ZERO PHASE SMOOTHING OF RADIO CHANNEL ESTIMATES

Michael McGuire, Alireza Movahedian, Mihai Sima

University of Victoria, Department of Electrical and Computer Engineering
P.O. Box 3055 Stn CSC, Victoria, BC V8W 3P6, Canada
{mmcguire,amovahed,msima}@ece.uvic.ca

ABSTRACT

Modern radio systems require accurate radio channel estimates in the presence of fast fading. This paper proposes an iterative radio channel estimation system where the detected and decoded symbols from the previous iteration are used as 'virtual' pilot symbols to improve estimation accuracy. In each iteration of the receiver algorithm, a low order Kalman filter provides a rough estimate of the channel gains. This rough estimate is refined to a more accurate radio channel estimate with a zero phase de-noising filter. This two stage channel estimation system is capable of providing near optimal channel estimation at a much lower computational cost than prior art. It is demonstrated that the proposed radio channel estimation is capable of supporting low error data reception at high fading rates. The algorithm can be implemented efficiently on modern multicore computing systems.

***Index Terms*—** iterative signal processing, channel estimation, radio communications

1. INTRODUCTION

For efficient data detection, radio receivers require accurate estimates of the channel gains. Classic channel estimation systems transmit known pilot signals and measure the resulting signal at the receiver, employing interpolation to estimate the channel gains between the pilot symbols [1]. In fast fading channels, the density of the pilot symbols required for sufficiently accurate channel estimation to support low error performance does not permit high data rates. To address this problem, iterative receiver algorithms have been proposed [2]. These methods use sparse pilots signals to initially estimate the radio channel for preliminary data detection and decoding. To improve data detection, these methods then re-estimate the radio channel using the previously detected data values as 'virtual' pilots which are much denser in time than the dedicated pilot symbols. This procedure is iterated until the detected data values converge. With iterative channel estimation/data detection, the required density of dedicated pilots symbols to achieve acceptable error performance is greatly reduced. The computational costs of prior art iterative methods of channel estimation under fast fading are too large for commercial field deployment [2].

Kalman filters for channel estimation have a computation cost per sample proportional to the length of the state vectors, S, squared [2, 3]. The state vector length is given by $S = L \cdot M$ where L is the number of propagation paths and M is a constant proportional to the memory of the radio channel estimation system. Channel fading processes require large memory M for accurate estimation [4], which makes the computational cost of pure Kalman filter-based channel estimation large. This paper proposes the use of a low-memory Kalman filter combined with a smoother for radio channel estimation. The Kalman filter has a low memory M so it provides a rough estimate of the channel gains for each propagation path at a low cost. To reduce the channel estimation error, a zero phase filter is applied to the estimated channel gains for each propagation path. Zero phase filters use digital infinite impulse response (IIR) filters with special processing to remove the phase distortions [5]. The channel gain estimates for N signal samples are buffered. The buffered channel gain estimates are passed through an IIR filter first in the forward direction. This filtered signal is then reversed and passed through the IIR filter again. The initial state of the reverse IIR filter is determined from the final state of the forward IIR filter to avoid negative effects from signal truncation. The reverse IIR filter output is reversed again to obtain the smoothed channel estimate. This procedure allows the system to enjoy the high selectivity and low computational cost of an IIR filter without phase distortion. The application of the zero-phase filter allows low cost channel estimation with near optimal estimation error.

Section 2 discusses the models of the radio channel. Section 3 will describe the structure of the proposed radio receiver and the channel smoother. Section 4 contains simulation results demonstrating the effectiveness of the proposed receiver. The conclusions of the paper and future research directions will be described in Section 5.

Notation: Matrices and vectors are denoted with bold uppercase and lowercase letters such as $\boldsymbol{M}$ and $\boldsymbol{x}$. Entry k of vector $\boldsymbol{x}$ is denoted $\boldsymbol{x}_k$. Entry n of row k of matrix $\boldsymbol{M}$ is denoted $\boldsymbol{M}_{k,n}$. The Kronecker matrix product is denoted as $\otimes$. The size N identity matrix is denoted as $\boldsymbol{I}_N$. The all zero matrix of with n rows and k columns is denoted as $\mathbf{0}_{n\times k}$. The matrix transpose and Hermitian transpose operations are denoted with superscript T and H symbols respectively.

© EURASIP, 2012 - ISSN 2076-1465

2. CHANNEL MODEL

The received radio signal at baseband is given as

$$\mathrm{y}(n) = \sum_{l=0}^{L-1} \mathrm{g}_l(n)\,\mathrm{s}(n-l) + \mathrm{v}(n) \tag{1}$$

where L is the number of propagation paths, $\mathrm{g}_l(n)$ is the gain of path l at sample time n, $\mathrm{s}(n)$ is the transmitted symbol at time n and $\mathrm{v}(n)$ is additive white Gaussian measurement noise with a variance of ${\sigma_v}^2$. The channel gains for each path is assumed to be subject to Rayleigh fading following Jake's model with the autocorrelation for propagation path l given by $\mathrm{R}_{gg}^l(d) = \mathrm{E}[\mathrm{g}_l(n)\,\mathrm{g}_l(n+d)] = P_l \mathcal{J}_0(2\pi f_d T_s d)$ where P_l is the mean power gain for propagation path l, $\mathcal{J}_0(\cdot)$ is the zeroth order Bessel function of the first kind, f_d is the Doppler frequency, and T_s is the sampling period [6]. The channel gains for different propagation paths are assumed to be independent. From $\mathrm{R}_{gg}^l(d)$, the power spectral density of the radio channel gains for path l is defined as

$$\mathrm{P}_{gg}^l(f) = \begin{cases} \dfrac{\mathrm{P}_l}{\pi f_d \sqrt{1-(f/f_d)^2}} & |f| < f_d \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

The optimal mean square error for channel gain estimation is $\epsilon^2 = \int_{-f_d}^{f_d} N_0 \mathrm{P}_{gg}^l(f) / \left[N_0 + \mathrm{P}_{gg}^l(f)\right] \mathrm{d}f$, calculated from Weiner filter theory, where N_0 is the noise density [7].

In prior art, radio channel gains have been accurately estimated via the use of basis expansion models (BEMs) of the channel process [2, 8]. For a block of T_p samples such that $n = 1 \cdots T_p$ the channel gain for a single tap is given as weighted sum of B basis functions; a vector of T_p samples of channel gains for propagation path l is given by $\boldsymbol{g}^l = \boldsymbol{E}\boldsymbol{x}^l$ where $\boldsymbol{E}$ is a $T_p \times M$ matrix containing the BEM basis vectors in each of its columns, and $\boldsymbol{x}^l$ is a length M vector holding the BEM coefficients for propagation path l. The discrete prolate spheroidal sequence (DPSS) BEM has been employed for channel modelling with the columns of $\boldsymbol{E}$ specified as the B eigenvectors associated with the highest magnitude eigenvalues of the matrix $\boldsymbol{C}$ defined as $\boldsymbol{C}_{r,c} = \sin(2\pi(r-c) f_d T_s) / [\pi(r-c)]$ with $B \geq 2\lceil f_d T_p T_s \rceil + 1$ [3, 9]. Channel estimation systems need to know the covariance of the channel coefficients, $\boldsymbol{R}_{XX}$, for each propagation path. The DPSS provides basis functions which are maximally concentrated in the Doppler frequency band. They can be calculated as $\boldsymbol{R}_{XX} = \mathrm{E}\left[\boldsymbol{x}^l (\boldsymbol{x}^l)^H\right] = \boldsymbol{E}^H \boldsymbol{R}_{gg} \boldsymbol{E}$ where $\boldsymbol{R}_{gg}$ is the covariance matrix of the channel gains within a BEM block, so that $[\boldsymbol{R}_{gg}]_{r,c} = \mathrm{R}_{gg}(r-c)$ assuming $P_l = 1$.

The accuracy of radio channel BEM estimation improves as T_p and/or B increases. Unfortunately, the cost of estimating the radio channel coefficients per input sample is proportional to the number of BEM coefficients B which is itself proportional to T_p. To obtain accurate channel gain estimates at lower values of B which permits low computational cost, we propose to combine channel state estimates between frames. This is discussed in the next section.

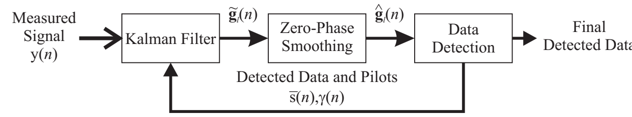

Fig. 1. Receiver configuration

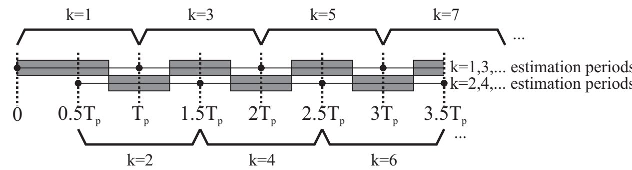

Fig. 2. Overlapping channel estimation blocks

3. RADIO RECEIVER

The structure of the radio channel estimation system we propose is shown in Figure 1. During each iteration, the radio channel is estimated using a Kalman filter based system using the value of the pilot signals and detected data symbol values from the previous iteration. A low order Kalman filter is employed to minimize computational cost. As a result of this, the Kalman filter provides only a noisy estimate of the radio channel gains, $\tilde{\mathrm{g}}_l(n)$ for $n = 1 \cdots N$ and $l = 0 \cdots L-1$. A zero phase filter, described below in more detail, is then employed to calculate improved channel gain estimates $\hat{\mathrm{g}}_l(n)$. These channel gain estimates are used for data detection and decoding. The estimates of the transmitted data symbols are also fed back to the channel estimation as 'virtual' pilots for the next iteration. The data symbol feedback consists of vectors of the mean estimated data symbols, $\bar{s}(n)$ and the variance of the estimated data symbols, $\gamma(n)$ which provides information about the confidence of the data detection/decoding system in each data symbol value. In the first iteration, only the values of pilot symbols are known and $\bar{\mathrm{s}}(n) = 0$ and $\gamma(n) = 1$ for all other symbols.

It is known that BEMs provide less accurate descriptions of signals near the start and end of each modeled period. To avoid this problem, the previous work proposed a channel estimation system with BEM periods of length T_p overlapping by $T_p/4$ samples at each end so that samples near the start or end of any given BEM period are not used for data detection [3]; only the samples from $T_p/4$ to $3T_p/4$ within a given BEM block are used for channel estimation except at the end or beginning of the measurement interval. For example, the first BEM block, $k = 1$, estimates the channel BEM coefficients for samples $1 \cdots T_p$ but the only the es-

Algorithm 1 Channel Estimation

Inputs:

- Number of channel samples: N
- Number of propagation paths: L
- Signal measurements: $y(n)$ for $n = 1 \cdots N$
- For $n = -L + 1 \cdots N$:
  - Estimated data symbols: $\bar{s}(n)$
  - Variance of estimated data symbols: $\gamma(n)$

Outputs:

- Channel gains vectors: $\tilde{g}_l(n)$ for $n = 1 \cdots N$

$n_s \leftarrow 1$
while $n_s + T_p \le N$ **do**
    if $n_s + T_p \ge N$ **then**
        $n_s \leftarrow N - T_p + 1$
    end if
    $b \leftarrow T_p/4$
    if $n_s = 1$ **then**
        $b \leftarrow 1$
    end if
    $e \leftarrow 3T_p/4$
    if $n_s + T_p - 1 = N$ **then**
        $e \leftarrow T_p$
    end if
    $\tilde{\boldsymbol{x}} \leftarrow$ Channel Kalman Filter $(n_s, \mathrm{y}(n), \bar{s}(n), \gamma(n))$
    for $l = 1 \cdots L$ **do**
        $\boldsymbol{v} \leftarrow \boldsymbol{E}\tilde{\boldsymbol{x}}$
        $\tilde{g}_l(n_s + b - 1 \cdots n_s + e - 1) \leftarrow \boldsymbol{v}_{b \cdots e}$
    end for
    $n_s \leftarrow n_s + T_p/2$
end while
for $l = 0 \cdots L - 1$ **do**
    Apply zero phase filter to $\tilde{g}_l(n)$ to obtain $\hat{g}_l(n)$
end for

Algorithm 2 Channel Kalman Filter

Inputs:

- Starting index: n_s
- Signal measurements: $\mathrm{y}(n)$ for $n = n_s \cdots n_s + T_p - 1$
- For $n = n_s - L + 1 \cdots n_s + T_p$:
  - Estimated data symbols: $\bar{s}(n)$
  - Estimated data variance: $\gamma(n)$

Outputs:

- Estimated channel gain BEM coefficients: $\tilde{\boldsymbol{x}}$

$\hat{\boldsymbol{x}}(0|0) \leftarrow \mathbf{0}_{(M\dot{L})\times 1}$
$\boldsymbol{P}(0|0) \leftarrow \boldsymbol{I}_L \otimes \boldsymbol{R}_{XX}$
for $n = 1 \cdots T_p$ **do**
    $\hat{\boldsymbol{x}}(n|n-1) \leftarrow \hat{\boldsymbol{x}}(n-1|n-1)$
    $\boldsymbol{P}(n|n-1) \leftarrow \boldsymbol{P}(n-1|n-1) + {\sigma_e}^2 \boldsymbol{I}_{M \cdot L}$
    $\boldsymbol{e} \leftarrow$ row n of $\boldsymbol{E}$
    $\hat{\boldsymbol{s}} \leftarrow [\bar{s}(n + n_s - 1) \;\cdots\; \bar{s}(n + n_s - L)]$
    $\boldsymbol{H} \leftarrow \bar{\boldsymbol{s}} \otimes \boldsymbol{e}$
    $c \leftarrow \boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^H + {\sigma_v}^2$
    for $l = 0 \cdots L - 1$ **do**
        $\boldsymbol{x}^l \leftarrow [\boldsymbol{x}(n|n-1)]_{l \cdot M + 1 \cdots (l+1)M}$
        $\boldsymbol{P}^l \leftarrow [\boldsymbol{P}(n|n-1)]_{l \cdot M + 1 \cdots (l+1)M, l \cdot M + 1 \cdots (l+1)M}$
        $v_l \leftarrow \gamma(n_s + n - 1 - l)\, \boldsymbol{e} \left[\boldsymbol{x}^l \left(\boldsymbol{x}^l\right)^H + \boldsymbol{P}^l\right] \boldsymbol{e}^H$
        $c \leftarrow c + v_l$
    end for
    $\boldsymbol{K} \leftarrow \boldsymbol{P}(n|n-1)\, \boldsymbol{H}^H / c$
    $\boldsymbol{z} \leftarrow \mathrm{y}(n + n_s - 1) - \boldsymbol{H}\hat{\boldsymbol{x}}(n|n-1)$
    $\hat{\boldsymbol{x}}(n|n) \leftarrow \hat{\boldsymbol{x}}(n|n-1) + \boldsymbol{K}\boldsymbol{z}$
    $\boldsymbol{P}(n|n) \leftarrow \left[\boldsymbol{I}_{(M \cdot L)} - \boldsymbol{K}\boldsymbol{H}\right] \boldsymbol{P}(n|n-1)$
end for
$\tilde{\boldsymbol{x}} \leftarrow \hat{\boldsymbol{x}}(T_p|T_p)$

timated channel gains for $1 \cdots 3T_p/4$ are used for data detection. The second BEM block, $k = 2$, estimates channel gain coefficients for samples $T_p/2 \cdots 3T_p/2 - 1$ with the estimated channel gains for $3T_p/4 \cdots 5T_p/4$ used for data detection. The third BEM block, $k = 3$, estimates channel gains for samples for $T_p \cdots 2T_p - 1$ with only the estimated values for $5T_p/4 \cdots 7T_p/4$ used for data detection. The setup of BEM blocks is shown in Figure 2 where the darker sections show the estimated channel gains for each BEM block used for data detection. The full description of the algorithm to estimate the channel gains using a Kalman filter to is described in Algorithm 1 and Algorithm 2.

The proposed channel estimation algorithm uses a Kalman filter algorithm [10] as shown in Algorithm 2 to estimate the channel gain BEM coefficient state vector $\boldsymbol{x}(n)$ which is the concatenation of the L BEM coefficient vectors, $\boldsymbol{x}^l(n)$ for the propagation paths $l = 0 \cdots L - 1$. The Kalman filter algorithm assumes that the channel BEM coefficients remain nearly constant over each BEM period; i.e., $\boldsymbol{x}(n+1) = \boldsymbol{x}(n) + \boldsymbol{w}(n)$ where $\boldsymbol{w}(n)$ is a white Gaussian vector process with the covariance of each random vector $\mathrm{Cov}[\boldsymbol{w}(n)] = {\sigma_e}^2 \boldsymbol{I}_{M \cdot L}$ with ${\sigma_e}^2$ being a small non-zero value. Ideally, ${\sigma_e}^2$ should be zero but using a small value such as 10^{-4} improves the numerical stability of the filtering algorithm. An equation relating the BEM coefficients to the observed signal is obtained by substituting the calculation of the channel gain from BEM coefficients into Eqn (1). The resulting measurement equation relating the state vector to the received signal is then $\mathrm{y}(n) = \boldsymbol{H}(n) \cdot \boldsymbol{x}(n) + \mathrm{v}(n)$ where $\boldsymbol{H}(n) = [\bar{\boldsymbol{s}}(n) \otimes \boldsymbol{e}(n)]$ when $\boldsymbol{e}(n)$ is the row of the basis matrix $\boldsymbol{E}$ for the sample n of the current BEM block, and $\bar{\boldsymbol{s}}(n) = [\bar{s}(n) \;\cdots\; \bar{s}(n - L + 1)]$.

The variance of the estimated data symbols adds ambiguity to the measurement equation. The channel estimation algorithm adjusts for this by increasing the assumed variance of the measurement noise vector $\mathrm{v}(n)$ for measurements where the influencing samples are not known perfectly. Since the channel gain process is independent of the transmitted data sequence and assuming the channel gain process for each propagation path is independent, the variance of $\mathrm{y}(n)$ is calculated as

$$\mathrm{Var}\left[\mathrm{y}\left(n\right)\right] = \boldsymbol{H}\left(n\right) \cdot \mathrm{Cov}\left[\boldsymbol{x}\left(n\right)\right] \cdot \left[\boldsymbol{H}\left(n\right)\right]^{H} + \sigma_v^{\,2} + \sum_{l=0}^{L-1} \boldsymbol{e}\left(n\right) \boldsymbol{R}_{xx}^{l} \left[\boldsymbol{e}\left(n\right)\right]^{H} \gamma\left(n-l\right) \quad (3)$$

where $\boldsymbol{R}_{xx}^{l} = \hat{\boldsymbol{x}}^{l}\left(n\right)\left[\hat{\boldsymbol{x}}^{l}\left(n\right)\right]^{H} + \mathrm{Cov}\left[\boldsymbol{x}^{l}\left(n\right)\right]$. The first line of (3) is just the measurement covariance calculation of the standard Kalman filter [10]. The second line computes the extra variance caused by ambiguity in the knowledge of the transmitted symbols. This calculation is performed in lines 10-15 of the Algorithm 2. The effect of this addition to the calculation is to speed up the convergence of the algorithm since it causes the channel estimation to emphasis measurements dominated by symbols which are known well and deemphasizes measurements dominated by symbols which are unknown. This adjustment is the main reason for the faster convergence of this channel estimation algorithm compared to the algorithm presented in [3].

The key difficulty with estimation of the radio channel with linear filters is that the radio channel gain processes are bandlimited, as shown in Eqn. (2), which requires the filters for optimal channel estimation to have extremely long impulse responses [4]. The long impulse response requirement translates into large memory requirements. For BEM channel estimation techniques, this memory requirement maps to using long BEM periods with [3] suggesting BEM periods of $T_p = 500$ data samples for acceptable BER results. Since the computational cost of the Kalman filter calculation per sample is on the order of $\mathcal{O}\left(J^2\right)$ where $J = B \cdot L$ is the length of the state vector and B is proportional to T_p, long BEM periods map to large computational costs.

This proposed method reduces the need for long BEM periods by employing the zero phase filter noise removal method stage. A Kalman filter calculates noisy estimates of the channel gains and the gain estimates for N samples are buffered. The channel gain estimate vectors for each propagation path are passed through an IIR filter first in the forward direction. This filtered signal is then reversed and passed through the IIR filter again with the final output time reversed again to obtain the final channel estimate. The initial state of the reverse IIR filter is calculated from the final state of the forward IIR filter to reduce truncation effects [11]. This procedure removes phase distortion, providing high selectivity at low computational cost [5]. For fast fading channels with normalized fading rates up to $f_d T_s = 0.01$, a zero phase filter with a component Elliptical approximation IIR filter of order $O = 6$ with a maximum passband ripple of $A_p = 0.001$ dB, a minimum stopband attenuation of $A_a = 14$ dB, a normalized passband edge of $W_p = 0.02$ and a normalized stopband edge of $W_a = 0.025$ allows good channel estimation performance with BEM periods as low as $T_p = 100$. The zero phase filter is applied to the estimated channel gain vector for each propagation path separately. The cost per sample of the zero phase filter is proportional to $\mathcal{O}\left(O \cdot L\right)$. To achieve comparable results without the zero phase filter, a BEM period of $T_p = 500$ was needed [3], showing the zero phase implementation gives a cost reduction of a factor of ≈ 25.

Another advantage of the proposed algorithm is that it exhibits a large degree of parallelism, and thus it is compatible with multicore implementations. Specifically, the Kalman filter calls on Line 14 of Algorithm 1 does not have any data interdependencies, so independent filter instances (one instance for each block of symbols shown in Figure 2) can be run in parallel subject to the available hardware capacity. In addition, since the covariance matrix c has the size 1×1 (thus, it is a scalar), the Kalman filter gain $\mathbf{K}$ can be calculated by a simple division rather by matrix inversion. The large majority of the remaining operations in the Kalman filter algorithm are matrix multiplications, that are very well supported in current Field-Programmable Gate Arrays (FPGA) architectures. The intrinsic parallelism and the match with existing FPGA architectures makes the proposed algorithm commercially viable.

For the simulations results described in Section 4, the data detection is performed using a Kalman filter equalizer followed by a soft input/soft output data detector/decoder. The full details of this data detection/decoder system are described in [3]. However, the channel estimation results are not specific to this specific data detection system, any data detection system may be used which can provide soft output feedback of the data symbol values to the channel estimation system.

4. RESULTS

This section presents results from the application of the preceding algorithm for a multipath propagation radio channel with $L = 3$ equally powerful propagation paths. The normalized fading rate is set to $f_d T_s = 0.01$. The data transmission signal considered is a standard binary data sequence modulated onto a 16-QAM gray-coded constellation after being encoded with a rate 1/2 convolutional code with the octal generator of (133,171). To assist with the initial radio channel estimation, we insert $l_p = 5$ pilots symbols before every $l_s = 20$ data symbols. The overall data rate is $1/2 \cdot l_s/(l_s + l_p) \cdot 4$ bits per sample. We assume that power control is employed so that the mean power of the received signal is unity: $\sum_{l=0}^{L-1} P_l = 1$. The data is encoded into blocks of $D = N \cdot l_s/\left(l_s + l_p\right)$ symbols with interleaving performed over the block $4 \cdot D$ coded bits. For our simulations, $D = 10000$ symbols so that the channel estimation block length was $N = 12500$. The channel estimation algorithm uses $T_p = 100$ and $B = 5$. The proposed algorithm's receiver was limited to 5 iterations.

The BER results from the application of the proposed algorithm, compared with the Extended Kalman Filter of [2] and an receiver operating the ideal knowledge of the CSI are shown in Figure 3. The algorithm from [3] is not plotted but it gives nearly identical BER performance to the proposed algorithm, though it requires 15 iterations to converge, as op-

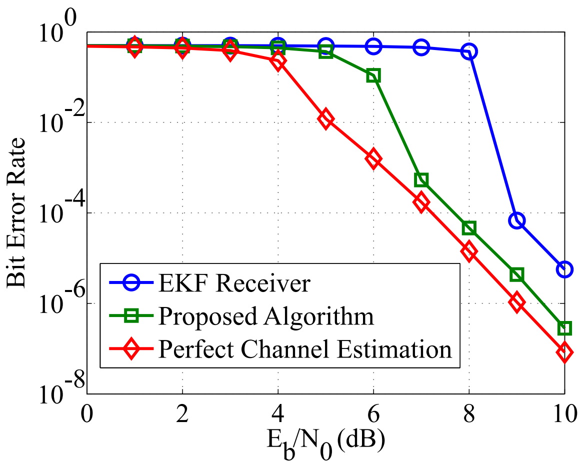

Fig. 3. Bit Error Rate Results

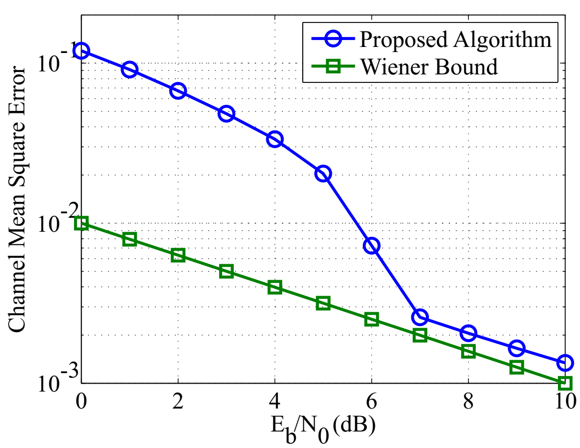

Fig. 4. Mean Square Error Results

posed to 5 iterations for the proposed algorithm . As described above, the computational cost of each iteration for the proposed algorithm is much lower than the other two algorithms. It can be seen that the proposed algorithm gives BER results within 0.5 dB of the ideal CSI case for $E_b/N_0 > 7$ dB.

The mean square error of the proposed channel estimation algorithm for one of the channel taps is shown in Figure 4 where the algorithms performance is compared with the ideal Weiner filter's mean square error. It can be seen that the proposed algorithms MSE is nearly as good as the ideal case at higher E_b/N_0 ratio values. This suggests that the proposed algorithm is nearly optimal in terms of channel estimation performance as well as having low computational cost.

5. CONCLUSIONS

We presented a new channel estimation algorithm which uses a Kalman filter to estimate the parameters of a fast fading radio channel using a zero phase filter to de-noise the channel gain estimates. This allows a lower memory Kalman filter to be used reducing the overall computational cost. The new algorithm provides nearly optimal BER and channel MSE performance and can be implemented on multicore processing units. Future work will investigate the relationship between the parameters of the component IIR filter and the required BEM period for the Kalman filter channel estimation.

6. REFERENCES

[1] X. Ma, G.B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *IEEE Trans. Signal Processing*, vol. 51, no. 5, pp. 1351–1366, May 2003.

[2] H. Kim and J.K. Tugnait, "Turbo equalization for doubly-selective fading channels using nonlinear Kalman filtering and basis expansion models," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 2076–2087, June 2010.

[3] A. Movahedian and M. McGuire, "An iterative receiver for fast fading channels using two Kalman filters," in *IEEE BWCCA 2011*, October 2011, pp. 122–127.

[4] R.J. Lyman, "Optimal mean-square prediction of the mobile-radio fading envelope," *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 819–824, March 2003.

[5] A. Antoniou, *Digital Signal Processing: Signals, Systems, and Filters*, McGraw-Hill, 2006.

[6] A.F. Molisch, *Wireless Communications*, John Wiley & Sons, 2005.

[7] P.Z. Peebles, *Probability, Random Variables, and Random Signal Principles*, McGraw Hill, 2001.

[8] Ping Wan, M. McGuire, and Xiaodai Dong, "Near-optimal channel estimation for OFDM in fast-fading channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3780–3791, October 2011.

[9] T. Zeman and Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Processing*, vol. 53, no. 9, pp. 3597–3607, September 2005.

[10] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 2002.

[11] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 988–992, April 1996.