RECURSIVE LEAST-SQUARES DOUBLY-SELECTIVE CHANNEL ESTIMATION USING EXPONENTIAL BASIS MODELS AND SUBBLOCK-WISE TRACKING

Jitendra K. Tugnait and *Shuangchi He*

Department of Electrical & Computer Engineering, Auburn University, Auburn, AL 36849

ABSTRACT

An adaptive channel estimation scheme, exploiting the oversampled complex exponential basis expansion model (CE-BEM), is presented for doubly-selective channels where we track the BEM coefficients. We extend/modify the subblock-wise tracking method using time-multiplexed (TM) training recently proposed by [1]. Two finite-memory recursive least-squares (RLS) algorithms, including the exponentially-weighted and the sliding-window RLS algorithms, are respectively applied to track the channel BEM coefficients. Simulation examples illustrate the superior performance of our scheme to the conventional block-wise channel estimator, and demonstrate its improvement on our previous work in [1].

***Index Terms*—** Doubly-selective channels, adaptive channel estimation, basis expansion models, recursive least-squares

1. INTRODUCTION

Due to multipath propagation and Doppler spread, wireless channels are characterized by frequency- and time-selectivity. Accurate modeling of time-variations of the channel plays a crucial role for estimation and tracking purposes. Among various models for channel time-variations, basis expansion models (BEM) depict evolutions of the channel over a period (block) of time, in which the time-varying channel taps are expressed as superpositions of time-varying basis functions in modeling Doppler effects, weighted by time-invariant coefficients [2].

In [1], a subblock-wise tracking approach was proposed for doubly-selective channels using time-multiplexed (TM) training. It exploits the complex exponential BEM (CE-BEM) for the overall channel variations of each (overlapping) block, and a first-order autoregressive (AR) model to describe the evolutions of the BEM coefficients. Since the time-varying nature of the channel can be well captured in the CE-BEM by (known) Fourier basis functions, the time-variations of the (unknown) BEM coefficients are likely much slower than those of the channel, and thus more convenient to track in fast-fading environments [1]. The slow-varying BEM coefficients are updated via Kalman filtering at each training session; during information sessions, channel estimates are generated by the CE-BEM using the estimated BEM coefficients [1]. This approach achieves better performance in fast-fading environments, than using conventional symbol-wise AR models or block-wise BEM representations [1].

This work was supported by NSF under Grant ECS-0424145.

The approach in [1], however, assumes that each BEM coefficient follows a first-order AR process, which is not necessarily true for a real-world channel, and possibly incurs modeling error in estimation. In this paper, we seek an adaptive channel estimation scheme with no a priori models for BEM coefficients. Two adaptive filtering algorithms with finite memory are considered for subblock-wise channel tracking: the exponentially-weighted recursive least-squares (RLS) algorithm and the sliding-window RLS algorithm.

Decision-directed channel tracking using a polynomial BEM has been investigated in [3], where the BEM coefficients are updated every block via the recursive least-squares (RLS) algorithm within a sliding window. Decision-directed channel estimation using Kalman filtering and polynomial or CE-BEM for OFDM systems has been explored in [4, 5]. All these contributions consider block-by-block updating unlike our contribution where we exploit subblock-wise updating. The distinction is as follows. Several subblocks comprise one block. For parameter identifiability one needs the number of subblocks at least as large as the number of basis functions used for channel modeling.

Notations: Superscripts $*$, T, and H denote the complex conjugation, transpose, and complex conjugate transpose respectively. $\mathbf{I}_N$ is the $N \times N$ identity matrix, and $\mathbf{0}_M$ is the M-column null vector. We use $\lceil \cdot \rceil$ for integer ceiling and $\lfloor \cdot \rfloor$ for integer floor. The symbol $E\{\cdot\}$ denotes expectation, and $\otimes$ denotes the Kronecker product. $\delta(\tau)$ is the Kronecker delta, i.e., $\delta(\tau) = 1$ for $\tau = 0$, and $\delta(\tau) = 0$ otherwise.

2. SYSTEM MODEL

Consider a single-input multi-output (SIMO), frequency- and time-selective, finite impulse response (FIR) linear channel with N outputs. Let $\{s(n)\}$ denote a scalar sequence that is input to the $(L+1)$-tap channel with discrete-time response $\{\mathbf{h}(n;l)\}$ (N-column vector channel response at time n to a unit input at time $n-l$). Then the symbol-rate noisy channel output is given by $(n = 0, 1, \ldots)$

$$\mathbf{y}(n) = \sum_{l=0}^{L} \mathbf{h}(n;l)\, s(n-l) + \mathbf{v}(n) \tag{1}$$

where $\mathbf{v}(n)$ is white complex Gaussian noise, with zero mean and autocorrelation $E\{\mathbf{v}(n+\tau)\mathbf{v}^H(n)\} = \sigma_v^2 \mathbf{I}_N \delta(\tau)$. In TM training schemes, $s(n)$ can be either a training or an information symbol.

In CE-BEM [2, 6], over the k-th block consisting of an observation window of T_B symbols, the channel is represented as ($n = (k-1)\,T_B, (k-1)\,T_B + 1, \ldots, kT_B - 1$ and $l = 0, 1, \ldots, L$)

$$\mathbf{h}(n;l) = \sum_{q=1}^{Q} \mathbf{h}_q^{(l)} e^{j\omega_q n}, \tag{2}$$

where one chooses ($q = 1, 2, \ldots, Q$ and $K \geq 1$ is an integer)

$$T := KT_B, \quad Q \geq 2\lceil f_d T T_s \rceil + 1, \tag{3}$$

$$\omega_q := \frac{2\pi}{T}\left[q - (Q+1)/2\right], \quad L := \lfloor \tau_d / T_s \rfloor, \tag{4}$$

τ_d and f_d are respectively the delay spread and the Doppler spread, and T_s is the symbol duration. The BEM coefficients $\mathbf{h}_q^{(l)}$'s remain invariant during this block, but are allowed to change at the next block; the Fourier basis functions $\{e^{j\omega_q n}\}$ are common for every block. Treating the basis functions as known, estimation of a time-varying process is reduced to estimating the invariant coefficients over a block of T_B symbols.

The BEM period is $T = KT_B$, whereas the block size is T_B symbols. If $K \geq 2$, the Doppler spectrum is over-sampled (therefore (2) is called an over-sampled CE-BEM) [6], compared with the critically-sampled case corresponding to $K = 1$ [2]. For $K = 1$, the rectangular window of this truncated discrete Fourier transform (DFT)-based model introduces spectral leakage, resulting in significant amplitude and phase distortion at the beginning and the end of the observation window [7]. To mitigate this leakage, the over-sampled CE-BEM with $K = 2$ or 3 has been explored in [6].

We employ the TM training scheme of [2], which is optimal for channels following critically-sampled CE-BEM representations, in our subblock tracking approach. In [2], each transmitted block consisting of T_B symbols is segmented into P subblocks of $m_b := T_B/P$ symbols each. Every subblock consists of an information session of m_d symbols together with a succeeding training session of $2L+1$ symbols ($m_b = m_d + 2L + 1$). The training session contains an impulse guarded by zeros (silent periods), which has the structure

$$\mathbf{c}_p := \begin{bmatrix} \mathbf{0}_L^T & \gamma & \mathbf{0}_L^T \end{bmatrix}^T, \quad \gamma > 0. \tag{5}$$

Note that training impulses only occur at time $n_p := pm_b + m_d + L$, ($p = 0, 1, \cdots$). By (1) at time $n_p + l$, ($l = 0, 1, \cdots, L$), the received signal is given by

$$\mathbf{y}(n_p + l) = \gamma \mathbf{h}(n_p + l; l) + \mathbf{v}(n_p + l). \tag{6}$$

3. SUBBLOCK-WISE RLS TRACKING

Consider two overlapping blocks (each of T_B symbols) that differ by just one subblock: the "past" block beginning at time n_0 and the "present" block beginning at time $n_0 + m_b$. Since the two blocks overlap so significantly, one would expect the BEM coefficients to vary only a little from the past block to the present overlapping one. As in [1], we propose to track the BEM coefficients (rather than the channel taps) subblock by subblock for their variations.

Stack the BEM coefficients in (2) into vectors

$$\mathbf{h}^{(l)} := \begin{bmatrix} \mathbf{h}_1^{(l)T} & \mathbf{h}_2^{(l)T} & \cdots & \mathbf{h}_Q^{(l)T} \end{bmatrix}^T, \tag{7}$$

$$\mathbf{h} := \begin{bmatrix} \mathbf{h}^{(0)T} & \mathbf{h}^{(1)T} & \cdots & \mathbf{h}^{(L)T} \end{bmatrix}^T \tag{8}$$

of size NQ and $M := NQ\,(L+1)$ respectively. The coefficient vectors of the p-th subblock will be denoted by $\mathbf{h}_q^{(l)}(p)$, $\mathbf{h}^{(l)}(p)$ and $\mathbf{h}(p)$. Defining

$$\mathbf{E}(n) := \begin{bmatrix} e^{-j\omega_1 n}\mathbf{I}_N & e^{-j\omega_2 n}\mathbf{I}_N & \cdots & e^{-j\omega_Q n}\mathbf{I}_N \end{bmatrix}^T,$$

the received signal at time $n_p + l$ is given by

$$\mathbf{y}(n_p + l) = \gamma \mathbf{E}^H(n_p + l)\,\mathbf{h}^{(l)}(p) + \mathbf{v}(n_p + l). \tag{9}$$

Further define

$$\mathbf{\Psi}(p) := \begin{bmatrix} \mathbf{E}(n_p) & & & \\ & \mathbf{E}(n_p+1) & & \\ & & \ddots & \\ & & & \mathbf{E}(n_p+L) \end{bmatrix}^H,$$

$$\mathbf{y}_s(p) := \begin{bmatrix} \mathbf{y}^T(n_p) & \mathbf{y}^T(n_p+1) & \cdots & \mathbf{y}^T(n_p+L) \end{bmatrix}^T,$$

and $\mathbf{v}_s(p)$ likewise. By (9), it follows that

$$\mathbf{y}_s(p) = \gamma \mathbf{\Psi}(p)\,\mathbf{h}(p) + \mathbf{v}_s(p), \tag{10}$$

which gives us a formulation to estimate the BEM coefficients using training sessions.

Our goal is to adaptively estimate the doubly-selective channel via subblock-wise tracking, by exploiting the invariance of the BEM coefficients over each block. Since CE-BEM is periodic with period T, the algorithm memory should be less than T to avoid periodicity of the BEM influencing the results. Therefore, an adaptive algorithm with finite memory is preferred which we implement via either exponentially-weighted RLS or sliding-window RLS approaches.

3.1. Exponentially-Weighted RLS Tracking

Based on (10), we can apply the exponentially-weighted RLS (EW-RLS) algorithm [8, Chapter 12] to tracking the channel BEM coefficients. Choose $\mathbf{h}$ to minimize the cost function

$$\beta\|\mathbf{h}\|^2 + \sum_{i=0}^{p} \lambda^{p-i}\|\mathbf{y}_s(p) - \gamma\mathbf{\Psi}(p)\mathbf{h}\|^2$$

where $\beta > 0$ is a regularization parameter and $0 < \lambda < 1$ is the forgetting factor. Unlike "standard" EW-RLS as applied to symbol-wise updating, in subblock-wise updating, one takes λ to be (much) smaller than one (0.6 or 0.7 for instance).

Mimicking [8, Chapter 12], the EW-RLS algorithm has the following steps:

1. Initialization:

$$\hat{\mathbf{h}}(-1) = \mathbf{0}_M, \quad \mathbf{P}(-1) = \beta\mathbf{I}_M.$$

2. RLS recursion: For $p = 0, 1, \ldots$

$$\boldsymbol{\Gamma}(p) = \lambda \mathbf{I}_{N(L+1)} + \gamma^2 \boldsymbol{\Psi}(p) \mathbf{P}(p-1) \boldsymbol{\Psi}^H(p),$$
$$\mathbf{K}(p) = \gamma \mathbf{P}(p-1) \boldsymbol{\Psi}^H(p) \boldsymbol{\Gamma}^{-1}(p),$$
$$\mathbf{e}(p) = \mathbf{y}_s(p) - \gamma \boldsymbol{\Psi}(p) \hat{\mathbf{h}}(p-1),$$
$$\hat{\mathbf{h}}(p) = \hat{\mathbf{h}}(p-1) + \mathbf{K}(p) \mathbf{e}(p),$$
$$\mathbf{P}(p) = \lambda^{-1} [\mathbf{P}(p-1) - \gamma \mathbf{K}(p) \boldsymbol{\Psi}(p) \mathbf{P}(p-1)]$$

where $\hat{\mathbf{h}}(p)$ is the estimate of $\mathbf{h}$ given the observations $\{\mathbf{y}_s(0), \mathbf{y}_s(1), \ldots, \mathbf{y}_s(p)\}$. Now we generate the channel for the entire p-th subblock by the estimate $\hat{\mathbf{h}}(p)$ via the CE-BEM (2) as

$$\hat{\mathbf{h}}(n; l) = \mathbf{E}^H(n) \hat{\mathbf{h}}^{(l)}(p) \tag{11}$$

for $n = pm_b, pm_b + 1, \ldots, (p+1) m_b - 1$. The definition of $\hat{\mathbf{h}}^{(l)}(p)$ is similar to (7).

3.2. Sliding-Window RLS Tracking

Compared with the EW-RLS algorithm that exponentially weakens the effects of all past data, the sliding-window RLS (SW-RLS) algorithm only utilizes the data in a sliding window of length W [8]. Since the BEM coefficients are invariant within a block of T_B symbols, we set $W = \lfloor T_B/m_b \rfloor$ subblocks so that only the present subblock and the past $W - 1$ subblocks within one block are used for adaptation. To this end, each iteration consists of a downdating stage that removes the oldest received subblock sample from the window, and an updating stage that inserts the next received subblock sample into the window [8, Problem 12.7, p. 750]. The cost function in this case is

$$\beta \|\mathbf{h}\|^2 + \sum_{i=p-W+1}^{p} \lambda^{p-i} \|\mathbf{y}_s(p) - \gamma \boldsymbol{\Psi}(p) \mathbf{h}\|^2$$

where if $p - W + 1 < 0$, we set $i = 0$.

Mimicking [8, Chapter 12], the SW-RLS algorithm has the following steps:

1. Initialization: For $p = 0, 1, \ldots, W - 1$, conduct channel tracking as in the EW-RLS algorithm, but with $\lambda = 1$. Then set

$$\hat{\mathbf{h}}_u(W-1) = \hat{\mathbf{h}}(W-1),$$
$$\mathbf{P}_u(W-1) = \mathbf{P}(W-1).$$

2. RLS recursion: For $p = W, W + 1, \ldots$

Downdating:

$$\boldsymbol{\Gamma}_d(p-1) = \mathbf{I}_{N(L+1)} - \gamma^2 \boldsymbol{\Psi}(p-W) \mathbf{P}_u(p-1) \boldsymbol{\Psi}^H(p-W),$$
$$\mathbf{K}_d(p-1) = \gamma \mathbf{P}_u(p-1) \boldsymbol{\Psi}^H(p-W) \boldsymbol{\Gamma}_d^{-1}(p-1),$$
$$\mathbf{e}(p-W) = \mathbf{y}_s(p-W) - \gamma \boldsymbol{\Psi}(p-W) \hat{\mathbf{h}}_u(p-1),$$
$$\hat{\mathbf{h}}_d(p-1) = \hat{\mathbf{h}}_u(p-1) - \mathbf{K}_d(p-1) \mathbf{e}(p-W),$$
$$\mathbf{P}_d(p-1) = \mathbf{P}_u(p-1) + \gamma \mathbf{K}_d(p-1) \boldsymbol{\Psi}(p-W) \mathbf{P}_u(p-1).$$

Updating:

$$\boldsymbol{\Gamma}_u(p) = \mathbf{I}_{N(L+1)} + \gamma^2 \boldsymbol{\Psi}(p) \mathbf{P}_d(p-1) \boldsymbol{\Psi}^H(p),$$
$$\mathbf{K}_u(p) = \gamma \mathbf{P}_d(p-1) \boldsymbol{\Psi}^H(p) \boldsymbol{\Gamma}_u^{-1}(p),$$
$$\mathbf{e}(p) = \mathbf{y}_s(p) - \gamma \boldsymbol{\Psi}(p) \hat{\mathbf{h}}_d(p-1),$$
$$\hat{\mathbf{h}}_u(p) = \hat{\mathbf{h}}_d(p-1) + \mathbf{K}_u(p) \mathbf{e}(p),$$
$$\mathbf{P}_u(p) = \mathbf{P}_d(p-1) - \gamma \mathbf{K}_u(p) \boldsymbol{\Psi}(p) \mathbf{P}_d(p-1).$$

Now $\hat{\mathbf{h}}_u(p)$ is the estimate of $\mathbf{h}$ based on the observations $\{\mathbf{y}_s(p-W+1), \mathbf{y}_s(p-W+2), \ldots, \mathbf{y}_s(p)\}$. Channel estimates are also generated for the p-th subblock using (11) by setting $\hat{\mathbf{h}}(p) = \hat{\mathbf{h}}_u(p)$.

4. SIMULATION EXAMPLES

A random doubly-selective Rayleigh fading channel is considered. We take $L = 2$ (3 taps) with output $N = 1$ in (1), and $\mathbf{h}(n; l)$ are zero-mean, complex Gaussian with variance $\sigma_h^2 = 1/(L+1)$. For different l's, $\mathbf{h}(n; l)$'s are mutually independent and satisfy Jakes' model. We consider a communication system with normalized Doppler spread $f_d T_s = 0.01$. The additive noise was zero-mean complex white Gaussian, and the (receiver) SNR refers to the average energy per symbol over one-sided noise spectral density. During information sessions the symbols are modulated by quadrature phase-shift keying (QPSK) with unit power. The training session is described by (5) with $\gamma = \sqrt{2L+1}$ so that the average symbol energy of training sessions is equal to that of information sessions. We select the period of the CE-BEM $T = 400$, and hence $Q = 9$ by (3). We set the subblock size $m_b = 20$ or 40 symbols.

Four channel estimation schemes are compared:

1. The block-wise channel estimation scheme in [2] (denoted by "block estimation [2]" in the figures), where the transmitted symbols are segmented into consecutive blocks each of T_B symbols. Every block consists of P ($P \geq Q$) subblocks as in Section 2. For each (non-overlapping) block, we estimate the BEM coefficients anew via a least-squares approach, and obtain the channel estimates over this block by the CE-BEM. We use an over-sampled CE-BEM with $T_B = T/2$ when $m_b = 20$ so as to suppress spectral leakage; whereas for $m_b = 40$, an over-sampled CE-BEM is not possible (cannot obtain $P \geq Q$), so that we choose $T_B = T$.
2. The subblock-wise Kalman tracking in [1] (denoted by "Kalman tracking [1]" in the figures): We assume the BEM coefficients follow a first-order AR model $\mathbf{h}(p) = \alpha \mathbf{h}(p-1) + \mathbf{w}(p)$, where the i.i.d. driving-noise vector $\mathbf{w}(p)$ is zero-mean complex Gaussian with auto-correlation $E\{\mathbf{w}(p) \mathbf{w}^H(p+\tau)\} = \sigma_w^2 \mathbf{I}_M \delta(\tau)$ and $\sigma_w^2 = \sigma_h^2(1 - |\alpha|^2)/Q$ [1]. For $m_b = 20$ and 40, we take $\alpha = 0.995$ and 0.97 respectively.
3. Proposed EW-RLS algorithm with $\beta = 1$. For $m_b = 20$ and 40, we take the forgetting factor $\lambda = 0.65$ and 0.5 respectively (those values were determined empirically: see also comments just before Sec. 3.1).

4. Proposed SW-RLS algorithm with $\beta = 1$. We take $T_B = T/2 = 200$, so that for $m_b = 20$ and 40, the window size $W = 10$ and 5 respectively.

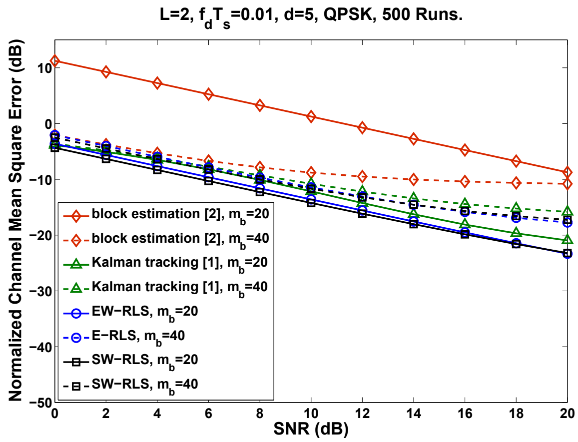

Fig. 1. *NCMSE vs SNR, under $f_dT_s = 0.01$, $m_b = 20$ and 40, with QPSK information symbols.*

The normalized channel mean square error (NCMSE) and the bit error rate (BER) of each scheme are studied. The NCMSE is defined as

$$\text{NCMSE} := \frac{\sum_{i=1}^{M_r}\sum_{n=0}^{T-1}\sum_{l=0}^{L}\left\|\hat{\mathbf{h}}^{(i)}(n;l) - \mathbf{h}^{(i)}(n;l)\right\|^2}{\sum_{i=1}^{M_r}\sum_{n=0}^{T-1}\sum_{l=0}^{L}\left\|\mathbf{h}^{(i)}(n;l)\right\|^2}$$

where $\mathbf{h}^{(i)}(n;l)$ is the true channel and $\hat{\mathbf{h}}^{(i)}(n;l)$ is the estimated channel at the i-th run, among total M_r runs. The BER's for the schemes are evaluated by employing a decision-feedback equalizer (DFE) [9] at the receiver, designed by using the channel estimates from each scheme, with feedforward length $l_f = 8$, feedback length $l_b = 2$, and delay $d = 5$. In each run, a symbol sequence of length 5000 is generated and fed into a random doubly-selective channel. The first 200 symbols are treated as training overhead and thus discarded in evaluations. All simulation results are based on 500 runs.

In Figs. 1 and 2, the performances of the four schemes under different SNR's are compared. Since the subblock-wise tracking methods update every subblock, superior performance has been achieved compared to the block-wise estimation scheme in [2]. [The block-wise scheme for m_b=20 exhibits numerical ill-conditioning perhaps because the oversampled CE-BEM basis functions are not orthogonal.] Assuming no a priori models of the BEM coefficients, the two proposed subblock RLS tracking schemes have similar NCMSE and BER performances, and they both outperform the Kalman subblock tracking approach we proposed earlier in [1].

Finally, we consider the computational complexity of each scheme, obtained by simply calculating the CPU time during the channel estimation process of each simulation program. The results are displayed in Table 1, all based on the $m_b = 20$ case and averaged over 100 Monte Carlo runs.

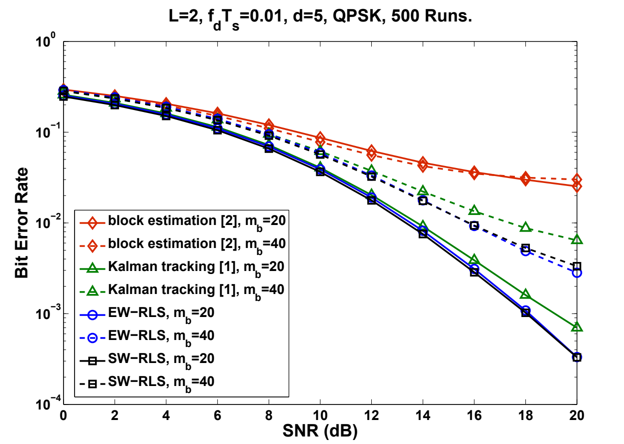

Fig. 2. *BER vs SNR, under $f_dT_s = 0.01$, $m_b = 20$ and 40, with QPSK information symbols.*

	CPU Time (s)
Block-wise Estimation [2]	0.026
Subblock-wise Kalman Tracking [1]	0.135
Subblock-wise EW-RLS Tracking	0.128
Subblock-wise SW-RLS Tracking	0.204

Table 1. *CPU times (seconds) per run for channel estimation.*

5. REFERENCES

[1] S. He and J.K. Tugnait, "Doubly-selective channel estimation using exponential basis models and subblock tracking," in *Proc. IEEE GLOBECOM '07*, Washington, DC, Nov. 26-30, 2007.

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

[3] D. Borah and B. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1863-1875, Nov. 1999.

[4] K.A.D. Teo, S. Ohno and T. Hinamoto, "Kalman channel estimation based on oversampled polynomial model for OFDM over doubly-selective channel," in *Proc. IEEE SPAWC*, New York, NY, June 2005, pp. 116-120.

[5] P. Banelli, R.C. Cannizzaro and L. Rugini, "Data-aided Kalman tracking for channel estimation in Doppler-affected OFDM systems," in *Proc. IEEE ICASSP 2007*, vol. 3, Honolulu, HI, April 2007, pp. 133-136.

[6] G. Leus, "On the estimation of rapidly time-varying channels," in *Proc. European Signal Process. Conf.*, Vienna, Austria, Sept. 6-10, 2004, pp. 2227-2230.

[7] T. Zemen and F. Mecklenbräuker, "Time-variant channel estimation using discrete prolate spheroidal sequences," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3597-3607, Sept. 2005.

[8] A.H. Sayed, *Fundamentals of Adaptive Filtering*. Wiley: New York, NJ, 2003.

[9] N. Al-Dhahir and J.M. Cioffi, "MMSE decision-feedback equalizers: finite-length results," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 961-975, July 1995.