Review of Unbiased FIR Filters, Smoothers, and Predictors for Polynomial Signals

Yuriy S. Shmaliy¹, Yrjö Neuvo² and Sanowar Khan³

¹Electronics Engineering Department, DICIS, Universidad de Guanajuato, Salamanca, Mexico
²Department of Communications and Networking, Aalto University, Aalto, Finland
³Department of Electrical and Electronic Engineering, City, University of London, London EC1V 0HB, UK
Email: shmaliy@ugto.mx

Abstract: Extracting an estimate of a slowly varying signal corrupted by noise is a common task. Examples can be found in industrial, scientific and biomedical instrumentation. Depending on the nature of the application the signal estimate is allowed to be a delayed estimate of the original signal or, in the other extreme, no delay is tolerated. These cases are commonly referred to as filtering, prediction, and smoothing depending on the amount of advance or lag between the input data set and the output data set. In this review paper we provide a comprehensive set of design and analysis tools for designing unbiased FIR filters, predictors, and smoothers for slowly varying signals, i.e. signals that can be modeled by low order polynomials. Explicit expressions of parameters needed in practical implementations are given. Real life examples are provided including cases where the method is extended to signals that are piecewise slowly varying. A critical view on recursive implementations of the algorithms is provided.

Keywords: Polynomial signal, unbiased FIR filter, trend analysis, signal prediction, signal smoothing, digital filter design

1 Introduction

There is a wide class of signals, which are smooth enough to be represented on finite time intervals with low degree polynomials. Examples can be found in target and multiple-target tracking provided by radars [1], positioning obtained with the Global Positioning System (GPS) [2], time synchronization of digital communication networks and local locks [3], biomedical engineering [4], etc. For such cases, finite impulse response (FIR) filtering, providing weighted averaging on a horizon of \( N \) points, typically gives good results owing to some important properties. In contrast to the infinite-impulse-response (IIR) structures, including the Kalman filter [5] and Kalman-Bucy filter [6], the FIR ones that are inherently bounded input/bounded output (BIBO) stable [7], have better robustness against temporary model uncertainties and round-off errors [8,9] and lower sensitivity to errors in the noise statistics.

The roots of unbiased FIR (UFIR) filtering of polynomial signals are in the Wiener filter theory [10] rather than in the regression-based Gauss’s ordinary least squares (OLS) [11] or best linear unbiased estimator (BLUE) and minimum variance unbiased (MVU) estimator [12]. Soon after Wiener has published his seminal result [10] related to noncausal signals, Zadeh and Ragazzini extended it in 1950 to the causal case [13]. Then Johnson developed in 1956 a discrete-time analog of the Zadeh-Ragazzini solution [14] considering it on a finite horizon of \( N \) points. About the same time, several other important findings were published. Developing the works by Johnson [14] and Blum [15], Trench introduced in [16] a concept of the noise power gain (NPG) as a ratio of the FIR filter output noise variance \( \sigma^2_{\text{out}} \) to the input noise variance \( \sigma^2_{\text{in}} \). Note that the NPG is akin to the noise figure in wireless communications. He also showed that the NPG for white Gaussian noise is equal to the sum of the squared coefficients of the filter impulse response \( h_m(n) \),

\[
\text{NPG} = \frac{\sigma^2_{\text{out}}}{\sigma^2_{\text{in}}} = \sum_{n=0}^{N-1} h_m^2(n), \tag{1}
\]

where \( n \) is a discrete time index and \( h_m(n) \) is considered in this paper as an \( m \)-degree polynomial function.

By introducing the NPG, Trench offered a simple rule of how to evaluate the efficiency of a FIR filter
in terms of noise reduction. Recently, a more general form of (1) was developed in state space as the generalized NPG (GNPG) [17].

In his work [14] Johnson also showed that the coefficients for \( h_m(n) \) can be defined analytically via the unbiasedness constraint using the method of Lagrange multipliers. That promised quick use with seemingly limitless applications. But, for whatever reason, researches worked out to do anything other than to develop Johnson’s idea and put it on a long hiatus, except for some extensions [18–20]. That was until 1988 when Heinonen and Neuvo developed the Johnson approach and found in [21] coefficients for the low-degree 1-step predictive FIR filters [22] while designing nonlinear FIR-median hybrid filters. During the following decades, the approach has been extended to differentiators and investigated by many authors [23–34].

The next steps put forward the unbiased FIR (UFIR) filters. Motivated by the necessity of having real-time estimates when prediction is not required, Shmaliy derived in [35, 36] the low-degree FIR filters in state space and showed that such structures obeying the unbiasedness condition belong to a class of UFIR estimators. Soon after a more general \( x \)-shift UFIR estimator was proposed in [37] which belongs to a two-parameter \((N, x)\) family of UFIR structures and for which the Heinonen and Neuvo 1-step predictive filter [21] is a special case of \( x = 1 \) and the UFIR filter [35, 36] a special case of \( x = 0 \). During the last decade, UFIR estimators have been developed extensively for filtering [38–45], predictive filtering [46, 47], smoothing filtering [48, 49], prediction [50, 51], and smoothing [49, 52] and used in diverse areas of signal processing, control, image processing, and communications. It was also shown that the \( m \)-degree UFIR functions are orthogonal [53, 54] and can thus be used for signals approximation, interpolation, and extrapolation.

The Johnson work has finally established a paradigm for linear statistical signal processing with already recognized advantages against Kalman filtering [3, 8, 55]: better robustness against uncertainties and lower sensitivity to noise. Below, we survey the state of the art in design of the unbiased FIR filters of polynomial signals emphasizing on major properties, advantages, and weaknesses.

## 2 Polynomial Signal and UFIR Filtering

Let us assume that there is a signal \( s(n) \triangleq s(nT) \) existing at discrete time point \( n \) and that its values follow up with a time step \( T \). In many applications, such a signal is measured as \( y(n) = s(n) + v(n) \), where \( v(n) \) is zero mean additive noise. Among physical processes there are signals which are smooth enough to be expanded on a finite horizon of \( N \) neighbouring time points to the finite degree Taylor series as

\[
s(n) = \sum_{g=0}^{K-1} z_{g+1} + T^g n^g \frac{g!}{g!}, \tag{2}
\]

where \( z_{g+1}, g \in [0, K-1] \), is the \((g+1)\)th state of \( s(n) \) at the initial point and \( K \) is the number of the states.

One may wish to estimate \( s(n) \) unbiasedly at time point \( n \) in the presence of noise \( v(n) \) utilizing available measurements on a horizon \([b, n]\) of \( N \) neighbouring points, from \( b = n - N + 1 \) to \( n \). The horizon \([b, n]\) corresponds to filtering. But it can also be shifted with respect to \( n \) on \( p \) points to the left to provide \( p \)-step predictive filtering or on \( q \) points to the right to organize \( q \)-lag smoothing filtering. The \( x \)-shift horizon thus stands for filtering when \( x = 0 \), for predictive filtering when \( x = p, p > 0 \), and for smoothing filtering if \( x = -q, q > 0 \).

It has been shown in [17, 47, 49] that signal \( s(n) \) having \( K \) states can unbiasedly be filtered at \( n \) on \([b-x, n-x]\) with arbitrary constant time-shift \( x \) via \( y(n) \) by the convolution

\[
\hat{s}(n|x) = h_m(n, x) * y(n), \tag{3}
\]

\(^1\) Here, \( \hat{s}(n|r) \) means an estimate of \( s(n) \) at \( n \) via measurements taken from past to \( r \).
if the filter \( m = K - 1 \) degree impulse response function \( h_m(n) \) is a monic polynomial

\[
h_m(n, x) = \sum_{j=0}^{m} a_{jm}(x)n^j
\]

(4)

existing from \( x \) to \( N - 1 + x \) and having the coefficients

\[
a_{jm}(x) = (-1)^j \frac{M_{(j+1)1}^{(m)}(x)}{|H_m(x)|},
\]

(5)

where \( |H_m| \) is the determinant and \( M_{(j+1)1}^{(m)} \) is the minor of the Hankel matrix

\[
H_m = \begin{bmatrix}
c_0 & c_1 & \cdots & c_m \\
c_1 & c_2 & \cdots & c_{m+1} \\
\vdots & \vdots & \ddots & \vdots \\
c_m & c_{m+1} & \cdots & c_{2m}
\end{bmatrix}.
\]

(6)

The components \( c_v, v \in [0, 2m] \), in (6) are power series

\[
c_v(x) = \sum_{i=x}^{N-1+x} i^v = \frac{1}{v+1} [B_{v+1}(N + x) - B_{v+1}(x)],
\]

(7)

where \( B_v(z) \) is the Bernoulli polynomial.

The above equations allow us to calculate explicit expressions and numerical values for \( h_m(n, x) \). For example, the ramp impulse response of the 1-degree with \( x = 0 \) is represented on a horizon of \( N = 7 \) points, from \( n = 0 \) to \( n = 6 \), as \( h_1(n) = [0.464, 0.357, 0.25, 0.143, 0.036, -0.071, -0.179] \).

### 3 Problems Solved with UFIR Structures

It has to be noticed now that the only cost function which \( h_m(n, x) \) uniquely obeys is the unbiasedness condition - the necessary condition of "good" estimate. That means that the average of the estimate is guaranteed by \( h_m(n, x) \) to be equal to the average of a signal: \( E\{h(n|n-x)\} = E\{s(n)\} \). Because the sufficient condition of "good" estimate - minimal variance - is not obeyed, the UFIR structure does not guarantee optimality in the mean square error (MSE) sense. But if noise is white Gaussian, the UFIR estimate is also the OLS estimate, BLUE, and MVU estimate and thus near optimal.

Figure 1 illustrates filtering of a noisy linear signal \( s(n) \) with the uniform \( h_0(n) \) and ramp \( h_1(n) \) to show key features of the FIR approach. First, the FIR filter requires \( N \) initial points to produce the first estimate. Second, \( h_0(n) \) produces a simple average which is biased, whereas the estimate obtained with \( h_1(n) \) is unbiased. The conclusion which can be made also follows from the Kalman filter theory:

- **The order of the optimal (and unbiased) filter is the same as that of a system.**

Thus, signals represented with an \( m \)-degree polynomial are filtered with minimal errors if the filter impulse response is also of \( m \)-degree.

The UFIR function \( h_m(n, x) \) demonstrates many unique properties resulting in near optimal performance of UFIR estimators. Aimed at discussing these properties in what follows, we first make some notations. For simplicity, we assign \( h_m(n) = h_m(n, x = 0) \) for filtering. For predictors and smoothers, we use another UFIR function \( h_m(n, x) \), which is coupled with \( h_m(n, x) \) by \( h_m(n, x) = h_m(n + x, x) \). Note that \( x = 0 \) makes both UFIR functions equal: \( h_m(n) = h_m(n) \).

The following estimation problems can be solved via discrete convolution using the UFIR functions as sketched in Fig. 2 with time diagrams.

\(^2\) Note that unbiasedness is also achieved with \( m > K - 1 \), but with lower noise reduction.
Figure 1. FIR filtering of a noisy linear signal with the uniform response $h_0(n)$ and ramp impulse response $h_1(n)$.

Figure 2. Estimation problems solved with FIR structures.
FIR filtering (Fig. 2a) provides a real time estimate at \( n \) given measurements from \( n - N + 1 \) to \( n \) as

\[
\hat{s}(n|n) = \sum_{i=0}^{N-1} h_m(i)y(n-i) .
\]  

(8)

Predictive FIR filtering (Fig. 2b) provides a \( p \)-step predictive estimate at \( n \) given measurements from \( n - N + 1 - p \) to \( n - p \) by

\[
\hat{s}(n|n-p) = \sum_{i=p}^{N-1+p} h_m(i,p)y(n-i) .
\]  

(9)

Predictive FIR filtering with \( p = 1 \) was originally proposed for polynomial models in [21]. In control theory, it is known as receding horizon filtering. The one-step predictive filtering is provided by

\[
\hat{s}(n|n-1) = \sum_{i=1}^{N} h_m(i,1)y(n-i) .
\]  

(10)

Smoothing FIR filtering (Fig. 2c) provides a \( q \)-lag smoothing estimate at \( n \) given data from \( n - N + 1 + q \) to \( n + q \) by

\[
\hat{s}(n|n+q) = \sum_{i=-q}^{N-1-q} h_m(i,-q)y(n-i) .
\]  

(11)

FIR prediction (Fig. 2d) provides a \( p \)-step prediction at \( n + p \), \( p > 0 \), given measurements from \( n - N + 1 \) to \( n \),

\[
\hat{s}(n+p|n) = \sum_{i=0}^{N-1} \hat{h}_m(i,p)y(n-i) .
\]  

(12)

FIR smoothing (Fig. 2e) provides a \( q \)-lag smoothed estimate at \( n - q \), \( q > 0 \), given measurements from \( n - N + 1 \) to \( n \) by

\[
\hat{s}(n-q|n) = \sum_{i=0}^{N-1} \hat{h}_m(i,-q)y(n-i) .
\]  

(13)

All kinds of smoothing problems can be solved using this relation. For example, one can organize smoothing with a fixed interval (\( N \) constant), fixed lag (\( q \) constant), and fixed point (\( n - q \) constant).

4 UFIR Filters

Defined to be two-parametric with constant \( x \), function \( h_m(n,x) \) becomes universal for the design of filters, smoothing filters, and predictive filters. The coefficients for low-degree polynomial \( h_m(n,x) \) are given in Table 1 from which one can easily come up with analytic forms for a particular shift \( x \).

Besides, it has been shown in [53,54] that \( h_m(n) \) establishes a class of discrete orthogonal polynomials (DOP) which are coupled with the recurrence relation

\[
h_m(n) = \frac{2m^2(2N-1) - n(4m^2 - 1)}{m(2m-1)(N+m)} h_{m-1}(n) - \frac{(2m+1)(N-m)}{(2m-1)(N+m)} h_{m-2}(n) ,
\]  

(14)

where \( m \geq 1 \), \( h_{-1}(n) = 0 \), and \( h_0(n) = \frac{1}{N} \). Using (14), function \( h_m(n) \) of any degree can be extracted analytically or numerically via the low-degree ones. A family of low-degree orthogonal functions \( h_m(n) \) is sketched in Fig. 3.

It is worth to known that the NPG associated with \( h_m(n,x) \) is defined by the zero-order coefficient \( a_{0m}(x) \) and, for large \( N \), becomes

\[
\text{NPG}_{|N \gg 1} \sim \frac{(m + 1)^2}{N} .
\]  

(15)
Table 1. Coefficients for Low degree \( h_m(n, x) \)

<table>
<thead>
<tr>
<th>( h_m(n, x) )</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform: ( a_{00} = \frac{1}{N} )</td>
<td></td>
</tr>
<tr>
<td>Ramp: ( a_{01} = \frac{2(2N-1)(N-1)+12(N-1)x}{N(N^2-1)} ) ( a_{11} = \frac{6(N-1+2x)}{N(N^2-1)} )</td>
<td></td>
</tr>
<tr>
<td>Quadratic: ( a_{02} = \frac{3N^4-12N^3+17N^2-12N+4}{N(N^3-5N^2+4)} ) ( a_{12} = \frac{-12N^3-7N^2+7N-2+2(7N^2-15N+7)x+30(N-1)x^2+20x^3}{N(N^3-5N^2+4)} ) ( a_{22} = \frac{30(N^2-3N+2)(N-1)x+6x^2}{N(N^3-5N^2+4)} )</td>
<td></td>
</tr>
<tr>
<td>Cubic: ( a_{03} = 8 \left[ \frac{2N^5-15N^4+47N^3-90N^2+113N-57N+18+5(6N^5-42N^4+107N^3-132N^2+91N-30)x}{N(N^2-1)(N^2-4)(N^2-9)} + \frac{5(42N^4-213N^3+378N^2-288N+91)x^2+10(71N^3-246N^2+271N-96)x^3}{N(N^2-1)(N^2-4)(N^2-9)} \right] ) ( a_{13} = -20 \left[ \frac{6N^5-42N^4+107N^3-132N^2+91N-30+4(2N^5-213N^3+378N^2-288N+91)x}{N(N^2-1)(N^2-4)(N^2-9)} + \frac{5(246N^2-525N+271)x^2+4(246N^2-525N+271)x^3+1050(N-1)x^4+1620x^5}{N(N^2-1)(N^2-4)(N^2-9)} \right] ) ( a_{23} = 120 \left[ \frac{2N^4-11N^3+28N^2-23N+6+2(13N^3-48N^2+5N-23)x+2(48N^2-105N+58)x^2+140(N-1)x^3+70x^4}{N(N^2-1)(N^2-4)(N^2-9)} \right] ) ( a_{33} = -140 \frac{3N^3-6N^2+11N-6+2(6N^2-15N+11)x+30(N-1)x^2+20x^3}{N(N^2-1)(N^2-4)(N^2-9)} )</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3. A family of low-degree orthogonal polynomial functions \( h_m(n) \).
The following rule arises from (15):

- **An increase in the UFIR filter degree results in lower noise reduction.**

Below, we analyse in brief several special cases of the low-degree functions $h_m(n,x)$ and connect them to applications reported in the literature.

### 4.1 Uniform UFIR Function

The zero-degree uniform UFIR function $h_0(n,x)$ is constant

$$h_0(n,x) = a_{01}(x) = \frac{1}{N}$$  \hspace{1cm} (16)

as being associated with simple average. The UFIR filter with (16) suits best detecting constant signals and its $N_{PG} = \frac{1}{N}$ manifests the well known rules:

- **The variance of white Gaussian noise is reduced at the output of the zero-order (simple average) UFIR filter by the factor of $\frac{1}{N}$.**
- **Simple averaging is best in terms of noise reduction among other linear filters.**

No other UFIR filter of higher degree is more successful in noise reduction than the one with $h_0(n)$. 

### 4.2 Ramp UFIR Function

The first-degree ramp function $h_1(n,x) = a_{01}(x) + a_{11}(x)n$ suits for signals which change linearly on a horizon of $N$ points. This is the most widely used response due to its simplicity and high noise reduction. The coefficient for $h_1(n,x)$ can be found in Table 1 and the function has important special cases.

By $x = 0$, the ramp impulse response

$$h_1(n) = \frac{2(2N-1) - 6n}{N(N+1)}$$ \hspace{1cm} (17)

corresponds to filtering. Function (17) was originally derived in [35] using linear regression and then re-derived in [36] in state space. The UFIR filter with (17) suits well for slowly changing signals such as in the Global Positioning System (GPS)-based timekeeping [36] and precise navigation [44]. The $N_{PG}$ associated with (17) is $N_{PG} = \frac{2(2N-1)}{N(N+1)} \approx \frac{4}{N}$.

The 1-step predictive ramp response is obtained if we set $x = 1$,

$$h_1(n,1) = \frac{2(2N+1) - 6n}{N(N-1)}$$ \hspace{1cm} (18)

It was originally derived in [21] and is associated with receding horizon filtering [8]. Developments and applications of $h_1(n,1)$ can be found in many papers [23–28,46], as predicting the next coming sample is of importance in many applications.

An important special case of $h_1(n,-q)$ is associated with smoothing filtering with lag $q = N - 1$,

$$h_1(n,-N+1) = \frac{2N-1 + 3n}{N(N+1)}$$ \hspace{1cm} (19)

This response allows estimating the initial signal value as often required in systems modeling and optimal filtering.

Another special case referred to the Savitsky-Golay filtering strategy [56] is obtained with lag $q = (N - 1)/2$ at the horizon center. This lag makes the ramp function uniform,

$$h_1\left(n,\frac{-N+1}{2}\right) = \frac{1}{N}$$ \hspace{1cm} (20)

and noise reduction in the relevant smoothing filter thus becomes as good as in simple averaging (16).

An evolution of the ramp response $h_1(n,x)$ by changing $x$ around zero is sketched in Fig. 4 for $N = 21$. As can be seen, an increase in the smoother lag $q$ results in a smaller slope of $h_1(n,-q)$ that, by $q = (N - 1)/2$, makes $h_1(n,-q)$ uniform. Further increase in $q$ is a mirror reflection of the preceeding evolution. Just on the contrary, any increase in $p$ leads to larger $N_{PG}$ and bigger prediction errors.
The 1-step predictive filtering, by letting

\[ h_1(n, x) = \frac{3N^2 - 3N + 2 - 6(2N - 1)n + 10n^2}{N(N + 1)(N + 2)}. \] (21)

The 1-step predictive filtering, by \( x = 1 \), is due to [21]

\[ h_2(n, 1) = \frac{3N^2 + 3N + 2 - 6(2N + 1)n + 10n^2}{N(N - 1)(N - 2)}. \] (22)

Smoothing filtering of the initial value with lag \( q = N - 1 \) becomes available if to set \( x = -q \) and use

\[ h_2(n, -N + 1) = \frac{3N^2 - 3N + 2 + 6(2N + 1)n + 10n^2}{N(N + 1)(N + 2)} \] (23)

and smoothing with lag \( q = (N - 1)/2 \) at the horizon center is achieved with

\[ h_2\left(n, \frac{-N + 1}{2}\right) = \frac{3.3N^2 - 7 - 20n^2}{4N(N^2 - 4)}. \] (24)

An evolution of the quadratic response (22) by changing \( x \) is discussed in detail in [49].

### 4.3 Quadratic UFIR Function

Filtering of signals whose values change quadratically are better provided using the second-degree UFIR function \( h_2(n, x) = a_{02}(x) + a_{12}(x)n + a_{22}(x)n^2 \). The coefficients can be found in Table 1. This filter is more responsive to trend changes than the ramp one. But the side effect is in growing \( NPG_{N\geq1} \equiv \frac{a}{N} \) and lower noise reduction. As well as the ramp response, the quadratic one also offers several special cases of practical importance.

Filtering, by \( x = 0 \), is provided with the response [36]

\[ h_2(n) = \frac{3N^2 - 3N + 2 - 6(2N - 1)n + 10n^2}{N(N + 1)(N + 2)}. \] (21)

4.4 Cubic UFIR Function

The coefficients for the cubic response \( h_3(n, x) = a_{03}(x) + a_{13}(x)n + a_{23}(x)n^2 + a_{33}(x)n^3 \) of the 3-degree are also listed in Table 1. This response is even more “flexible” than the quadratic one \( h_2(n, x) \). Therefore, its \( NPG_{N\geq1} \equiv \frac{a}{N} \) often does not allow for acceptable noise reduction. This limits applications of the cubic response. Although \( h_3(n, x) \) suggests many special cases, their analytic forms are rather complex. Below we consider only two most interesting ones.

Filtering is obtained by setting \( x = 0 \) with [36]

\[ h_3(n) = \frac{8(2N^3 - 3N^2 + 7N - 3) - 20(6N^2

\[ -6N + 5)n + 120(2N - 1)n^2 - 140n^3}{N(N + 1)(N + 2)(N + 3)} \] (25)

and one-step predictive filtering is achieved by letting \( x = 1 \) using [21]

\[ h_3(n, 1) = \frac{4[2(2N^3 + 3N^2 + 7N + 3) - 5(6N^2

\[ +6N + 5)n + 30(2N + 1)n^2 - 35n^3]}{N(N - 1)(N - 2)(N - 3)}. \] (26)
4.5 Properties of \( h_m(n, x) \)

The UFIR filter impulse response \( h_m(n, x) \) demonstrates many properties which can be useful in the filter design. The most important ones are listed in Table 2 and discussed below.

<table>
<thead>
<tr>
<th>Property</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region of existence</td>
<td>( x \leq n \leq N - 1 + x )</td>
</tr>
<tr>
<td>( z )-transform at ( \omega = 0 ):</td>
<td>( H_m(z = 1, x) = 1 ) (UFIR filter is LP filter)</td>
</tr>
<tr>
<td>Unit area</td>
<td>( \sum_{n=x}^{N-1+x} h_m(n, x) = 1 )</td>
</tr>
<tr>
<td>Finite norm (NPG)</td>
<td>( \sum_{n=x}^{N-1+x} h_m^2(n, x) = a_0m(x) = h_m(0, x) )</td>
</tr>
<tr>
<td>Value at zero</td>
<td>( h_m(0, x) &gt; 0 )</td>
</tr>
<tr>
<td>Zero moments</td>
<td>( \sum_{n=x}^{N-1+x} h_m(n, x) n^u = 0 ), ( 1 \leq u \leq m )</td>
</tr>
<tr>
<td>Orthogonality (DOP)</td>
<td>( \sum_{n=x}^{N-1+x} h_k(n, x) h_m(n, x) n_u = 0 ), ( 1 \leq u \leq</td>
</tr>
<tr>
<td>Unbiasedness</td>
<td>( \int_0^{2\pi} H_m(e^{j\omega T}) d(\omega T) = \int_0^{2\pi}</td>
</tr>
<tr>
<td></td>
<td>( \sum_{k=0}^{N-1}</td>
</tr>
</tbody>
</table>

**Region of existence** Function \( h_m(n, x) \) exists from \( x \) to \( N - 1 + x \) and it is equal to zero otherwise. For filtering, \( x = 0 \), the response ranges from 0 to \( N - 1 \). For \( p \)-step predictive filtering \( (x = p) \), it is shifted on \( p \) points to the right and, for \( q \)-lag smoothing filtering \( (x = -q) \), on \( q \) points to the left. Each shift is accompanied with an evolution in the response shape that is illustrated in Fig. 4 for the ramp case and considered in detail in [49, 50] for low-degree filters.

**Any UFIR filter is an LP filter** \( z \)-transform applied to \( h_m(n, x) \) produces \( H_m(z) = 1 \) at zero frequency \( \omega = 0 \) and we thus conclude that:

> The UFIR filter is essentially a low-pass (LP) filter.

The frequency response of the UFIR filter, to be considered later, is rather complex and reminiscent of a sinc function.

**Unit area** The sum of values of \( h_m(n, x) \) is equal to one for all \( N \) and \( x \); that is, this function has a unit area, \( \sum_{n=x}^{N-1+x} h_m(n, x) = 1 \). In addition to the inherent BIBO stability, this property manifests that the UFIR filter is strictly stable.

**Finite norm (NPG)** The squared norm of \( h_m(n, x) \) is finite. It is equal to the free coefficient \( a_0m(x) \) and response value at \( n = x \). Therefore, the squared norm of \( h_m(n, x) \) is also equal to the filter NPG, \[
\sum_{n=x}^{N-1+x} h_m^2(n, x) = a_0m(x) = h_m(0, x) = \text{NPG} > 0.
\]
Figure 5. Noise power gain of low-degree UFIR filters around \( x = 0 \).

Figure 5 sketches the NPG for low-degree filters around \( x = 0 \) and one notices an easily seen difference between the smoothing filtering and predictive filtering. It also shows that the best noise reduction is achieved at the horizon center with \( q = (N - 1)/2 \) for odd \( m \) and that it is not the case for even \( m \) which suggest another lags for best denoising.

**Value at zero** It follows from the finite norm property that the value of \( h_m(n, x) \) at zero is always positive: \( h_m(0, x) > 0 \).

**Zero moments** Function \( h_m(n, x) \) has zero moments, \( \sum_{n=-\infty}^{N-1+x} h_m(n, x)n^u = 0 \) for \( 1 \leq u \leq m \). The moments of the products of these functions are also zeros, \( \sum_{n=x}^{N-1+x} h_k(n, x)h_m(n, x)n^u = 0 \) for \( 1 \leq u \leq |k - m| \). The latter may surprise, if one forgets the property of orthogonality, to be considered next.

**Orthogonality** A set of functions \( h_m(n) \) of degrees from zero to infinity establish a class of DOP that can be used for signal approximation in line with the classical Meixner, Hahn, Krawtchouk, and Charlier polynomials [53,54]. The benefit of this approximation resides in the fact that \( h_m(n) \) belongs to a one parameter family of DOP suitable for blind approximation with embedded unbiasedness, whereas the classical polynomials have more than one coefficient to optimize. Although a more general set of functions \( h_m(n, x) \) seemingly has this property as well, the orthogonality of \( h_m(n, x) \) still has not been investigated.

**Unbiasedness condition** Practical implementations of unbiased digital filters have a strong criterion in the frequency domain. A filter will track a signal unbiasedly if its \( z \)-transform and discrete Fourier transform (DFT) obey the following unbiasedness conditions [38,39], respectively,

\[
\int_0^{2\pi} H_m(e^{j\omega T})d(\omega T) = \int_0^{2\pi} \left| H_m(e^{j\omega T}) \right|^2d(\omega T),
\]

\[
\sum_{k=0}^{N-1} |H_m(k)|^2 = \sum_{k=0}^{N-1} H_m(k).
\]

Measures of unbiasedness turn up from (27) and (28) as absolute differences between the left and right sides.
5 UFIR Predictors and Smoothers

Although the $x$-shift UFIR estimator (4) solves the problems of filtering, smoothing filtering, and predictive filtering at time index $n$ in a universal scheme, many applications require prediction at $n+p$ and smoothing at $n−q$. The modified response $\tilde{h}_m(n,x)$ suitable for predictors and smoothers is coupled with $h_m(n,x)$ by the relationship

$$\tilde{h}_m(n,x) = h_m(n+x,x)$$

such that $\tilde{h}_m(n) = h_m(n)$ when $x = 0$. A key difference between these responses is that $x$ affects both the time position and shape of $h_m(n,x)$, whereas $\tilde{h}_m(n,x)$ is affected only on shape and thus exists from 0 to $N−1$ for all $x$. Based upon (29), function $h_m(n,x)$ of any degree can easily be converted to $\tilde{h}_m(n,x)$. Therefore, below we discuss specifics of $\tilde{h}_m(n,x)$ based only on the ramp function.

The ramp function $\tilde{h}_m(n,x)$ is coupled with $h_1(n)$ given by (17) with the relationship

$$\tilde{h}_1(n,x) = h_1(n) + \frac{6x(N−1−2n)}{N(N^2−1)}.$$  

(30)

Effect of $x$ on the NPG of (30) is stronger than in $h_1(n,x)$. However, the difference is in second order terms which for large $N$ can be neglected, as well as in $h_1(n,x)$. An evolution of $\tilde{h}_m(n,x)$ by changing $x$ around zero is shown in Fig. 6. What can be said comparing Fig. 4 and Fig. 6 is that $x$ has equal effect on their shapes. But, in line with this, $h_1(n,x)$ undergoes time shifts.

Several special cases of this response are discussed next. One-step ($p = 1$) ramp prediction from $n$ to $n+1$ becomes

$$\tilde{h}_1(n,1) = \frac{4(N−1−6n)}{N(N−1)}.$$  

(31)

Linear prediction often gives more accuracy than any nonlinear one on a long baseline [57]. Therefore, responses (30) and (18) play a special role. The ramp smoother with lag $q = N−1$,

$$\tilde{h}_1(n,N+1) = -\frac{2N−2−3n}{N(N+1)},$$

(32)

provides a more straightforward estimation of the initial value than (19). Smoothing at the horizon center is obtained with the response function

$$\tilde{h}_1\left(n,\frac{−N+1}{2}\right) = \tilde{h}_0(n) = \frac{1}{N},$$

(33)

which becomes uniform, as well as in the case of (20).

With a little specifics, most of the properties of $\tilde{h}_m(n,x)$ listed in Table 3 are the same as for $h_m(n,x)$. The crucial difference is that $\tilde{h}_m(n,x)$ exists from 0 to $N−1$ for all $x$, whereas $h_m(n,x)$ is $x$-shifted in time.
Table 3. Major Properties of $\tilde{h}_m(n, x)$

<table>
<thead>
<tr>
<th>Property</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region of existence:</td>
<td>$0 \leq n \leq N - 1$</td>
</tr>
<tr>
<td>$z$-transform at $\omega = 0$:</td>
<td>$H_m(z = 1, x) = 1$ (UFIR smoothers and predictors are LP structures)</td>
</tr>
<tr>
<td>Unit area:</td>
<td>$\sum_{n=0}^{N-1} \tilde{h}_m(n, x) = 1$</td>
</tr>
<tr>
<td>Finite norm (NPG)</td>
<td>$\sum_{n=0}^{N-1} \tilde{h}_m^2(n, x) = a_0m(x) = h_m(0, x)$</td>
</tr>
<tr>
<td>Initial value:</td>
<td>$\tilde{h}_m(0, x) \geq 0$</td>
</tr>
<tr>
<td>Zero moments:</td>
<td>$\sum_{n=0}^{N-1} \tilde{h}_m(n, x)(n + x)^u = 0$, $1 \leq u \leq m$</td>
</tr>
<tr>
<td>Orthogonality (DOP):</td>
<td>$\sum_{n=0}^{N-1} \tilde{h}_k(n, x)\tilde{h}_m(n, x)(n + x)^u = 0$, $1 \leq u \leq</td>
</tr>
<tr>
<td>Unbiasedness:</td>
<td>$\int_0^{2\pi} \tilde{H}_m(e^{j\omega T}, x)d(\omega T) = \int_0^{2\pi}</td>
</tr>
<tr>
<td></td>
<td>$\sum_{k=0}^{N-1}</td>
</tr>
</tbody>
</table>

6 Transfer Function and Frequency Domain Properties

Valuable insight to the characteristics of the UFIR filters is obtained by looking at their $z$-domain transfer functions and their frequency responses. The UFIR filter can be implemented in the continuous-frequency domain via the $z$-transform of its impulse response or in the discrete-frequency domain utilizing DFT. The transfer function has some important specifics that we consider next.

6.1 Transfer Function of UFIR Filter

The transfer function $H_m(z)$ of the UFIR filter was investigated for polynomial signals in detail in [38,45]. Applied to $h_m(n)$, the the $z$-transform produces

$$H_m(z) = \sum_{n=0}^{N-1} h_m(n)z^{-n}$$

$$= \sum_{i=0}^{m} \beta_i z^{-i} + z^{-N} \sum_{i=0}^{m} \gamma_i z^{-i}$$

$$1 + \sum_{i=1}^{m+1} \alpha_i z^{-i}, \quad (34)$$

where $z = e^{j\omega T}$ and the coefficients are given in Table IV for low-degree responses. If to assign

$$H_{\beta m}(z) = \left( \sum_{i=0}^{m} \beta_i z^{-i} \right) \left( 1 + \sum_{i=1}^{m+1} \alpha_i z^{-i} \right)^{-1}, \quad (35)$$

$$H_{\gamma m}(z) = \left( \sum_{i=0}^{m} \gamma_i z^{-i} \right) \left( 1 + \sum_{i=1}^{m+1} \alpha_i z^{-i} \right)^{-1}, \quad (36)$$

then a generalized block-diagram of the $m$-degree UFIR filter can be represented with a structure shown in Fig. 7. By setting the numerical values of Table IV to the denominator of (34) one easily notices that, for $m = 0$, the transfer function has one pole at $z = 1$ and that the number of poles at $z = 1$
Table 4. Coefficients for Transfer functions of Low-Degree UFIR Filters

<table>
<thead>
<tr>
<th>m</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>β₀</td>
<td>0</td>
<td>1</td>
<td>a₀₁(N)</td>
<td>a₀₂(N)</td>
</tr>
<tr>
<td>β₁</td>
<td>0</td>
<td>-4/N</td>
<td>18(N-1)</td>
<td>N(N+1)</td>
</tr>
<tr>
<td>β₂</td>
<td>0</td>
<td>0</td>
<td>9/N</td>
<td>24(N-3)</td>
</tr>
<tr>
<td>β₃</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-16/N</td>
</tr>
<tr>
<td>γ₀</td>
<td>-1/N</td>
<td>-2/N</td>
<td>-2/N</td>
<td>2/N</td>
</tr>
<tr>
<td>γ₁</td>
<td>0</td>
<td>2(N-2)</td>
<td>6(N-3)</td>
<td>N(N+1)</td>
</tr>
<tr>
<td>γ₂</td>
<td>0</td>
<td>0</td>
<td>2(N-2)(N-3)</td>
<td>N(N+1)(N+2)</td>
</tr>
<tr>
<td>γ₃</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-4(N-2)(N-3)(N-4)</td>
</tr>
<tr>
<td>α₁</td>
<td>-1</td>
<td>-2</td>
<td>-3</td>
<td>-4</td>
</tr>
<tr>
<td>α₂</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>α₃</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-4</td>
</tr>
<tr>
<td>α₄</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 7. A generalized block-diagram of the m-degree UFIR filter.

increases as m increases. This means that even though the UFIR filter transfer function itself gives a stable input-to-output relationship, the block diagram implementation shown in Fig. 7 contains internal unstable structures.

The magnitude and phase responses associated with (34) are shown in Fig. 8a and Fig. 8b for low-degree UFIR filters with N = 20. It can be seen from Fig. 8a that unbiasedness with increasing m is maintained by increasing the passband width and amplifying the first side lobe of the frequency response. The obvious drawback is that the stopband attenuation simultaneously gets worse. One can improve the stopband attenuation by introducing a prefilter that attenuates the stopband further and has a minimal effect on the unbiasedness [32]. The phase responses of the UFIR filter (Fig. 8b) and predictive UFIR filters [22, 23, 29] behave similarly. This forces the group delays of these filters to oscillate around a small constant value.

In addition to the fundamental properties of 2π-periodicity, symmetry of |Hₘ(z)|, and asymmetry of arg Hₘ(z), the following ones may be useful in the UFIR filter design.

**Transfer function at ω = 0:** By ω = 0, we have

\[ Hₘ(z = 1) = 1 \]  
(37)

for all m, and see again that the UFIR filter is a LP filter.

**Impulse response at n = 0:** The counterclockwise circular integration in the inverse z-transform of Hₘ(z) always produces a positive imaginary value. That means that we always have hₘ(0) > 0.

**Finite norm:** The Parceval theorem and finite norm property of hₘ(n) suggest that

\[ \frac{1}{2\pi} \int_{0}^{2\pi} |Hₘ(e^{j\omega T})|^2 d\omega T = hₘ(0) < \infty \]  
(38)
to mean that the squared norm of the impulse response in the transform domain is equal to $h_m(n)$ at $n = 0$.

The unbiasedness condition in the transform domain: Given a digital FIR filter with the $m$-degree transfer function $H_m(e^{j\omega T})$. This filter could be said to be unbiased if the following unbiasedness condition is obeyed:

$$\int_0^{2\pi} H_m(e^{j\omega T})d(\omega T) = \int_0^{2\pi} |H_m(e^{j\omega T})|^2d(\omega T).$$

(39)

We have already listed this property in Table 2 and Table 3.

Noise power gain: Because the NPG is defined by the squared norm of $h_{in}(N)$, the Parceval theorem and (39) suggest that $\text{NPG} = h_m(0) = a_{0m}$.

UFIR Filter with ramp response As an example, we now consider an implementation of the first-degree UFIR filter. The transfer function for this filter can be found from (38) to be

$$H_1(z) = \frac{2}{N} \frac{a_{11}N^2 - 2z^{-1} + z^{-N} \left(1 - \frac{N-2}{N+1}z^{-1}\right)}{(1 - z^{-1})^2}.$$  

(40)

A simple analysis reveals that the region of convergence in (40) is for all $z$ and that the filter is both stable and causal. Fig. 9 sketches a relevant block-diagram whose structure is $N$-invariant, although some blocks must be tuned to $N$. The diagram utilizes 6 multipliers and 4 adders. A further optimized form requiring 3 multipliers and 5 adders is shown in [23]. As already mentioned, the recursive implementations of UFIR algorithms are internally unstable implying that infinite arithmetic accuracy would be needed in the implementation. However, if the feedback loop is implemented using two’s complement arithmetic and the forward path multiplications are implemented exactly the internal overflows will not manifest
themselves to the output and exact pole-zero cancellation is achieved [23]. On the other hand, the performance of modern signal processing hardware suggests that the straightforward implementation of UFIR structures in the direct form is likely to be the preferred one in most cases. Design and implementation of UFIR filters with degrees exceeding one can be found in [22, 23, 29, 38, 39].

6.2 UFIR Filter Properties in the DFT Domain

If a system is designed to have digital signal processing (DSP) units, the UFIR filter can be implemented as a software in the discrete Fourier transform (DFT) domain. The standard convolution computation algorithm implies transferring $h_m(n)$ and $y(n)$ to the DFT domain, finding the product of the transforms, and returning back to the time domain using the inverse DFT. Fast convolution algorithms can also be designed as that [40] based on the computation by parts. Further reduction of the computation complexity can be achieved if to find short analytic forms for the DFT of $h_m(n)$ as shown in [39].

The DFT applied to $h_m(t)$ gives the discrete transfer functions

$$H_m(k) = \sum_{n=0}^{N-1} h_m(n)W_N^{nk}, \quad (41)$$

where $W_N = e^{-j2\pi/N}$ and a discrete-frequency index $k$ ranges as $0 \leq k \leq N - 1$. Unlike the $z$-transform of $h_m(n)$, the DFT is a discrete smooth function having the magnitude and phase sketched in Fig. 10. This figure neatly shows that the UFIR filter is an LP filter. It also suggests that the phase in the DFT domain is basically nonlinear.

In addition to the fundamental properties of $N$-periodicity, symmetry of $|H_m(k)|$, and asymmetry of $\arg H_m(k)$, the following important ones can be listed.

**Unbiasedness condition in the DFT domain:** Similarly to (43), the unbiasedness condition in the DFT domain becomes [39]

$$\sum_{k=0}^{N-1} |H_m(k)|^2 = \sum_{k=0}^{N-1} H_m(k). \quad (42)$$

**DFT value at $k = 0$:** By $k = 0$, we have $W_N^0 = 1$ and the DFT is thus $H_m(0) = 1$ for all $m$ manifesting again that the UFIR filter is an LP filter.

**Impulse response at $n = 0$:** By the inverse DFT applied to $H_m(k)$, the value of $h_m(n)$ at $n = 0$ becomes

$$h_m(0) = \frac{1}{N} \sum_{k=0}^{N-1} H_m(k). \quad (43)$$

Because of $h_m(0) > 0$ [17], the sum of the DFT coefficients is real and positive, $\sum_{k=0}^{N-1} H_m(k) = Nh_m(0) > 0$. 
DFT at $k = N/2$: By setting $k = N/2$ with $N$ even, one has $W^{m0}_{2N} = (-1)^n$ and the DFT turns out to be

$$H_1k(N) = -\sum_{m=0}^{l} \frac{a_{lm}(N)}{2} [E_m(N) - E_m(0)],$$

where $E_m(x)$ is the Euler polynomial.

Noise power gain: Because the NPG is specified by the squared norm of $h_m(n)$, relation (42) brings it to

$$\text{NPG} = \frac{1}{N} \sum_{k=0}^{N-1} |H_{1k}(N)|^2 = h_{00}(N) = a_{00}(N).$$

Estimate error bound: Employing (45), the estimation error bound can be defined in the three-sigma sense as

$$\beta = 3\sigma_v \sqrt{\text{NPG}},$$

where $\sigma_v$ is the standard deviation of the measurement white Gaussian noise $v_n$.

UFIR filter with ramp response The UFIR filter properties hold true for any degree $m$. However, for arbitrary $m$ the DFT (41) does not exist in closed form. Instead, solutions can be found for low degrees. Referring to [39] for a detailed analysis of (41), we consider below only the ramp response.

The DFT of $h_1(n)$ is $H_1k = 1$ at $k = 0$ and it has a simple form for $0 < k \leq N - 1$ of [39]

$$H_1(k) = \frac{6}{N + 1} \left[ 1 - \frac{j}{2\sin(\pi k/N)} W_{2N}^{1i} \right] = \frac{\alpha^2 \pi}{(N + 1) \sin(\pi k/N)^2} \left( \frac{\pi}{2} - 1 \right).$$

Figure 10. DFT of the low-degree impulse responses of the UFIR filter: (a) magnitude response for $N = 20$, (b) Bode plot of the magnitude response for $N = 200$, and (c) phase response for $N = 20$. 

DFT at $k = N/2$: By setting $k = N/2$ with $N$ even, one has $W^{m0}_{2N} = (-1)^n$ and the DFT turns out to be

$$H_1k(N) = -\sum_{m=0}^{l} \frac{a_{lm}(N)}{2} [E_m(N) - E_m(0)],$$

where $E_m(x)$ is the Euler polynomial.

Noise power gain: Because the NPG is specified by the squared norm of $h_m(n)$, relation (42) brings it to

$$\text{NPG} = \frac{1}{N} \sum_{k=0}^{N-1} |H_{1k}(N)|^2 = h_{00}(N) = a_{00}(N).$$

Estimate error bound: Employing (45), the estimation error bound can be defined in the three-sigma sense as

$$\beta = 3\sigma_v \sqrt{\text{NPG}},$$

where $\sigma_v$ is the standard deviation of the measurement white Gaussian noise $v_n$.

UFIR filter with ramp response The UFIR filter properties hold true for any degree $m$. However, for arbitrary $m$ the DFT (41) does not exist in closed form. Instead, solutions can be found for low degrees. Referring to [39] for a detailed analysis of (41), we consider below only the ramp response.

The DFT of $h_1(n)$ is $H_1(k) = 1$ at $k = 0$ and it has a simple form for $0 < k \leq N - 1$ of [39]

$$H_1(k) = \frac{6}{N + 1} \left[ 1 - \frac{j}{2\sin(\pi k/N)} W_{2N}^{1i} \right] = \frac{\alpha^2 \pi}{(N + 1) \sin(\pi k/N)^2} \left( \frac{\pi}{2} - 1 \right).$$
The magnitude response $|H_1(k)|$ of (47) decreases monotonously (see Fig. 10a and Fig. 10b) from unity at $k = 0$ to $3/(N+1)$ at $k = N/2$ and tends toward zero at $k = N/2$ by $N \gg 1$. Note that $N$ must be even for DFT to exist at $k = N/2$. Fig. 10a illustrates $|H_1(k)|$ for $N = 20$. One can observe that the Bode plot of $|H_1(k)|^2$ shown in Fig. 10b has a fundamental slope of $k^{-2}$ that evolves to $k^0$ at $k = N/2$.

6.3 Filters with Embedded Unbiasedness

The benefit of the unbiasedness condition (42) is in its ability to form a measure of biasedness. In the median form, this measure is

$$\kappa = \frac{1}{N} \sum_{k=0}^{N-1} |H_k| - H_k|.$$

(48)

Based upon (48), filters of different designs and structures can be tested for unbiasedness.

The most common filters fulfilling $\kappa = 0$ are listed in Table 5. An ideal LP filter has the rectangular magnitude response and zero phase. Its DFT is

$$H_k = \begin{cases} A, & |k| \leq B \\ 0, & |k| > B \end{cases},$$

(49)

where $A = \text{const}$ and $B$ is associated with the filter bandwidth. Testing (49) with $A = 1$ by (48) gives us $\kappa = 0$ and we infer that an ideal LP filter is unbiased for all signals whose spectral contents fall within the filter bandwidth. But, the filter becomes biased if $A \neq 1$ [39]. Although an ideal LP filter is the most universal unbiased one, it has IIR and is thus noncausal. Any truncating leads to the Gibbs’ phenomenon [40] and the filter loses an ability of unbiased filtering, because its magnitude is no longer uniform and unit within the bandwidth. An important conclusion follows:

▶ There is no universal causal LP filter providing unbiased filtering of an arbitrary signal.

An antipode of the ideal LP filter has the uniform impulse response associated with simple averaging,

$$h_0(n) = \begin{cases} \frac{1}{N}, & 0 \leq n \leq N-1 \\ 0, & \text{otherwise} \end{cases}.$$

(50)

The DFT of (50) is

$$H_k = \begin{cases} 0, & 0 < k < N \\ 1, & k = 0, k = N \end{cases}$$

(51)

and this filter is also unbiased, because (51) produces $\kappa = 0$ for all $N \geq 2$. However, unlike (49) with $A = 1$ suitable for all signals, (51) fits only constant signals and its use is thus strongly limited. Furthermore, any magnification in (51) leads to bias. Other examples of filters listed in Table 5 [39], including the cases of $m < \infty$ and $m = \infty$, confirm what has been stated before:

▶ A signal of degree $m$ can unbiasedly be filtered only with a filter of the same or larger degree.

And we notice again that this finding is a cornerstone of the optimal Kalman filter theory.
7 State-Space Algorithm

The convolution-based form (3) often fails to operate in real-time. In line with the z-transform-based implementation and fast convolution computation using DFT, fast state-space Kalman-like algorithms can also be used. Unlike the polynomial model (2) projecting a signal from the horizon initial point \( b = n - N + 1 \) to the horizon last point \( n \), the state-space model makes it from \( n - 1 \) to \( n \). Accordingly, signal \( s(n) \) and its measurement \( y(n) \) are represented with the state and measurement equations, respectively, as

\[
\begin{align*}
x_n &= Fx_{n-1}, \\
y_n &= Hx_n + v_n,
\end{align*}
\]

where \( x_n = [z_1(n) \ z_2(n) \ldots z_K(n)]^T \) is the signal state vector having \( K \) components. The measurement matrix is \( H = \left[ 1 \ 0 \ldots 0 \right] \) and the state transition matrix is given by

\[
F = \begin{bmatrix}
1 & T & T^2 & \ldots & T^{K-1} \\
0 & 1 & T & \ldots & \frac{T^{K-1}}{K-1} \\
0 & 0 & 1 & \ldots & \frac{T^{K-2}}{K-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1
\end{bmatrix}.
\]

For this model, the batch UFIR filtering estimate \( \hat{x}_n \) of \( x_n \) is obtained by

\[
\hat{x}_n = (A_N^T A_N)^{-1} A_N^T y_n,
\]

where \( y_n = [y_b \ y_{b+1} \ldots y_n]^T \),

\[
A_N = \left[ H^T (HF)^T \ldots (HF^{N-1})^T \right]^T.
\]

By (55), the estimate \( \hat{s}(n) \) of \( s(n) \) is taken as the first component of \( \hat{x}_k \),

\[
\hat{s}(n) = \hat{z}_1(n).
\]

The UFIR filter (55) estimates the signal states similarly to the classical Gauss's OLS which find coefficients for the signal polynomial model [12] via a set of measurements. For Gaussian noise, the form (55) is the same as for the OLS and BLUE and the estimate provided by (55) is equal to that by the MVU estimator. That means that:

▶ The batch UFIR filter (55) belongs to a class of OLS.

Example 1: A linear signal is filtered with the first-degree UFIR filter having a ramp response (17). In state space, such a signal has two states and can thus be filtered using (55) if to set

\[
F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \end{bmatrix}.
\]

Practice dictates that (55) is inappropriate for real-time filtering due to the computational burden coupled with high dimensions of \( A_N \) and \( y_n \) and large computation time when \( N \gg 1 \). Fast computation of (55) can be provided using the iterative algorithm [37] whose pseudo code is given in Algorithm 1.

This Algorithm 1 is akin to the Kalman filter, since both utilize recursions. However, in contrast to the latter, Algorithm 1 does not require the noise statistics and therefore has better engineering features. The first estimate produced by this algorithm appears at \( n = N - 1 \) that is the last point of the initial horizon

\[\text{In [36], the components in extended vector and matrices have inverse numeration that slightly modifies the estimate.}\]
Algorithm 1: Fast Iterative UFIR Filtering Algorithm

Data: $N$
Result: $\hat{x}_n$

begin
for $n = N - 1, N, \cdots$ do

$s = n - N + K$;
$G_s = (A_k^T A_k)^{-1}$;
$\tilde{x}_s = G_s A_k y_s$;
for $l = s + 1 : n$ do

$G_l = [H_l^T H_l + (FG_l F_l^T)^{-1}]^{-1}$;
$K_l = G_l H_l^T$;
$\tilde{x}_l = F \tilde{x}_{l-1} + K_l (y_l - HF \tilde{x}_{l-1})$;
end for

$\hat{x}_n = \tilde{x}_n$;
end for
end

[0, $N - 1$]. Each subsequent time point $n$ moves the horizon ahead and the estimate at $n$ is provided over measurements taken from $[n - N + 1, n]$. Recursions are repeated here from $n - N + K + 1$ to $n$. The GNPG $G_n$ and estimate are defined at $n - N + K$ using short batch forms. In this manner, an estimate is iteratively updated until it reaches the best value taken at $l = n$. Due to the iterations, Algorithm 1 operates about $N$ times slower than the Kalman filter. However, if the noise statistics are defined imprecisely as it usually is in practice, Algorithm 1 may produce more accuracy than in the Kalman filter.

8 Applications of UFIR Estimators

The unbiasedness of the UFIR algorithms and their optimality properties make them attractive for analyzing slowly varying signals encountered in various real life instrumentation and measurement applications. The closed form expressions for the multiplier values allow easy and fast implementation without the need to use a parameter optimization program. In this section we present as examples a few real life applications where the advantages of the UFIR algorithms can be clearly noticed.

8.1 Stepwise Signal Filtering Using FIR-median Hybrid Structures

The FIR-median hybrid structures broaden the application domains of the UFIR estimators to cases that cannot be properly tackled with linear signal processing algorithms. Good examples are noise corrupted staircase and sawtooth type signals whose abrupt level or slope changes cannot be accurately recovered with linear algorithms.

The FIR-median filter shown in Fig. 11 comprises of two linear predictive UFIR subfilters that produce the forward $\hat{s}_{FW}(n)$ and backward $\hat{s}_{BW}(n)$ estimates. The median is then found between $\hat{s}_{FW}(n)$, $\hat{s}_{BW}(n)$, and the measurement $y(n)$ to produce the estimate $\hat{s}(n)$. The FMH filter can be designed using different kinds of UFIR substructures whose number can be large. It is interesting to notice that in the case of forward and backward range predictors this construction allows signals composed of ramps of varying slopes. An example of denoising of the systolic blood pressure signal using the FMH structure with the 1-step predictive UFIR filters is given in Fig. 12, after [21]. Even a quick look at this figure shows a high efficiency of the FMH approach.

It is easy to notice that the filter structure (Fig. 11) has some unconventional properties. In the presence of no noise both staircase and sawtooth shaped signals pass the filter unchanged assuming the UFIR subfilters are shorter than the levels and linear slopes in the signal.

In the median filter literature signals that are not changed by filtering operation are called root signals. It is also known that repetitive filtering of a signal by a median filter eventually generates root signals. The FMH filter can be designed for many different types of applications by modifying the UFIR structures.
Figure 11. General structure of the FMH filter with 1-step predictive forward and backward UFIR substructures [50].

Figure 12. FMH filtering of a systolic blood pressure signal [21]: (a) original measured signal and (b) filtered signal.

Detailed analysis of the noise attenuation capabilities of FIR-median structures is difficult. The UFIR noise attenuation was discussed earlier and the properties of the median operation are also known. The median operation gives the value that minimizes the sum of absolute differences of the samples in the operation window and it is optimal for reducing Laplacian noise in a similar fashion as averaging is optimal for Gaussian noise. For Gaussian noise median is only about 2 dB worse than the average [58].

As an example of the denoising capabilities of the FMH structure we use the systolic blood pressure signal measured from a test person on an exercise bicycle where the load level is increased stepwise. The measurement is taken invasively and, as can be seen, the signal has a strong noise component. The signal was filtered repeatedly with an FMH structure up to the stage where the output was very close to a root signal. The blood pressures corresponding to the different load levels are clearly visible. In another similar blood pressure measurement it was noticed that the blood pressure was 10 mmHg higher during a lecture in front of a class than immediately before or after the class.

Although the FMH structure was originally proposed in [21] to utilize the 1-step predictive UFIR filters, it can also be designed using the UFIR filters and smoothing filters. Images tend to have sharp edges and slowly varying levels of intensity suggesting that FMH structures can be used in image enhancement. A promising example of applications can be found in [59].

8.2 Filtering of Clock Errors Using UFIR Filters

In digital communication networks, imbedded clocks are responsible for keeping time. If the frequency of a local clock differs from that of a master oscillator, then the time interval error (TIE) will grow with time and frequency steering will be required. With this aim, measurements of the TIE are provided.

Three main clock states are recognized [60,61]: TIE (first state $x_{1n}$), fractional frequency offset (second state $x_{2n}$), and linear fractional frequency drift rate (third state $x_{3n}$). Accordingly, clocks are commonly represented in state space either with the three-state polynomial model and state vector $x_n = [x_{1n} \ x_{2n} \ x_{3n}]$ or two-state model and $x_n = [x_{1n} \ x_{3n}]$. Measurements $y_n$ of the TIE $x_{1n}$ are often obtained using the time-interval counter for the Global Positioning System (GPS) time. The GPS time is accurate, but not precise in view of the sawtooth noise induced by the commercially available receivers.
Moreover, the clock oscillator noise has flicker (not white) components with the power spectral density of $1/f$ type. In such an environment, the Kalman filter which requires the noise components to be white Gaussian and uncorrelated may produce extra errors.

To provide more robust estimation of the clock state, the UFIR filter is used [3] as shown in Fig. 13. Here, the GPS one pulse per second (1PPS) timing signal from the SynPaQ/III GPS Sensor is used as a commercially available time. The time difference between the oven controlled crystal oscillator (OCXO)-based local clock and the GPS 1PPS is measured as $y_n$ by the Stanford Frequency Counter SR620. The local clock TIE measured in such a way goes to the UFIR filter which provides estimates $\hat{x}_n$ of the clock state. Simultaneously, the actual TIE $x_{1n}$ is measured for the reference Cesium Frequency Standard CsIII (Symmetricom, Inc., San Jose, CA).

Figure 14 gives typical errors produced by the two-state UFIR and Kalman filters [3]. 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 the GPS time temporary uncertainties. It also produces smaller noise and lower regular errors.

### 8.3 Filtering of Clock State Using UFIR Ladder Algorithm

Signal estimation is usually provided for a constant sampling time $T$ applied to all states. Because the higher order states in the polynomial model typically change more slowly, they may require larger
sampling intervals. In such cases, thinning of digital flaws can be made and the ladder UFIR algorithm designed in [62] used. A block diagram of the ladder algorithm shown in Fig. 15 for the three state model has been used with different purposes in [45, 63]. Here the first state $x_{1n}$ is measured in a regular way as $y_{1n} = y_n$. Measurements for the second and third states are formed by the time derivatives of the estimates of the first and second states, respectively.

$$y_{1n} = y_n.$$ Measurements for the second and third states are provided by the discrete time derivatives of the estimates of the first and second states, respectively, with individual sampling intervals $T_i$.

In [45], the ladder algorithm was used in the experimental setup for oscillator frequency error measurement and comparison. The frequency instability measured using a phase comparator and software Compass receiver utilizing the UFIR filter is illustrated in Fig. 16. As mentioned in [45], the high frequency components are efficiently mitigated by the UFIR filter. It is also seen that the UFIR filter-based software and cost phase comparator produce almost identical estimates of the oscillator frequency errors in terms of the Allan deviation (ADEV).

**8.4 Clock Error Steering Using Predictive UFIR Structures**

Once synchronized, two independent clocks will walk away from one another without limit. In order to keep the clock TIE at an acceptable level, it would be necessary to synchronize them again at regular
intervals. The performance between synchronizations depends on the quality of the clock oscillator and
errors in the clock state estimator and correction mechanism [64].

A GPS-based clock synchronization structure utilizing the one-step predictive UFIR filter (18) with
the ramp response $h_1(n, 1)$ has been designed in [46] as shown in Fig. 17. Typical steering errors achieved
by implementing this structure for the OCXO-based clock are sketched in Fig. 18. The results are given in
terms of the time deviation (TDEV) which is the standard measure in timekeeping [60]. As can be seen,

![Figure 17](image-url)  
**Figure 17.** Block-diagram of the GPS-based local clock synchronization loop utilizing the one-step predictive
UFIR filter with the ramp response $h_1(n, 1)$ [46].

![Figure 18](image-url)  
**Figure 18.** Typical steering errors of the GPS-locked OCXO-based clock obtained with the UFIR filter having
the ramp response (18) [46].

the TDEV of the unlocked crystal clock increases with large averaging time, whereas the reference GPS
pulse per second (1PPS) signal with sawtooth has a uniform TDEV function. The UFIR filter included
to the loop allows for an efficient steering of clock errors with large averaging times. An additional
smoothing filter (1-order LP, $T = 1000$ s) completes the picture, providing a substantial noise reduction
at small averaging times. As a result, the clock TDEV traces much lower the mask specified in the ITU-T
Recommendation G.811 (Timing characteristics of primary reference clocks) for digital communications
networks.

Other efficient applications of the structure shown in Fig. 17 were reported in several papers [65–67].
For example, Fig. 19 shows the pseudolite time synchronization method using a GPS timing receiver
as developed in [65]. In GPS ground applications, the term pseudolite (or pseudo satellite) is used to
refer to some equipment that is not a satellite but which performs a function commonly in the domain
of satellites. When a GPS timing receiver is used, operation is enabled using a pseudolite clock source
with low time accuracy, compared to when a pseudolite time synchronization station is used, because the
process of error data accumulation and processing is not required. The pseudolite time synchronization
method utilizing GPS timing 1PPS signals is illustrated in Fig. 19. A simulation platform that consists of
a clock generator, a clock synchronization algorithm processor, and a performance analyzer was designed
in [65] as shown in Fig. 20. Different kinds of pseudolite clocks were selected and an UFIR synchronization
Figure 19. GPS timing receiver 1PPS reference pseudolite time synchronization method [65].

Figure 20. Pseudolite time synchronization platform utilizing the UFIR filter-based algorithm [65].
8.5 Holdover in GPS-Based Timekeeping Using Predictive UFIR Filters

In synchronization networks, there is an operating condition under which a clock loses its controlling input and uses stored data, acquired while in locked operation, to control its output. The condition is termed holdover which begins when the clock output no longer reflects the influence of a connected external reference, or transition from it. Holdover terminates when the output of the clock reverts to locked mode condition [61].

Applications of the $p$-step UFIR predictive filters and predictors to bridge the holdover gap in clock synchronization when the GPS signal is unavailable or has an insufficient accuracy were reported in [50, 63]. Figure 21 sketches the best and worst cases of the holdover organized using the $p$-step ramp UFIR predictor as shown in [50]. Definitely the prediction error will grow with each new step $p$. But this is, nevertheless, a solution when measurement is unavailable.

8.6 GPS Carrier Tracking Using $P$-step UFIR Predictor

In GPS applications, a GPS receiver acquires and then tracks a GPS signal. After acquiring the GPS signal while in an acquisition mode, the GPS receiver then operates in a tracking mode where the carrier frequency and code phase of the signal are estimated using the Costas loop and delay lock loop, respectively [68]. When the carrier power-to-noise density ($C/N_0$) of the GPS signal drops below a threshold level the receiver is forced to operate in tracking mode. It returns back to the acquisition mode once the $C/N_0$ ratio again exceeds the threshold.

An efficient scheme for GPS carrier tracking utilizing the $p$-step UFIR predictor was recently patented as [68]. Here, the UFIR linear predictor with ramp response was modified to operate on a horizon of $N + 1$ points, from 0 to $N$,

$$
\tilde{h}_1(n, x) = 2^{2N^2 + N(1 + 3x) - 3n(N + 2x)} N(N + 1)(N + 2).
$$

Beyond a high accuracy inherent to linear predictors, some other benefits of (59) were recognized in [68]: BIBO stability, ease of use and implementation in hardware, and that it does not require any information besides the tap inputs.

8.7 Target Tracking Using UFIR Filter

Tracking of moving targets and objects is another field of applications for UFIR filters with polynomial impulse responses $h_k(n, x)$. In [69] [70], the UFIR filter was applied to 3D target tracking with three
states in each coordinate. That resulted in \((9 \times 9)\) dimensions of the system matrix \(F\) in (41) and \((3 \times 9)\) dimensions of the observation matrix \(H\) in (42). As can be seen in Fig. 22 which demonstrates tracking errors, the Kalman filter loses to the UFIR filter under the unknown noise statistics and in the presence of uncertainties. Besides, the authors were successful in modifying the UFIR filter to provide better accuracy than in the original solution.

9 Conclusions

Now 60 years after Johnson has shown how to derive the UFIR filter for polynomial signals in white Gaussian noise, the fundamental properties of UFIR estimation have been characterized and straightforward design and implementation tools developed. There is no universal UFIR structure providing unbiased filtering of arbitrary signals. However, polynomial signals of degree \(m\) can be filtered unbiasedly if the UFIR filter has the same degree. The filter degree can be larger, but noise reduction will be less efficient. Taking this into account, the order \(m\) is more or less limited to 3 and many of the practical applications use \(m\) equal 1 or 0.

The family of the polynomial impulse responses of UFIR filters is orthogonal which allows using this technique for signals approximation, interpolation, and extrapolation.

The UFIR method has been extended to cover a variety of relevant signal processing tasks including filtering, smoothing, predicting, and differentiating. The examples taken from high accuracy instrumentation and measurement environments show the utility of the method. In addition, the UFIR method has opened up a new class of nonlinear signal processing methods, FIR-median hybrid structures, which combine the best properties of linear and median worlds, the design flexibility and noise attenuation of linear operations and the root signal characteristics of median operations.

References

61. “EN 300 462-4-1 v1.1.1, European Standard (Telecommunications series) Transmission and Multiplexing (TM); Generic requirements for synchronization networks; Part 4-1: Timing characteristics of slave clocks suitable for synchronization supply to Synchronous Digital Hierarchy (SDH) and Plesiochronous Digital Hierarchy (PDH) equipment.”


