Ultimate iterative UFIR filtering algorithm

Yuriy S. Shmaliy [a,*], Sanowar Khan [b], Shunyi Zhao [c]

[a] Department of Electronics Engineering, Universidad de Guanajuato, 36885 Salamanca, Mexico
[b] Department of Electronics Engineering, City University London, London, UK
[c] Key Laboratory of Advanced Process Control for Light Industry, Jiangnan University, Wuxi, China

A B S T R A C T

Measurements are often provided in the presence of noise and uncertainties that require optimal filters to estimate processes with highest accuracy. The ultimate iterative unbiased finite impulse response (UFIR) filtering algorithm presented in this paper is more robust in real world than the Kalman filter. It completely ignores the noise statistics and initial values while demonstrating better accuracy under the mismodeling and temporary uncertainties and lower sensitivity to errors in the noise statistics.

Keywords:
Unbiased FIR filter
Kalman filter
Noisy measurements
State estimation

1. Introduction

Optimal estimation of system state is often required when measurements are provided in the presence of noise. If the process and its measurement are both linear and noise is white Gaussian, then the Kalman filter (KF) is recognized to be the best estimator. However, real life dictates that the conditions cannot always be satisfied for the KF. Therefore, it may produce extra errors, even unacceptable. In order to improve the KF performance big efforts were made over decades. Cox in [1] has derived the extended KF (EKF) for nonlinear models by linearizing the state-space equations. For highly nonlinear systems, Julier and Uhlmann employed the unscented transform [2] and proposed the unscented KF (UKF). Both EKF and UKF have then been used extensively. In [3], the EKF was developed to the invariant EKF for nonlinear systems possessing symmetries (or invariances). For high-dimensional systems, the ensemble KF was proposed by Evensen in [4] and, for systems with sparse matrices, the fast KF applied by Lange in [5]. The robust Kalman filter was designed by Masreliez [6,7] for linear state-space relations with non-Gaussian noise referred to as heavy tailed noise or Gaussian one mixed with outliers. Essential contributions to robust and highly robust estimation were also made in [8–10]. Most recently, Li et al. have developed in [11] an advanced robust unscented KF with an adaptive ability to changes in noise covariances.

Many other modifications have also attracted researches attention, just a few to mention. An efficient restoration Kalman-like algorithm was designed by Basseville et al. in [12] for hidden Markov trees. In [13], the KF was developed by Soken et al. for the conditions of small satellite attitude estimation with missed measurements. Ait-El-Fquih and Desbouvries applied in [14] the Kalman-like approach to triple Markov chains. Kalman filtering was also developed by Vaswani in [15] and Carmi et al. in [16] for sparse signals with unknown and time-varying sparsity patterns. An efficient Kalman-like tracking algorithm [17] was applied to the autoregressive channel process estimation with fading by Stefanatos and Katsaggelos. In [18], Becis-Aubry et al. discussed the two-step estimation algorithm with a switching gain matrix. In [19], Carli et al. employed the concept of the centralized KF for state estimation in the complex wireless sensor networks. Rawicz et al. reported in [20] explicit Kalman–Bucy-like solutions for two-state H_∞ and H_2 filters. In [21], Jwo and Cho have made several useful remarks regarding the linearized and extended Kalman filters, and we meet new useful solutions each year.

Notwithstanding the fact that the KF is definitely the best linear real-time estimator and in spite of a number of its essential improvements, the KF still may suffer of unacceptable errors when the model does not match the system exactly, the process implies uncertainties, noise is not white Gaussian, and the noise statistics are defined imprecisely. Jazwinski has resumed in [22] that poor performance of the KF in real world has a lot to do with its infinite impulse response [22] and that finite impulse response (FIR) filtering is more successful in accuracy under the real-world conditions. He also formulated the key motivation to develop FIR filters:

* Corresponding author.
E-mail address: shmaliy@ugto.mx (Y.S. Shmaliy).

A limited memory filter appears to be the only device for preventing divergence in the presence of unbounded perturbation in the system.

That means that the FIR filter is inherently more robust than the IIR (Kalman) filter as it does not project "old" errors to the estimate. In the subsequent decades, FIR filtering has been investigated profoundly to offer several important solutions. Kwon and Han have developed the theory of the bias-constrained receding horizon (one-step predictive) control [23] utilizing measurement data from the horizon $[k-N, k-1]$, where k is the current discrete time index and N is the horizon length. Diverse receding horizon FIR solutions were proposed in [24–28]. Shmaliy et al. have developed FIR filtering on the horizon $[k-N+1, k]$ that has resulted in several optimal, unbiased, and in-between solutions [29–33]. Most recently, receding horizon FIR filtering has been developed by Ahn et al. in [34–38] and Zhao et al. have found fast iterative forms for FIR filters in [39,40].

Among different kinds of FIR filters developed during decades, the unbiased FIR (UFIR) filter is most robust. This filter [32,28] appears as a solution to the unbiasedness constraint [24] or as a special case of the optimal FIR (OFIR) filter [30] when the model is noiseless. The UFIR filter has the embedded bounded input/bounded output (BIBO) stability and its algorithm completely ignores the noise statistics and initial conditions. The UFIR filter does not guarantee optimality, but the variance of its estimate diminishes as a reciprocal of the averaging horizon of N points. Accordingly, it becomes practically optimal when $N \gg 1$ and this is in agreement with the Gauss's ordinarily least squares (OLS).

Advantages of the iterative UFIR algorithm go along with the requirements of the optimal horizon of N_{opt} points which is applied to minimize the mean square error (MSE). Of practical importance is that a single tuning factor N_{opt} can be determined in a way much easier than for the noise statistics: using reference measurements or even via observations without a reference model [31]. The bad side is that iterations require about N_{opt} times more computation time and this price paid for the advantages may be an issue in real-time applications. But using parallel computing can make the UFIR algorithm as fast as the KF.

Although diverse forms of the UFIR filter have been discussed in many papers [32,33,41–43], its most demanded linear algorithm still has not been addressed to the engineer in a simple form. In this technical note, we introduce the ultimate iterative UFIR filtering algorithm suitable for immediate use.

2. Measured process and estimates

A big class of engineering problems can be solved if to represent the process and its measurement with the K-state space linear model as

$$\mathbf{x}_k = \mathbf{F}_k\mathbf{x}_{k-1} + \mathbf{w}_k, \tag{1}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \tag{2}$$

where $\mathbf{x}_k \in \mathbb{R}^K$ is the system state vector, $\mathbf{F}_k \in \mathbb{R}^{K\times K}$ is the state transition matrix, $\mathbf{y}_k \in \mathbb{R}^M$ is the observation vector, and $\mathbf{H}_k \in \mathbb{R}^{M\times K}$ is the measurement matrix. The process noise $\mathbf{w}_k \in \mathbb{R}^K \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ and the observation noise $\mathbf{v}_k \in \mathbb{R}^M \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ are both zero mean and white Gaussian with covariances, $\mathbf{Q}_k$ and $\mathbf{R}_k$. Vectors $\mathbf{w}_k$ and $\mathbf{v}_k$ and the initial state are uncorrelated and independent at each time step. For our purposes, we assign $\hat{\mathbf{x}}_{k|r}$ to be an estimate of $\mathbf{x}_k$ at time-index k via measurements from past up to and including at time-index r. We will also employ the following variables:

$\hat{\mathbf{x}}_k^- \triangleq \hat{\mathbf{x}}_{k|k-1}$, the *a priori* (prior or predicted) state estimate given observations up to and including at time $k-1$;

$\mathbf{P}_k^- \triangleq \mathbf{P}_{k|k-1} = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-)^T\}$, the *a priori* (prior or predicted) estimate covariance given observations up to and including at time $k-1$;

$\hat{\mathbf{x}}_k \triangleq \hat{\mathbf{x}}_{k|k}$, the *a posteriori* or posterior state estimate given observations up to and including at k;

$\mathbf{P}_k \triangleq \mathbf{P}_{k|k} = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T\}$, the *a posteriori* or posterior error covariance matrix given observations up to and including at k.

The KF applied to (1) and (2) produces estimates recursively in one step and is the best optimal engineering solution among other known real-time estimators. It is also simple in programming. But, in order for the KF estimate to be optimal, the initial $\mathbf{x}_0$ and $\mathbf{P}_0$ as well as the noise covariances, $\mathbf{Q}_k$ and $\mathbf{R}_k$, must be given that is typically not the case in practice. Experience dictates that the implementation of optimal KF is difficult due to the inability in getting good estimates of these values and the KF is thus suboptimal for all practical purposes.

2.1. Unbiased FIR filtering estimate

In contrast to the KF, the UFIR filter operates at once with N measurements on a horizon from $m = k-N+1$ to k and does not require neither the initial state $\mathbf{x}_m$ and error $\mathbf{P}_m$ nor the noise statistics $\mathbf{Q}_i$ and $\mathbf{R}_i$, $i \in [m, k]$. Instead it claims that N must be optimal, N_{opt}, in order to minimize the MSE and produce near optimal estimate. To run the iterations, the UFIR algorithm self determines the state $\hat{\mathbf{x}}_{m+K-1}$ at $m+K-1$ in a short batch form, where K is the number of the system states. It then updates estimates iteratively using recursions to reach the best value at k. The estimate $\hat{\mathbf{x}}_k$ obtained in such a way is called in [33] the optimal UFIR (OUFIR) estimate. Note that the estimation error covariance $\mathbf{P}_k$ is not involved to the algorithm that is an important advantage against the KF.

2.1.1. Batch UFIR estimate

The cost function for the UFIR filter is the unbiasedness condition

$$E\{\hat{\mathbf{x}}_k\} = E\{\mathbf{x}_k\}, \tag{3}$$

which means that the average of the state estimate is required to be equal to the average of the state. In the OLS format, the UFIR estimate can be written on a horizon $[m, k]$ as [30]

$$\hat{\mathbf{x}}_k = \left(\mathbf{C}_{m,k}^T\mathbf{C}_{m,k}\right)^{-1}\mathbf{C}_{m,k}^T\mathbf{Y}_{m,k}, \tag{4}$$

where the extended observation vector $\mathbf{Y}_{m,k}$ and mapping matrix $\mathbf{C}_{m,k}$ are represented as

$$\mathbf{Y}_{m,k} = \left[\mathbf{y}_m^T\mathbf{y}_{m+1}^T \ldots \mathbf{y}_k^T\right]^T, \tag{5}$$

$$\mathbf{C}_{m,k} = \begin{bmatrix} \mathbf{H}_m\left(\mathcal{F}_k^{m+1}\right)^{-1} \\ \mathbf{H}_{m+1}\left(\mathcal{F}_k^{m+2}\right)^{-1} \\ \vdots \\ \mathbf{H}_{k-1}\mathbf{F}_k^{-1} \\ \mathbf{H}_k \end{bmatrix}, \tag{6}$$

and the product of system matrices is depicted as

$$\mathcal{F}_r^g = \begin{cases} \mathbf{F}_r\mathbf{F}_{r-1}\ldots\mathbf{F}_g, & g \leqslant r, \\ I, & g = r+1 \\ 0, & \text{otherwise.} \end{cases}, \tag{7}$$

The standard convolution-based form for the batch UFIR filter [24] differs from (4) and is the following

$$\hat{\mathbf{x}}_k = \overline{\mathbf{K}}_{m,k}\mathbf{Y}_{m,k}, \tag{8}$$

where $\overline{\mathbf{K}}_{m,k}$ is the *UFIR filter gain* given by

$$\overline{\mathbf{K}}_{m,k} = (\mathbf{C}_{m,k}^T\mathbf{C}_{m,k})^{-1}\mathbf{C}_{m,k}^T \tag{9}$$
$$= \mathbf{G}_k\mathbf{C}_{m,k}^T. \tag{10}$$

Matrix $\mathbf{G}_k$ is known as the *generalized noise power gain* (GNPG) [44] computed as

$$\begin{aligned}\mathbf{G}_k &= \overline{\mathbf{K}}_{m,k}\overline{\mathbf{K}}_{m,k}^T \\ &= (\mathbf{C}_{m,k}^T\mathbf{C}_{m,k})^{-1}.\end{aligned} \tag{11}$$

The GNPG plays an important role in FIR filtering. It originates from the noise power gain (NPG) introduced by Trench in [45] as a ratio of the FIR filter output noise variance σ_{out}^2 to the input noise variance σ_{in}^2. As such, the NPG is akin to the noise figure in wireless communications. Trench has also shown that the NPG for white Gaussian noise is equal to the sum of the squared coefficients of the FIR filter impulse response function $h(k)$,

$$\text{NPG} = \frac{\sigma_{\text{out}}^2}{\sigma_{\text{in}}^2} = \sum_{n=0}^{N-1} h^2(k), \tag{12}$$

which is the squared norm of $h(n)$. In state space, $\overline{\mathbf{K}}_{m,k}$ represents coefficients of the filter impulse response. Therefore, $\overline{\mathbf{K}}_{m,k}\overline{\mathbf{K}}_{m,k}^T$ is called the GNPG.

2.1.2. Iterative UFIR estimate

The fast iterative form for (4) was found in [32]. Below, we give a more simple derivation to the iterative UFIR filtering algorithm.

Consider (4) and represent the inverse of the GNPG $\mathbf{G}_k = (\mathbf{C}_{m,k}^T\mathbf{C}_{m,k})^{-1}$ as

$$\begin{aligned}\mathbf{G}_k^{-1} &= \mathbf{C}_{m,k}^T\mathbf{C}_{m,k} \\ &= \sum_{i=0}^{N-1}(\mathcal{F}_k^{m+1+i})^{-T}\mathbf{H}_{m+i}^T\mathbf{H}_{m+i}(\mathcal{F}_k^{m+1+i})^{-1}\end{aligned} \tag{13}$$

$$\begin{aligned}&= \mathbf{H}_k^T\mathbf{H}_k + \sum_{i=0}^{N-2}(\mathcal{F}_k^{m+1+i})^{-T}\mathbf{H}_{m+i}^T\mathbf{H}_{m+i}(\mathcal{F}_k^{m+1+i})^{-1} \\ &= \mathbf{H}_k^T\mathbf{H}_k + \mathbf{F}_k^{-T}\left[\sum_{i=0}^{N-2}(\mathcal{F}_{k-1}^{m+1+i})^{-T}\mathbf{H}_{m+i}^T\mathbf{H}_{m+i}(\mathcal{F}_{k-1}^{m+1+i})^{-1}\right]\mathbf{F}_k^{-1}.\end{aligned} \tag{14}$$

From (7), we have $\mathcal{F}_r^g = 0$ if $g > r+1$ and one can find $\mathbf{G}_{k-1}^{-1}$ from (13) as

$$\mathbf{G}_{k-1}^{-1} = \sum_{i=0}^{N-2}(\mathcal{F}_{k-1}^{m+1+i})^{-T}\mathbf{H}_{m+i}^T\mathbf{H}_{m+i}(\mathcal{F}_{k-1}^{m+1+i})^{-1} \tag{15}$$

that is exactly the term in brackets of (14). By combining (14) and (15), we arrive at the recursion for the GNPG,

$$\mathbf{G}_k = [\mathbf{H}_k^T\mathbf{H}_k + (\mathbf{F}_k\mathbf{G}_{k-1}\mathbf{F}^T)^{-1}]^{-1}. \tag{16}$$

Similarly, represent the product $\mathbf{C}_{m,k}^T\mathbf{Y}_{m,k}$ as

$$\begin{aligned}\mathbf{C}_{m,k}^T\mathbf{Y}_{m,k} &= \mathbf{H}_k^T\mathbf{y}_k + \mathbf{F}_k^{-T}\sum_{i=0}^{N-2}(\mathcal{F}_{k-1}^{m+1+i})^{-T}\mathbf{H}_{m+i}\mathbf{y}_{m+i} \\ &= \mathbf{H}_k^T\mathbf{y}_k + \mathbf{F}_k^{-T}\mathbf{C}_{m,k-1}^T\mathbf{Y}_{m,k-1}.\end{aligned} \tag{17}$$

Combine (4) with (17) and write

$$\hat{\mathbf{x}}_k = \mathbf{G}_k(\mathbf{H}_k^T\mathbf{y}_k + \mathbf{F}_k^{-T}\mathbf{C}_{m,k-1}^T\mathbf{Y}_{m,k-1}). \tag{18}$$

From (4) taken at $k-1$ find $\mathbf{C}_{m,k-1}^T\mathbf{Y}_{m,k-1} = \mathbf{G}_{k-1}^{-1}\hat{\mathbf{x}}_{k-1}$ and go from (18) to

$$\hat{\mathbf{x}}_k = \mathbf{G}_k(\mathbf{H}_k^T\mathbf{y}_k + \mathbf{F}_k^{-T}\mathbf{G}_{k-1}^{-1}\hat{\mathbf{x}}_{k-1}). \tag{19}$$

By simple manipulations with (16), find $\mathbf{G}_{k-1}^{-1} = \mathbf{F}_k^T(\mathbf{G}_k^{-1} - \mathbf{H}_k^T\mathbf{H}_k)\mathbf{F}_k$, substitute into (19), provide the transformations, and arrive at the recursive Kalman-like form of [32],

$$\hat{x}_k = \mathbf{F}_k\hat{\mathbf{x}}_{k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{F}_k\hat{\mathbf{x}}_{k-1}), \tag{20}$$

in which $\mathbf{K}_k = \mathbf{G}_k\mathbf{H}_k^T$ is the bias correction gain that is not the Kalman gain and the GNPG $\mathbf{G}_k$ is computed recursively by (16).

As well as in the KF, each recursion in the UFIR filter implies two phases: "Predict" and "Update". Following the UFIR filter strategy, the estimate at time-index k is obtained iteratively with an auxiliary variable l beginning with $l = m+K$ and ending when $l = k$. To run iterations, the initial estimate at $l = m+K-1$ is provided using (4) in a shortest available batch form on a horizon $[m, m+K-1]$, because the inverse in (4) does not exist otherwise.

Since the UFIR algorithm does not require the noise statistics, the predict phase computes only one value – the *a priori* state estimate $\hat{\mathbf{x}}_l^- = \mathbf{F}_l\hat{\mathbf{x}}_{l-1}$. Combined with the current state observation to refine the state estimate, the *a priori* state estimate is iteratively updated to the *a posteriori* state estimate via the following values: the GNPG (16), the measurement residual $\mathbf{z}_l = \mathbf{y}_l - \mathbf{H}_l\hat{\mathbf{x}}_l^-$, the bias correction gain $\mathbf{K}_l = \mathbf{G}_l\mathbf{H}_l^T$, which is not the Kalman gain, and the *a posteriori* state estimate $\hat{\mathbf{x}}_l = \hat{\mathbf{x}}_l^- + \mathbf{K}_l\mathbf{z}_l$.

The iterative UFIR filtering algorithm is given in Table 1 with a pseudo code. The only tuning parameter required by this algorithm is the horizon length N which must be optimal, N_{opt}, in order to minimize MSE. There are at least two ways of how to find N_{opt}. The value of N_{opt} can be ascertained at the early stage by minimizing the trace of the estimation error covariance $\mathbf{P}_k$ as

$$N_{\text{opt}} = \arg\min_N\{\text{tr}\,\mathbf{P}(N)\}. \tag{21}$$

In this case, the system state $\mathbf{x}_k$ is measured using test equipment and supposed to be known. Otherwise, N_{opt} can be estimated via measurement $\mathbf{y}_k$ with no reference as shown in [31]:

$$N_{\text{opt}} \cong \arg\min_N\left\{\frac{\partial}{\partial k}\text{tr}\,\mathbf{L}_k(N)\right\}, \tag{22}$$

by minimizing the derivative of the trace of the mean square value

$$\mathbf{L}_k = E\{(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_k)(\mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_k)^T\}.$$

Beyond these approaches, the correlation method was employed in [46] to find N_{opt} and an advanced technique was developed in [47] to specify the minimum acceptable N. Note that, for nonlinear and time-variant systems, N_{opt} must be specified at each time index k.

Provided N, the initial values of the GNPG and estimate are computed at $k - N + K$ and Algorithm (Table 1) produces the first estimate at $N-1$. Recursions are organized based on (14)–(16) and the output is taken when the iterative variable reaches the current time index, $l = k$. One may thus deduce that a lack of information about noise, initial values, and error covariances is compensated in the UFIR algorithm by setting properly the GNPG which tunes the bias correction gain most closely to the Kalman gain, by N_{opt}. Note that the minimal horizon length for the initial batch estimate is equal to the number K of the states that makes the GNPG almost unity. Therefore, the initial GNPG in many cases can be substituted with an identity matrix without essential loss in accuracy.

Table 1
Iterative UFIR filtering algorithm.

Data: $\mathbf{y}_k$
Result: $\hat{x}_k$
begin
  for $k = N-1 : \infty$ **do**
    $m = k-N+1, \quad s = m+K-1;$
    $\mathbf{G}_s = (\mathbf{C}_{m,s}^T\mathbf{C}_{m,s})^{-1};$
    $\tilde{\mathbf{x}}_s = \mathbf{G}_s\mathbf{C}_{m,s}^T\mathbf{Y}_{m,s};$
    for $l = s+1 : k$ **do**
      $\mathbf{G}_l = [\mathbf{H}_l^T\mathbf{H}_l + (\mathbf{F}_l\mathbf{G}_{l-1}\mathbf{F}_l^T)^{-1}]^{-1}$
      $\tilde{\mathbf{x}}_l = \mathbf{F}_l\tilde{\mathbf{x}}_{l-1} + \mathbf{G}_l\mathbf{H}_l^T(\mathbf{y}_l - \mathbf{H}_l\mathbf{F}_l\tilde{\mathbf{x}}_{l-1})$
    end for
    $\hat{\mathbf{x}}_k = \tilde{\mathbf{x}}_k$
  end for
end

2.2. Estimation errors

Even though the UFIR filter does not require the error covariance $\mathbf{P}_k$ to update the estimate, the latter may be needed to evaluate estimation accuracy and precision. An iterative form still has not been proposed to compute $\mathbf{P}_k$ precisely in view of mathematical issues. But the error upper bound has a computationally simple form [33] as shown in Table 2.

Provided $\mathbf{P}_s$, recursions in this algorithm are organized with one value in each phase. The algorithm first predicts the estimation error over the system noise covariance $\mathbf{Q}_l$ and then updates it over the measurement noise covariance $\mathbf{R}_l$ and the bias correction gain $\mathbf{K}_l = \mathbf{G}_l\mathbf{H}_l^T$ taken from Table 1. The value produced by the algorithm (Table 2) is overestimated for the worst case and thus an actual error should be expected to be smaller.

2.3. Properties of UFIR filter

Looking for a good estimator we naturally desire to have one that is optimal, robust, universal for different signals in diverse environments, and (!!!) simple. As well as the KF, the UFIR one is a universal linear estimator, so let us look at its ability to fulfil other requirements and measure up to optimal Kalman estimate. Before providing a comparative analysis, it is a proper place to stress again that the optimal KF is challenged by the unbiased FIR filter that is potentially lesser accurate. But this is when the operation conditions are ideal. In real world any filter falls short of the requirements in some items and the question arises of how far. The KF has the recursive IIR structure in which the feedback plays a key role. On the contrary, the UFIR filter is a transversal estimator relying on the input-to-output relation with no feedback. Many properties of these estimators are due to their different structures.

Table 2
UFIR filter estimation error upper bound.

Data: $N,\ \mathbf{P}_s,\ \mathbf{Q}_k,\ \mathbf{R}_k,\ \mathbf{K}_k,$
Result: $\mathbf{P}_k$
begin
  for $k = N-1 : \infty$ **do**
    $s = k-N+K;$
    for $l = s+1 : k$ **do**
      $\widetilde{\mathbf{P}}_l^- = \mathbf{F}_l\widetilde{\mathbf{P}}_{l-1}\mathbf{F}_l^T + \mathbf{Q}_l$
      $\widetilde{\mathbf{P}}_l = (\mathbf{I} - \mathbf{K}_l\mathbf{H}_l)\widetilde{\mathbf{P}}_l^-(\mathbf{I} - \mathbf{K}_l\mathbf{H}_l)^T + \mathbf{K}_l\mathbf{R}_l\mathbf{K}_l^T$
    end for
    $\widehat{\mathbf{P}}_k = \widetilde{\mathbf{P}}_k$
  end for
end

Unbiasedness vs. optimality: Unbiasedness is the necessary condition of any good estimate. But the unbiased estimate may suffer of extra noise because the sufficient condition – minimal variance – is not applied. In turn, the optimal estimate guarantees the minimal MSE as a compromise between the bias and variance. Optimality in the KF and OFIR filter results in very consistent outputs. In turn, the UFIR filter produces a bit different and lesser precise estimate, although the difference between the unbiased and optimal estimates may practically be insignificant.

System model: The KF applies only to stochastic models. That is when both $\mathbf{Q}_k$ and $\mathbf{R}_k$ have zeroth components, the KF cannot operate. In contrast, the UFIR filter serves equally well for both the stochastic and deterministic systems. Both the KF and UFIR filter have extensions to nonlinear systems in the first and second order of approximations [1,48,49]. As was observed in Section 1, the KF has been extended to deal with many other practical issues. In this regard, the UFIR filter is still under the development.

Initial conditions: The UFIR filter does not require the initial conditions but produces the first estimate at $N-1$. On the contrary, the KF requires $\mathbf{x}_k$ and $\mathbf{P}_k$ at $k=0$ and updates these values at each next time step. Theory argues that there are no transients in the KF under the ideal conditions. However, transients in the KF may last much longer than in the UFIR filter in real world applications.

Noise environments: The key requirement that follows behind the derivation of the Bayesian KF estimator is that noise must be Gaussian and uncorrelated in both the state and observation equations. The Kalman filter is best under such conditions provided that the noise statistics are known exactly. Otherwise, it does not guarantee optimality. In contrast, the UFIR filter does not impose any restrictions on noise distribution and covariance and produces unbiased estimates if noise is just zero mean.

In order to produce optimal estimates, the KF requires the noise statistics at each time point. Because of this requirement and practical inability to obtain good estimates of $\mathbf{Q}_k$ and $\mathbf{P}_k$, the KF is suboptimal for all practical purposes. Fig. 1 gives an idea about the KF immunity to errors in the noise statistics in the worst case. Actual covariances are substituted here with $\mathbf{Q}_k/\alpha^2$ and $\alpha^2\mathbf{P}_k$, where α indicates an error in the noise standard deviation. As can be seen, errors in the KF grow rapidly with α and may become unacceptable, especially if a system is nonlinear [50].

In contrast, the UFIR filter ignores the noise statistics (except for the zero-mean assumption) but requires N_{opt} points in order to minimize the MSE. Note that N_{opt} can be measured in a way much easier than for the noise statistics [31]. Furthermore, the difference

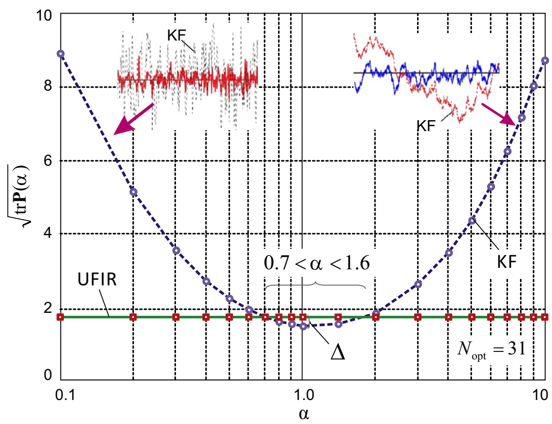

Fig. 1. Effect of errors in the noise covariances on KF estimates. The KF and UFIR estimates are poorly distinguishable when $\alpha = 1$ and Δ becomes insignificant with $N_{\text{opt}} \gg 1$.

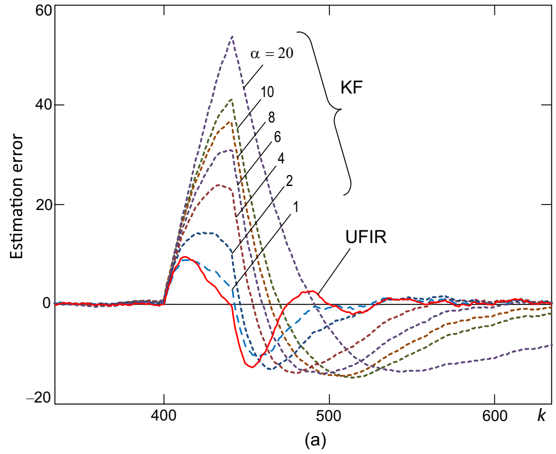

Fig. 2. Robustness of the UFIR filter against temporary model uncertainty in a gap of $400 \leqslant k \leqslant 440$ for a two-state polynomial model. As can be seen, the KF demonstrates poor robustness.

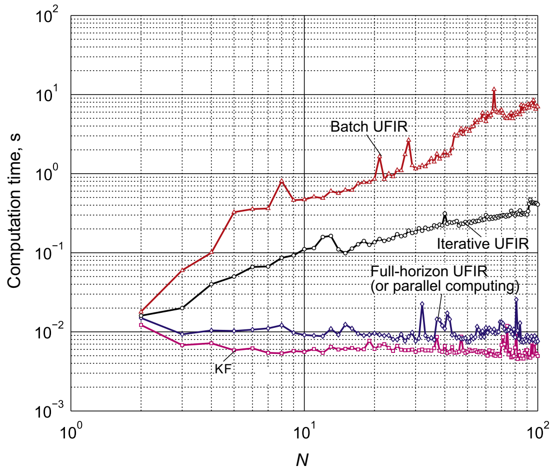

Fig. 3. Computation time measured as a function of N in different UFIR filtering algorithms and KF using the same computer and software.

between the optimal and unbiased estimates becomes negligible if $N_{\text{opt}} \gg 1$. This means that, invoking no information about noise, the UFIR filter is able to produce virtually optimal estimates.

Robustness: Two issues require robustness from an estimator. The model may not match a process accurately that causes mismodeling errors. Temporary uncertainties in the process and measurements may also lead to errors. Fig. 2 illustrates typical responses of the KF and UFIR filter to temporary uncertainties in the 2-state polynomial model. As can be seen, under the ideal conditions ($\alpha = 1$) both filters act quite similarly, except that the KF may have lasting transients owing to IIR. A situation changes dramatically if to admit errors in the noise statistics with $\alpha > 1$. In this case, the KF falls very short of the UFIR filter and its performance becomes particularly poor. An overall conclusion that can be made is that in real world the UFIR filter may demonstrate much better robustness than KF.

Stability: Both the KF and UFIR filter are stable filters but the transversal UFIR filter structure has the imbedded BIBO stability. For linear systems, neither the KF nor UFIR filter can pretend to be essentially advantageous in this plane under the ideal conditions. For nonlinear systems, extended versions of both these filters become unstable close to borders of high nonlinearities. It has also been observed in many nonlinear systems that the KF is more stable under the ideal conditions and it loses to the UFIR filter and may even diverge otherwise.

Computational complexity: Any digital estimator links its computation time directly to the computational complexity. From this standpoint, the KF has the lowest complexity $\mathcal{O}(1)$ and is the fastest one. The full-horizon UFIR algorithm [30,33] which iteratively updates the estimate over all time also has low complexity $\mathcal{O}(1)$ and thus operates as fast as the KF. The iterative UFIR algorithm (Table 1) has medium complexity $\mathcal{O}(N)$. It operates much faster than the batch UFIR algorithm but loses to the KF. Finally, the batch UFIR filter (7) having highest complexity $\mathcal{O}(N^2)$ is definitely not a real-time estimator when $N \gg 1$. Fig. 3 sketches the computation time measured as function of N in all these estimators using the same computer and software. The dependence on N is clearly seen here and we notice that this picture is typical.

Memory consumption: Memory required to complete operation in digital filtering often depends on the computational complexity. The batch UFIR filter which process simultaneously N measurements needs about N^2 more memory and the iterative one about N times more memory than the KF. If to design the iterative UFIR filer using parallel computing, then the algorithm will require about N^2 more memory than KF. The most "economic" UFIR filter is full-horizon which does not require much memory. But memory is no longer an issue in view of the tremendous progress in the computational resources.

3. Examples of applications

In spite of its engineering potential, UFIR filtering is still a relatively new technique. We thus give two practical examples which demonstrate advantages of the FIR approach against KF.

3.1. GPS-based UFIR filtering of clock state

Clocks are typically modeled using the two-state or three-state polynomial model [51]. The first state represents the time-interval error, the second state the fractional frequency offset, and the third state the linear frequency drift rate. To estimate the clock state, the time-interval error can be measured using a time-interval counter for the reference time provided by the Global Positioning System (GPS) timing receivers. The GPS time is accurate, but not precise in view of the GPS time uncertainties and sawtooth noise induced by the timing receivers. To filter the sawtooth noise out, an optimal filter can be used as shown in Fig. 4a. But the uniformly distributed measurement sawtooth noise is not Gaussian and the Kalman filter can thus produce extra errors. Furthermore, noise in the clock oscillator has flicker components with the power spectral density of the $1/f^n$ type that makes hardly possible to specify correctly the system noise covariance matrix for the Kalman filter. More details about GPS-based clock estimation and steering using UFIR filtering can be found in [52].

It was reported in [30,46] that the UFIR filter which ignores the noise statistics is an efficient alternative to KF. Fig. 4b gives typical errors produced by the UFIR filter and two-state KF, both applied to the GPS-based measurements of the time errors of a crystal clock imbedded in the Stanford Frequency Counter SR620. As can be seen, there is no time delay between the estimates and the filters thus have similar time constants. Herewith, the FIR filter demonstrates better robustness against mismodeling and the GPS time temporary uncertainty. It also produces smaller noise and lower regular errors.

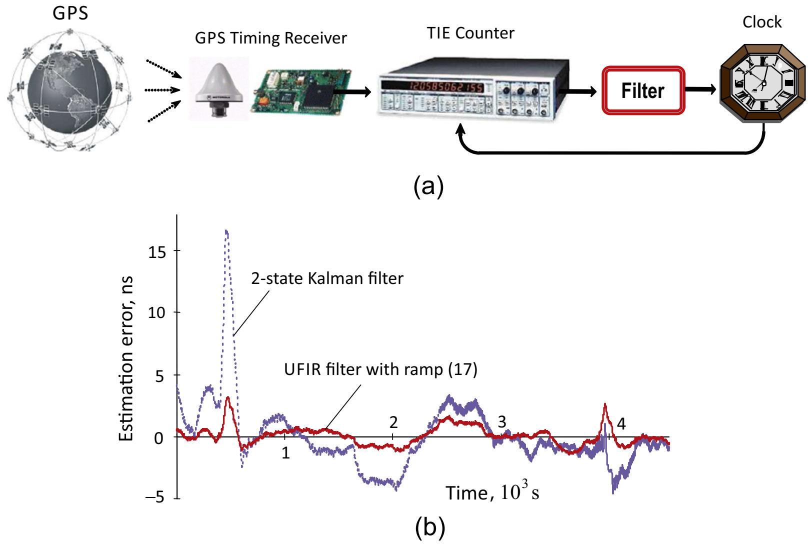

Fig. 4. GPS-based measurements and steering of local clock time errors using the UFIR filter, after [30]: (a) measurement set and (b) estimation errors.

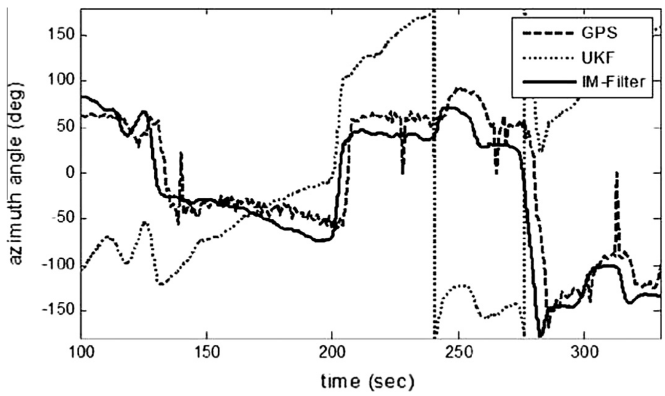

Fig. 5. Experimental estimation of heading in the INS/GPS system using the IM filter for the initially set zero heading, after [53].

3.2. Error reduction in navigation systems

To avoid large navigation errors and instability in the inertial navigation system (INS) integrated with GPS, an interacting multiple (IM) filter was proposed in [53] using a multifilter fusion technique in order to combine advantages of the FIR filter and other kinds of filters. Experimental estimation of heading using the unscented KF (UKF) and the IM filter has been provided in [53] for the GPS-based reference estimate. The unknown initial heading was set to zero. Fig. 5 shows estimation errors produced by the UKF and IM filter. As can be seen, the IM filter provides error reduction at acceptable levels, whereas the performance of the UKF is poor.

Several other applications of UFIR filtering in diverse electronic systems can be found in [54–59].

4. Conclusions

Unbiased FIR filtering introduced in this article is another opportunity to provide fast near optimal estimation beyond the KF. The rules of thumb are the following: (1) the Kalman filter is best under the ideal conditions and (2) if noise is nonwhite and/or non-stationary, the noise statistics are not known exactly, and/or the model undergoes temporary uncertainties and/or implies mismodeling, then the UFIR filter may produce smaller errors. That means that the FIR approach is more robust in real-world. Besides, a discrepancy between the KF and UFIR filter outputs reduces with N_{opt} and practically vanishes when $N_{\text{opt}} \gg 1$. Let us also notice that the UFIR approach can easily be applied to provide smoothing and prediction. As long as the UFIR filter ignores noise in the algorithm, a smoothed estimate with lag $q > 0$ is also the prior estimate $\tilde{\mathbf{x}}_{k-q} = (\mathbf{F}_k \ldots \mathbf{F}_{k-q+1})^{-1}\hat{\mathbf{x}}_k$. In turn, prediction with step $p > 0$ can be organized as $\tilde{\mathbf{x}}_{k+p} = \mathbf{F}_{k+p} \ldots \mathbf{F}_{k+1}\hat{\mathbf{x}}_k$.

Acknowledgment

This investigation was supported by the Royal Academy of Engineering under the Newton Research Collaboration Programme NRCP/1415/140.

References

[1] H. Cox, On the estimation of state variables and parameters for noisy dynamic systems, IEEE Trans. Autom. Contr. 9 (1) (1964) 5–12.
[2] S.J. Julier, J.K. Uhlmann, A new extension of the Kalman filter to nonlinear systems, Proc. SPIE 3068 (1997) 182–193.
[3] S. Bonnabel, P. Martin, E. Salaün, Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem, in: Proc. of 48th IEEE Conf. on Decision and Contr., 2009, pp. 1297–1304.

[4] G. Evensen, Sequential data assimilation with nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res. 99 (C5) (1994) 143–162.
[5] A.A. Lange, Simultaneous statistical calibration of the GPS signal delay measurements with related meteorological data, Phys. Chem. Earth Part A: Solid Earth Geodesy 26 (6–8) (2001) 471–473.
[6] C.J. Masreliez, Approximate non-Gaussian filtering with linear state and observation relations, IEEE Trans. Autom. Contr. AC-20 (1) (1975) 107–110.
[7] R.D. Martin, C.J. Masreliez, Robust estimation via stochastic approximation, IEEE Trans. Inform. Theory IT-21 (3) (1975) 263–271.
[8] R.A. Maronna, V.J. Yohai, Robust Statistics Theory and Methods, J. Wiley, New York, 2006.
[9] G. Chang, M. Liu, M-estimator-based robust Kalman filter for systems with process modeling errors and rank deficient measurement models, Nonlinear Dyn. 80 (2015) 1431–1449.
[10] V. Stojanovic, N. Nedic, Robust Kalman filtering for nonlinear multivariable stochastic systems in the presence of non-Gaussian noise, Int. J. Robust Nonlinear Contr. 26 (3) (2016) 445–460.
[11] W. Li, S. Sun, Y. Jia, J. Du, Robust unscented Kalman filter with adaptation of process and measurement noise covariances, Digital Signal Process. 48 (2016) 93–103.
[12] M. Basseville, A. Benveniste, K.C. Chou, S.A. Golden, R. Nikoukhah, A.S. Willsky, Modeling and estimation of multiresolution stochastic processes, IEEE Trans. Inf. Theory 38 (2) (1992) 766–784.
[13] H.E. Soken, C. Hajiyev, S. Sakai, Robust Kalman filtering for small satellite attitude estimation in the presence of measurement faults, Euro. J. Control 20 (2) (2014) 64–72.
[14] B. Ait-El-Fquih, F. Desbouvries, Kalman filtering in triplet Markov chains, IEEE Trans. Signal Process. 54 (8) (2006) 2957–2963.
[15] N. Vaswani, Kalman filtered compressed sensing, in: Proc. 15th IEEE Int. Conf. on Image Process., 2008, pp. 893–896.
[16] A. Carmi, P. Gurfil, D. Kanevsky, Methods for sparse signal recovery using Kalman filtering with embedded pseudo-measurement norms and quasi-norms, IEEE Trans. Signal Process. 58 (4) (2010) 2405–2409.
[17] S. Stefanatos, A.K. Katsaggelos, Joint data detection and channel tracking for OFDM systems with phase noise, IEEE Trans. Signal Process. 56 (9) (2008) 4230–4243.
[18] Y. Becis-Aubry, M. Boutayeb, M. Darouach, State estimation in the presence of bounded disturbances, Automatica 44 (7) (2008) 1867–1873.
[19] R. Carli, A. Chiuso, L. Schenato, S. Zampieri, Distributed Kalman filtering based on consensus strategies, IEEE J. Sel. Areas Commun. 26 (4) (2008) 622–633.
[20] P.L. Rawicz, P.R. Kalata, K.M. Murphy, T.A. Chmielewski, Explicit formulas for two state Kalman, H_2 and H_∞ target tracking filters, IEEE Trans. Aerospace Electron. Syst. 39 (1) (2003) 53–69.
[21] D.J. Jwo, T.S. Cho, Critical remarks on the linearised and extended Kalman filters with geodetic navigation examples, Measurement 43 (9) (2010) 1077–1089.
[22] A.H. Jazwinski, Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
[23] W.H. Kwon, S. Han, Receding Horizon Control: Model Predictive Control for State Models, Springer, London, 2005.
[24] W.H. Kwon, P.S. Kim, P. Park, A receding horizon Kalman FIR filter for discrete time-invariant systems, IEEE Trans. Autom. Contr. 44 (9) (1999) 1787–1791.
[25] W.H. Kwon, P.S. Kim, S.H. Han, A receding horizon unbiased FIR filter for discrete-time state space models, Automatica 38 (3) (2002) 545–551.
[26] S.H. Han, W.H. Kwon, P.S. Kim, Quasi-deadbeat minimax filters for deterministic state-space models, IEEE Trans. Automat. Control 47 (11) (2002) 1904–1908.
[27] C.K. Ahn, S. Han, W.H. Kwon, H_∞ FIR filters for linear continuous-time state-space systems, IEEE Signal Process. Lett. 13 (9) (2006) 557–560.
[28] P.S. Kim, M.E. Lee, A new FIR filter for state estimation and its applications, J. Comput. Sci. Technol. 22 (5) (2007) 779–784.
[29] Y.S. Shmaliy, L. Arceo-Miquel, Efficient predictive estimator for holdover in GPS-based clock synchronization, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 55 (10) (2008) 2131–2139.
[30] Y.S. Shmaliy, O. Ibarra-Manzano, Time-variant linear optimal finite impulse response estimator for discrete state-space models, Int. J. Adapt. Control Signal Process. 26 (2) (2012) 95–104.
[31] F. Ramirez-Echeverria, A. Sarr, Y.S. Shmaliy, Optimal memory for discrete-time FIR filters in state-space, IEEE Trans. Signal Process. 62 (2014) 557–561.
[32] Y.S. Shmaliy, An iterative Kalman-like algorithm ignoring noise and initial conditions, IEEE Trans. Signal Process. 59 (6) (2011) 2465–2473.
[33] Y.S. Shmaliy, D. Simon, Iterative unbiased FIR state estimation: a review of algorithms, EURASIP J. Adv. Signal Process. 113 (1) (2013) 1–16.
[34] C.K. Ahn, A new solution to the induced l_∞ finite impulse response filtering problem based on two matrix inequalities, Int. J. Control 87 (2) (2014) 404–409.
[35] C.K. Ahn, Strictly passive FIR filtering for state-space models with external disturbance, Int. J. Electron. Commun. 66 (11) (2012) 944–948.
[36] J.M. Pak, C.K. Ahn, Y.S. Shmaliy, P. Shi, M.T. Lim, Switching extensible FIR filter bank for adaptive horizon state estimation with applications, IEEE Trans. Control Syst. Technol. 24 (3) (2016) 1052–1058.
[37] I.H. Choi, J.M. Pak, C.K. Ahn, S.H. Lee, M.T. Lim, M.K. Song, Arbitration algorithm of FIR filter and optical flow based on ANFIS for visual object tracking, Measurement 75 (2015) 338–353.
[38] I.H. Choi, J.M. Pak, C.K. Ahn, Y.H. Mo, M.T. Lim, M.K. Song, New preceding vehicle tracking algorithm based on optimal unbiased finite memory filter, Measurement 73 (2015) 262–274.
[39] S. Zhao, Y.S. Shmaliy, Fast computation of discrete optimal FIR estimates in white Gaussian noise, IEEE Signal Process. Lett. 22 (6) (2015) 718–722.
[40] S. Zhao, Y.S. Shmaliy, F. Liu, Fast Kalman-like optimal unbiased FIR filtering with applications, IEEE Trans. Signal Process. 64 (9) (2016) 2284–2297.
[41] J. Pomarico-Franquiz, S. Khan, Y.S. Shmaliy, Combined extended FIR/Kalman filtering for indoor robot localization via triangulation, Measurement 50 (2014) 236–242.
[42] J. Contreras-Gonzalez, O. Ibarra-Manzano, Y.S. Shmaliy, Clock state estimation with the Kalman-like UFIR algorithm via TIE measurement, Measurement 46 (1) (2013) 476–483.
[43] J.M. Pak, C.K. Ahn, Y.S. Shmaliy, M.T. Lim, Improving reliability of particle filter-based localization in wireless sensor networks via hybrid particle/FIR filtering, IEEE Trans. Ind. Inform. 11 (5) (2015) 1089–1098.
[44] Y.S. Shmaliy, O. Ibarra-Manzano, Noise power gain for discrete-time FIR estimators, IEEE Signal Process. Lett. 18 (4) (2011) 207–210.
[45] W.F. Trench, A general class of discrete time-invariant filters, J. Soc. Ind. Appl. Math. 9 (3) (1961) 405–421.
[46] Y. Kou, Y. Jiao, D. Xu, M. Zhang, Ya. Liu, X. Li, Low-cost precise measurement of oscillator frequency instability based on GNSS carrier observation, Adv. Space Res. 51 (6) (2013) 969–977.
[47] J.M. Pak, C.K. Ahn, Y.S. Shmaliy, P. Shi, M.T. Lim, Switching extensible FIR filter bank for adaptive horizon state estimation with applications, IEEE Trans. Control Syst. Technol. 24 (3) (2016) 1052–1058.
[48] M. Athans, R.P. Wishner, A. Bertolini, Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements, IEEE Trans. Autom. Control AC-13 (1968) 504–514.
[49] Y.S. Shmaliy, Suboptimal FIR filtering of nonlinear models in additive white Gaussian noise, IEEE Trans. Signal Process. 60 (10) (2012) 5519–5527.
[50] F. Daum, Nonlinear filters: beyond the Kalman filter, IEEE A& E Syst. Mag. 20 (8) (2005) 57–69.
[51] IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology – Random Instabilities, IEEE Standard 1139-1999; IEEE, Piscataway, NJ, 1999.
[52] Y.S. Shmaliy, GPS-based Optimal FIR Filtering of Clock Models, Nova Science Publ., New York, 2009.
[53] S.Y. Cho, IM-filter for INS/GPS-integrated navigation system containing low-cost gyros, IEEE Trans. Aerospace Electron. Syst. 50 (4) (2014) 969–977.
[54] J.H. Lee, S. Hwang, D.H. Yu, C. Park, S.J. Lee, Software-based performance analysis of a pseudolite time synchronization method depending on the clock source, J. Position. Nav. Tim. 3 (4) (2014) 163–170.
[55] Y. Zhang, W. Lu, D. Lei, Y. Huang, D. Yu, Effective PPS signal generation with predictive synchronous loop for GPS, IEICE Trans. Commun. E97-B (8) (2014) 1742–1749.
[56] Y.S. Shmaliy, O. Ibarra-Manzano, L. Arceo-Miquel, J. Munoz-Diaz, A thinning algorithm for GPS-based unbiased FIR estimation of a clock TIE model, Measurement 41 (5) (2008) 538–550.
[57] Y. Chen, S. Ding, Z. Xie, Z. Qi, X. Liang, Design study for a quasisynchronous CDMA sensor data collection system: an LEO satellite uplink access technique based on GPS, Int. J. Distrib. Sensor Networks 2015 (2015) 1–15. ID 421745.
[58] R.Y. Ramlall, Method for Doppler-Aided GPS Carrier-Tracking Using p-Step Ramp Unbiased Finite Impulse Response Predictor, U.S. Patent 8 773 305, July 8, 2014.
[59] J. Fu, J. Sun, S. Lu, Y. Zhang, Maneuvering target tracking with modified unbiased FIR filter, J. Beijing Univ. Aeronaut. Astronaut. 41 (1) (2015) 77–82.