma_{w_1}^2 = \sigma_{w_2}^2 = \sigma_{w_4}^2$ and

\[
R = \sigma_n^2 \begin{bmatrix} \tau^2 / 3 & \tau / 2 & 0 & 0 \\ \tau / 2 & 1 & 0 & 0 \\ 0 & 0 & \tau^2 / 3 & \tau / 2 \\ 0 & 0 & \tau / 2 & 1 \end{bmatrix}. \tag{23}
\]

In such a model we have

\[
h_n(x_n) = \sqrt{\frac{x_{1n}^2 + x_{2n}^2}{(x_{1n}^2 + x_{2n}^2)}} \tag{24}
\]

and allow noise $v_n = [v_{1n} v_{2n}]$ to have the variance $\sigma_v^2 = \sigma_{v_1}^2 = \sigma_{v_2}^2$ and covariance

\[
Q = \begin{bmatrix} \sigma_v^2 & 0 & 0 & \sigma_v^2 \\ 0 & \sigma_v^2 & 0 & \sigma_v^2 \end{bmatrix}. \tag{25}
\]
Simulation has been conducted at 1000 points with step $\tau = 0.1$ s for $\sigma_w = 0.01$ m and $\sigma_v = 0.2$ m. The result shown in Fig. 4 confirms a statement made in [24] for linear filtering: EKF is much less robust against the temporary model uncertainties in the presence of errors in noise covariances.

4. CONCLUSIONS

In this paper, we have developed and studied an efficient iterative EFIR filtering algorithm. We have also shown several important advantages of this algorithm against the recursive EKF. It is only within a narrow range around the ideal conditions that EKF has better accuracy than EFIR. Otherwise, errors in EKF grow rapidly and may cause divergence, whereas the EFIR filter ignoring noise statistics remains at the same error level. Of practical importance is that the only tuning scalar value $N_{opt}$ required by the EFIR filter can easily be specialized via test measurements or even using regular measurements with no reference. Moreover, the determination of $N_{opt}$ requires much smaller efforts and cost than for the noise statistics, especially if the process is time-varying. A payment for this is an $N_{opt} = 1$ times longer operation required by the EFIR algorithm to complete iterations. This drawback can be circumvented using parallel computing.

REFERENCES