EURASIP Journal on Applied Signal Processing 2002:8, 818–830

An Adaptive Channel Estimation Algorithm Using Time-Frequency Polynomial Model for OFDM with Fading Multipath Channels

Xiaowen Wang
Wireless Systems Research Department, Agere Systems, Murray Hill, NJ 07974, USA
Email: xiaowenw@agere.com

K. J. Ray Liu
Electrical and Computer Engineering Department, University of Maryland, College Park, MD 20742, USA
Email: kjrliu@eng.umd.edu

Received 1 August 2001 and in revised form 7 March 2002

Orthogonal frequency division multiplexing (OFDM) is an effective technique for the future 3G communications because of its great immunity to impulse noise and intersymbol interference. The channel estimation is a crucial aspect in the design of OFDM systems. In this work, we propose a channel estimation algorithm based on a time-frequency polynomial model of the fading multipath channels. The algorithm exploits the correlation of the channel responses in both time and frequency domains and hence reduce more noise than the methods using only time or frequency polynomial model. The estimator is also more robust compared to the existing methods based on Fourier transform. The simulation shows that it has more than 5 dB improvement in terms of mean-squared estimation error under some practical channel conditions. The algorithm needs little prior knowledge about the delay and fading properties of the channel. The algorithm can be implemented recursively and can adjust itself to follow the variation of the channel statistics.

Keywords and phrases: channel estimation, OFDM, polynomial approximation.

1. INTRODUCTION

The 3G wireless communication system is the next generation mobile cellular system that aims to provide high rate data communications of a bit rate up to 2 Mbit/s. Among many technical challenges in this broadband system, the severe intersymbol interference (ISI) caused by multipath effect of wireless channels is an essential one. One effective technique to deal with this problem is the orthogonal frequency division multiplexing (OFDM) [1, 2]. In OFDM systems, the entire bandwidth is partitioned into parallel subchannels by dividing the transmit data into several parallel low bit rate data streams to modulate the carriers corresponding to those subchannels. By doing so, the OFDM system has a relatively longer symbol duration, thus provides a great resistance to ISI and impulse noise. When the number of subchannels is large enough, the subchannels can be treated as independent of each other and only a one-tap equalizer is needed for each subchannel. Because of these advantages, OFDM has become a promising technique for broadband wireless communications.

Channel estimation is a key issue in a communication system, as is the case for the OFDM system. Without the knowledge of channel information, noncoherent detection, such as differential modulation, has to be used and results in some performance loss compared to the coherent detection. The channel estimation problem becomes more important for the 3G systems because many sophisticated signal processing techniques that require the knowledge of the channel information are expected to be used to meet the challenge of throughput and performance. For example, the independence of the subchannels in OFDM systems provides an easy way to optimize the transmitter design by adjusting the bit rate and transmit power across subchannels according to their channel conditions [3], which implies that the channel information has to be known at the transmitter.

The channel estimation problem is also more challenging in the 3G system because both the multipath effect and the fading effect have to be considered in this mobile broadband system. The important observation to solve the channel estimation problem in the OFDM systems is that the fading multipath channel in the 3G system is correlated in both time and frequency domain, even though subchannels are treated independently when performing the signal detection. The channel estimation algorithms should exploit such correlation to improve the accuracy of the estimation. Van de Beek et al. [4] tried to exploit the correlation of the channel parameters in frequency domain while Mignone and Morello [5] used the correlation in time domain. Li et al. [6] considered the correlation in both time and frequency domains. The estimators designed in these literatures are all Fourier-transform-based approaches, which implicitly assumed that the channel power spectrum can be viewed as band limited. The assumption is true when we consider the ensemble statistics. However, in practice, we can only get finite discrete samples of the channel response of the time varying channel. The leakage can be very severe and then degrade the performance dramatically.

In this work, we consider the problem from another point of view. Because of the correlation of the fading multipath channel, it can be viewed as a smoothly varying function of both time and frequency. It has been stated in the approximation theory that such a smoothly varying function can be approximated by a set of basis functions [7], for example, the polynomial basis [8]. Borah and Hart [9, 10] used the time domain polynomial approximation while Luise et al. in [11] used the frequency domain polynomial approximation. However, the channel responses used in coherent detection of OFDM are located in the time-frequency plane. Therefore, it is naturally to exploit the channel correlation in both time and frequency domains using a time-frequency polynomial model. The noise can then be greatly suppressed by estimating a smaller number of coefficients of the basis functions over a large number of observations. Moreover, it also make the estimator design more flexible and robust to the variation of channel statistics. We can also view Fourier transform as a type of model basis function and hence Fourier-transform-based method is the same type of method as the polynomial-model-based channel estimation scheme but with different model accuracy and different noise reduction capability. These two methods compared, the model error of the Fourier basis is very sensitive to the channel statistics and works only for some very specific system parameters and channel statistics. On the contrary, the polynomial-model-based method performs more consistently and robustly for variety of channels.

A key problem in using the polynomial model to estimate the channel responses is to decide the model order and time-frequency window dimensions of observations. The model approximation error of polynomial model decreases when increasing model order or decreasing the window dimensions. On the other hand, the noise is reduced more when decreasing the model order or increasing the window dimensions. It is important to reach a tradeoff between the model error and noise reduction. In this paper, we propose an adaptive algorithm that adjusts the window dimensions to balance the tradeoff. With this adaptive algorithm, the channel correlation function or the fading and delay characteristics are no longer that essential in the design of the channel estimator. The estimator can adapt its settings to the variation of the channel statistics.

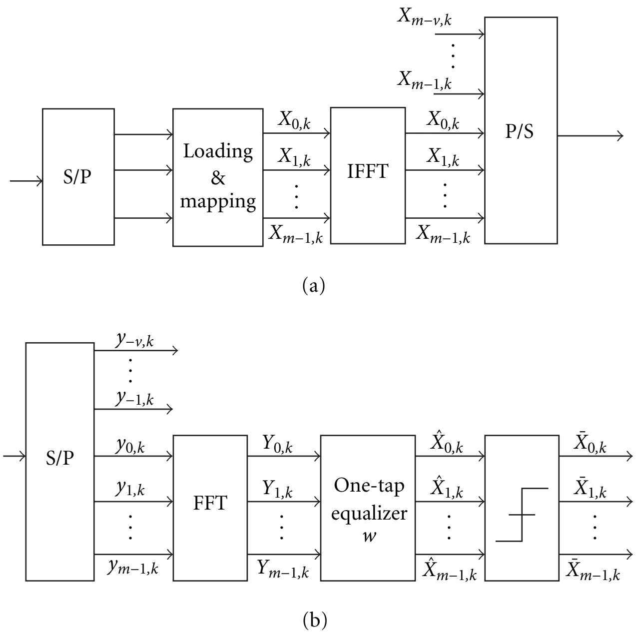

Figure 1: OFDM transmitter and receiver. (a) Transmitter, (b) receiver.

The rest of the paper is organized as follows. First, we introduce the OFDM system in Section 2 and the fading multipath channel in Section 3.1. Then, we discuss the time-frequency polynomial model Section 3.2 and derive the corresponding recursive channel estimation algorithm in Section 4. The performance analysis is discussed on the general-model-based estimation approach in Section 5. Then the window dimension adaptive algorithm is derived in Section 6 based on the performance analysis. Finally, the simulation results are presented to demonstrate the performance in Section 7 and the conclusion is drawn in Section 8.

2. OFDM SYSTEMS

Figures 1a and 1b show the transmitter and receiver of an OFDM system, respectively. The OFDM system divides the whole bandwidth B_d into m subchannels by buffering the input data to blocks, and then partitions the block into m lower rate bit streams. In most of OFDM systems, the subchannels are divided evenly, the bandwidth of the subchannels or the rate of the bit streams is $\Delta f = B_d/m$. The bit streams may contain different amount of bits and use different transmit energy according to the channel condition. The bit and energy allocation is done by a loading algorithm. Then the bit streams are mapped to some complex constellation points $X_{i,k}$, $i = 0, \ldots, m-1$ at the kth block. The modulation is then implemented by m-point inverse discrete Fourier transform (IDFT). Then the modulated data go through P/S converter to form the serial data $x_{i,k}$. A cyclic prefix which is constructed using the last v samples of $x_{i,k}$'s is inserted before sending the $x_{i,k}$'s to the channel. Now it follows that the symbol duration is m/B_d, however, the actual block duration is

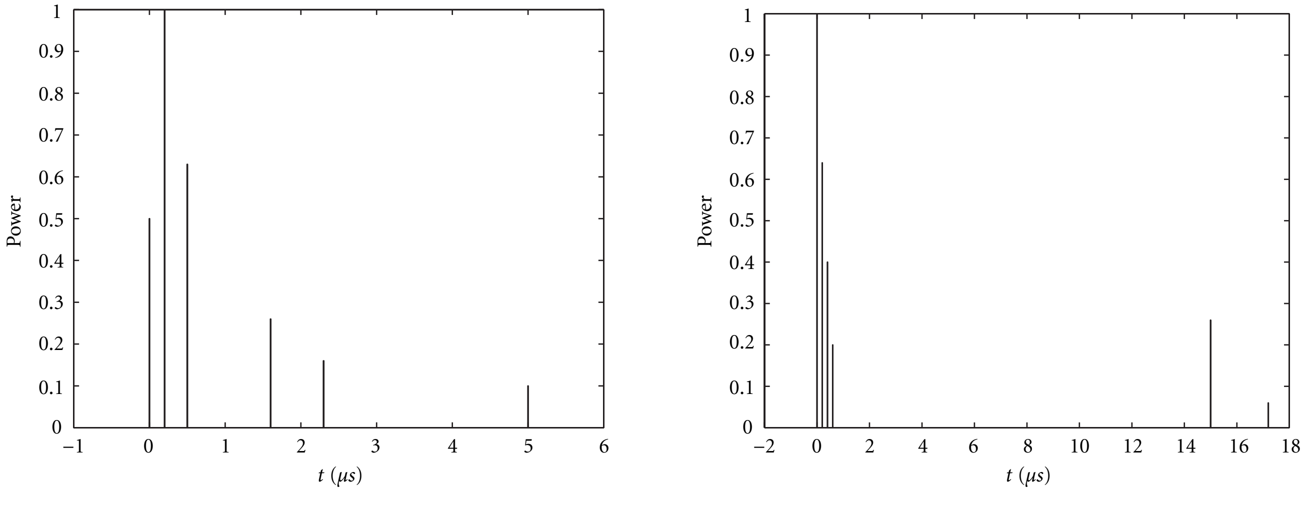

FIGURE 2: Two typical delay profiles (a) TU and (b) HT.

$T_f = (m+\nu)/B_d$. For a system with $B_d = 800$ kHz, $m = 128$, and $\nu = 16$, the block duration is $T_f = 180$ microseconds. Such a system is used in the rest of this paper.

At the receiver, the prefix part is discarded. The demodulation is performed by the discrete Fourier transform (DFT) operation. If the cyclic prefix is long enough, then the interference between two OFDM blocks is eliminated and the subchannels can be viewed as independent of each other, that is, the demodulated data $Y_{i,k}$ can be expressed as

$$Y_{i,k} = H_{i,k}X_{i,k} + N_{i,k}, \tag{1}$$

where $H_{i,k}$ is the channel frequency response at $i\Delta f$ of kth block and $N_{i,k}$ is the corresponding channel noise that is assumed to be white Gaussian process with zero mean and variance σ^2.

Because of the simple relation of (1), only a one-tap equalizer is needed for each subchannel at the receiver, that is,

$$\hat{X}_{i,k} = Y_{i,k}W_{i,k}, \tag{2}$$

where the equalizer coefficient $W_{i,k}$ is some function of $H_{i,k}$. For example, the zero-forcing equalizer is constructed as $W_{i,k} = 1/H_{i,k}$. Then the decision or decoding is made upon $\hat{X}_{i,k}$.

3. POLYNOMIAL CHANNEL MODEL

3.1. Fading multipath channel

In a mobile broadband wireless communication system such as 3G, the transmission is impaired by both fading that is due to the mobility, and multipath that is due to the wide bandwidth. This fading multipath channel has long been known to be modeled as a time-varying linear filter [12],

$$h(t,\tau) = \sum_i \gamma_i(t)\delta(\tau - \tau_i), \tag{3}$$

where $\gamma_i(t)$'s are independent complex Gaussian processes with zero mean and variance p_i's. For OFDM systems, we can assume that the channel is time varying for different blocks but time-invariant within one block. The channel frequency response $H_{i,k}$'s are samples of the continuous channel response $H(t,f) = \int h(t,\tau)e^{-j2\pi f\tau}\,d\tau$, that is,

$$H_{i,k} = H(kT_f, i\Delta f). \tag{4}$$

The correlation function of $H(t,f)$ is defined as $r_H(t,f) \triangleq \mathrm{E}[H(t_1,f_1)H^*(t_1-t,f_1-f)]$. Assume that the correlation function of $\gamma_i(t)$ follows $\mathrm{E}[\gamma_i(t_1)\gamma_i^*(t_1-t)] = p_i r(t)$, then we have

$$r_H(t,f) = r_t(t)r_f(f). \tag{5}$$

For the Rayleigh fading channel [12], $r_t(t) = J_0(2\pi f_D t)$ and $r_f(f) = \sum_i p_i e^{-j2\pi\Delta f\tau_i}$ with $J_0(\cdot)$ denoting the zero-order Bessel function, f_D being the *Doppler frequency* describing the channel variation along t, and p_i's together with τ_i's being delay profiles describing the channel dispersion which is also often characterized by the *maximum delay spread* $T_d \triangleq \max_i \tau_i$. Three types of delay profiles are used in this work, TU, HT, and 2-ray [13]. The TU and HT delay profiles are shown in Figure 2. The 2-ray delay profile has two equal power paths and the delay between two paths is T_d. We also assume that the channel is normalized in our simulation, that is, $\sum_i p_i = 1$.

From (5), the power spectrum of the channel response is

$$S_H(\xi,\nu) = \int\int r_H(t,f)e^{-j(t\xi+f\nu)}\,d\xi\,d\nu = S_t(\xi)S_f(\nu), \tag{6}$$

where $S_t(\xi) = \int r_t(t)e^{-jt\xi}\,d\xi$ and $S_f(\nu) = \int r_f(f)e^{-jf\nu}\,d\nu$. Because of the physical mechanism of the propagation, the channel varies smoothly and most of the energy is concentrated in a finite bandwidth in both time and frequency domains. The bandwidths are f_D for $S_t(\xi)$ and T_d for $S_f(\nu)$, respectively.

3.2. Time-frequency polynomial channel model of OFDM systems

We know from the approximation theory [7, 8] that the smoothly varying channel responses can be approximated by projecting to a finite set of basis functions. In [14], it was shown that the channel responses in a small time domain window around a center point k_0 of dimension $2K+1$ can be closely approximated by a small set of polynomial basis functions, that is,

$$H_{i,k} = \sum_{m=0}^{M-1} H_{i,k_0}(m)(k-k_0)^m + R_M, \quad \text{for } k_0 - K \le k \le k_0 + K, \tag{7}$$

where

$$\begin{aligned} H_{i,k_0}(m) &= \frac{T_f^m}{m!}\frac{\partial^m H(t,f)}{\partial t^m}\bigg|_{t=k_0T_f}, \\ R_M &= \frac{((k-k_0)T_f)^M}{M!}\frac{\partial^M H(t,f)}{\partial t^M}\bigg|_{t=t'} \end{aligned} \tag{8}$$

with $k_0T_f \le t' \le kT_f$.

For such an approximation, it can be proved that the mean-squared model error is bounded by

$$\begin{aligned} \mathrm{E}\left[\|R_M\|^2\right] &\le \left(\frac{((k-k_0)T_f)^M}{M!}\right)^2 \int_0^{f_D} (2\pi\xi)^{2M} S_t(\xi)\,d\xi \\ &\le f_D\left(\frac{2\pi(k-k_0)f_DT_f}{M!}\right)^{2M}. \end{aligned} \tag{9}$$

Here we assume that $\int_0^{f_D} S_t(\xi)\,d\xi = 1$. It can be seen from (10) that the sufficient condition for this error to converge to zero is $f_DT_f \ll 1$, that is, if $f_DT_f \ll 1$ then $\lim_{M\to\infty} \mathrm{E}[\|R_M\|^2] = 0$.

Similarly, if the channel delay spread T_d satisfies $T_d\Delta f \ll 1$, which means that the frequency variation of $H_{i,k}$'s is smooth enough along frequencies, then $H_{i,k}$'s in a frequency domain window of dimension $2I+1$ around i_0, $[i_0-I, i_0+I]$, can be approximated by the polynomial bases, that is,

$$H_{i,k} = \sum_{n=0}^{N-1} H_{i_0,k}(n)(i-i_0)^n + R_N, \quad \text{for } i_0 - I \le i \le i_0 + I, \tag{10}$$

where

$$\begin{aligned} H_{i_0,k}(n) &= \frac{\Delta f^n}{n!}\frac{\partial^n H(t,f)}{\partial f^n}\bigg|_{f=i_0\Delta f}, \\ R_N &= \frac{((i-i_0)\Delta f)^N}{N!}\frac{\partial^N H(t,f)}{\partial f^N}\bigg|_{f=f'} \end{aligned} \tag{11}$$

with $i_0\Delta f \le f' \le i\Delta f$.

The mean-squared model error of this approximation is bounded by

$$\mathrm{E}\left[\|R_N\|^2\right] \le \left(\frac{((i-i_0)\Delta f)^N}{N!}\right)^2 \int_0^{T_d} (2\pi\nu)^N S_f(\nu)\,d\nu. \tag{12}$$

The time domain expansion (7) is used for channel estimation in [10], while the frequency domain expansion (10) is adopted in [11]. For a given channel, the selection of the above two types of expansions depends on the channel statistics, f_D and T_d, and the system parameters, T_f and Δf. Moreover, it is naturally to expand the channel responses in both time and frequency domains [15] for the OFDM system, since its signal is distributed in a time frequency plane. The expansion in the time-frequency window of dimensions $(2I+1)\times(2K+1)$ around i_0 and k_0 is

$$H_{i,k} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} H_{i_0,k_0}(nm)(k-k_0)^m(i-i_0)^n + R_{MN}, \tag{13}$$

for $k_0 - K \le k \le k_0 + K$ and $i_0 - I \le i \le i_0 + I$, where

$$\begin{aligned} H_{i_0,k_0}(nm) &= \frac{T_f^m \Delta f^n}{m!n!}\frac{\partial^m\partial^n H(t,f)}{\partial t^m \partial f^n}\bigg|_{t=k_0T_f,\, f=i_0\Delta f}, \\ R_{MN} & \\ &= R_M + R_N \\ &\quad - \frac{((k-k_0)T_f)^M((i-i_0)\Delta f)^N}{M!N!}\frac{\partial^M\partial^N H(t,f)}{\partial t^M \partial f^N}\bigg|_{t=t',\, f=f'} \end{aligned} \tag{14}$$

with $k_0T_f \le t' \le kT_f$ and $i_0\Delta f \le f' \le i\Delta f$.

The mean-squared model error is then bounded by

$$\begin{aligned} \mathrm{E}\left[\|R_{MN}\|^2\right] \le{}& \left(\frac{(KT_f)^M}{M!}\right)^2 \int_0^{f_D} (2\pi\nu)^M S_t(\xi)\,d\xi \\ &+ \left(\frac{(I\Delta f)^N}{N!}\right)^2 \int_0^{T_d} (2\pi\nu)^N S_f(\nu)\,d\nu \\ &+ \left(\frac{(KT_f)^M}{M!}\right)^2 \left(\frac{(I\Delta f)^N}{N!}\right)^2 \\ &\times \int_0^{f_D}\int_0^{T_d} (2\pi\xi)^M (2\pi\nu)^N S_t(\xi) S_f(\nu)\,d\xi\,d\nu. \end{aligned} \tag{15}$$

Without loss of generality, assuming $M = N$ and using the multipath Rayleigh fading channel described in Section 3.1, we can show that [15]

$$\begin{aligned} \mathrm{E}\left[\|R_{MM}\|^2\right] \le{}& \frac{2M!(2\pi KT_f f_D)^{2M}}{2^{2M}(M!)^4} + \frac{(2\pi I\Delta f T_d)^{2M}}{(M!)^2} \\ &+ \frac{2M!(4\pi^2 KIT_f\Delta f f_D T_d)^{2M}}{2^{2M}(M!)^6}. \end{aligned} \tag{16}$$

Again the sufficient conditions for convergence of the above expansion are $f_DT_f \ll 1$ and $T_d\Delta f \ll 1$. It is noticed that these two conditions are usually satisfied in a practical

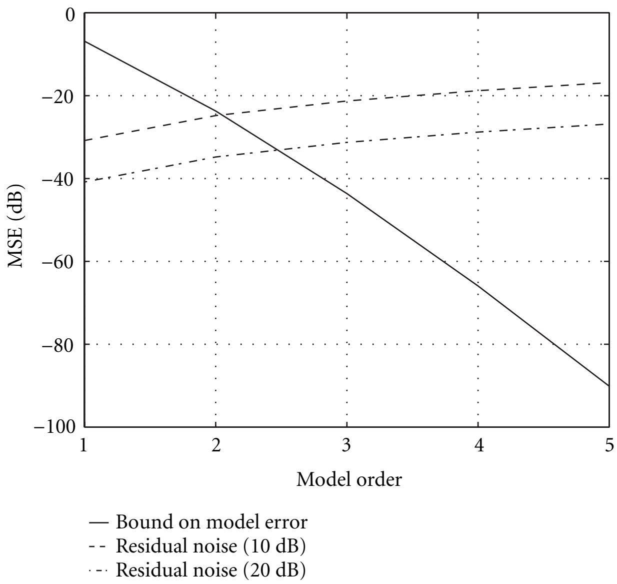

Figure 3: Bound on mean-squared model error.

OFDM systems. When $f_D T_f > 1$, the block duration is so long or the channel changes are so fast that the channel cannot be viewed as invariant during one block and the system suffers large interchannel interference (ICI). On the other hand, when $T_d \Delta f > 1$, the block duration is so short or the channel dispersion is so large that the subchannels can no longer be treated independently and the system would suffer both ISI and ICI. The OFDM system cannot work in either case. Hence, it is reasonable to assume that both conditions are satisfied in a well designed OFDM system.

Now we take a close look at the upper bound of the model error. Suppose that the length of the cyclic prefix v can be ignored compared to the number of the subchannels. Then the first term in (16) is determined by $f_D T_f = m f_D / B_d$, while the second term is determined by $\Delta f T_d = B_d T_d / m$. The third term is actually determined by $f_D T_d$ and is much smaller than the first two terms, since they both are smaller than one. For the first two terms, when m is large, the first term is dominating, then we should choose smaller K or larger M. While m is small, then the second term is dominating and I should be smaller or the model order N should be larger. If the Doppler frequency f_D, maximum delay T_d, and bandwidth B_d are fixed, we can adjust the window dimensions according to m to keep the time-frequency model error to certain level but we still have a small MN/IK. However, if only time or frequency domain expansion is used, the model error cannot be adjusted to maintain a small level with the same M/K or N/I when the number of subchannels m varies.

Figure 3 shows the upper bound of mean the squared model error with $f_D T_f = T_d \Delta f = 10^{-2}$ and $I = K = 5$ according to model order M. It shows that the model error is under -40 dB as the model order is 3. This means that we only need to estimate 9 model coefficients to get the 121 channel responses. In this figure, we also show the residual noise for SNR of 10 dB and 20 dB. It shows that the noise can be greatly reduced with very small penalty on model error. Moreover, such a model approximation does not need to know the actual channel correlation function.

4. CHANNEL ESTIMATION ALGORITHM WITH POLYNOMIAL MODEL

4.1. Estimator structure

The channel estimation problem in OFDM systems is to estimate the channel response $H_{i,k}$ based on the transmitted signal $X_{i,k}$ and the received signal $Y_{i,k}$. The information of the transmitted signal $X_{i,k}$'s is obtained either from training or from detected feedback. In OFDM systems, an instantaneous estimate can be easily constructed as $\tilde{H}_{i,k} = Y_{i,k}/X_{i,k}$. Then suppose that we have chosen the model order and window dimensions such that the model error is small and can be ignored, we can approximate (13) in a matrix form

$$\mathbf{H}_{i_0,k_0} \simeq \mathbf{Q}_{M,N}(I,K)\mathbf{b}_{i_0,k_0}, \tag{17}$$

where

$$\mathbf{H}_{i_0,k_0} = \left[H_{-I+i_0,-K+k_0} \cdots H_{I+i_0,-K+k_0} \cdots H_{-I+i_0,K+k_0} \cdots H_{I+i_0,K+k_0}\right]^T,$$

$$\mathbf{b}_{i_0,k_0} = \left[H_{i_0,k_0}(0,0) \cdots H_{i_0,k_0}(N-1,0) \cdots H_{i_0,k_0}(0,M-1) \cdots H_{i_0,k_0}(N-1,M-1)\right]^T,$$

$$\mathbf{Q}_{M,N}(I,K) = \begin{bmatrix} q_{0,0}^{-I,-K} & \cdots & q_{0,N-1}^{-I,-K} & \cdots & q_{M-1,0}^{-I,-K} & \cdots & q_{M-1,N-1}^{-I,-K} \\ \vdots & & \vdots & & \vdots & & \vdots \\ q_{0,0}^{I,-K} & \cdots & q_{0,N-1}^{I,-K} & \cdots & q_{M-1,0}^{I,-K} & \cdots & q_{M-1,N-1}^{I,-K} \\ \vdots & & \vdots & & \vdots & & \vdots \\ q_{00}^{I,K} & \cdots & q_{0,N-1}^{I,K} & \cdots & q_{M-1,0}^{I,K} & \cdots & q_{M-1,N-1}^{I,K} \end{bmatrix}, \tag{18}$$

with $q_{m,n}^{i,k} = i^n k^m$, for $i = -I, \ldots, 0, \ldots, I$, $k = -K, \ldots, 0, \ldots, K$, $m = 0, \ldots, M-1$, and $n = 0, \ldots, N-1$.

Define

$$\tilde{\mathbf{H}}_{i_0,k_0} = \left[\tilde{H}_{-I+i_0,-K+k_0} \cdots \tilde{H}_{I+i_0,-K+k_0} \cdots \tilde{H}_{-I+i_0,K+k_0} \cdots \tilde{H}_{I+i_0,K+k_0}\right]^T, \tag{19}$$

then

$$\begin{aligned} \tilde{\mathbf{H}}_{i_0,k_0} &= \mathbf{H}_{i_0,k_0} + \mathbf{N}_{i_0,k_0} \\ &\simeq \mathbf{Q}_{M,N}(I,K)\mathbf{b}_{i_0,k_0} + \mathbf{N}_{i_0,k_0}, \end{aligned} \tag{20}$$

where

$$\mathbf{N}_{i_0,k_0} = \left[\frac{N_{-I+i_0,-K+k_0}}{X_{-I+i_0,-K+k_0}} \cdots \frac{N_{I+i_0,-K+k_0}}{X_{I+i_0,-K+k_0}} \cdots \frac{N_{-I+i_0,K+k_0}}{X_{-I+i_0,K+k_0}} \cdots \frac{N_{I+i_0,K+k_0}}{X_{I+i_0,K+k_0}}\right]^T. \tag{21}$$

Using least square (LS) methods, we can get the estimation of the coefficients of the polynomial basis from the instantaneous estimates

$$\hat{\mathbf{b}}_{i_0,k_0} = \mathbf{Q}^{\dagger}_{M,N}(I,K)\tilde{\mathbf{H}}_{i_0,k_0}, \tag{22}$$

where $\mathbf{Q}^{\dagger}_{M,N}(I,K)$ is the pseudoinverse of $\mathbf{Q}_{M,N}(I,K)$. The channel estimation then can be constructed as

$$\begin{aligned}\hat{H}_{i,k} &= \mathbf{q}_{M,N}(i-i_0,k-k_0)^T\hat{\mathbf{b}}_{i_0,k_0}\\ &= \mathbf{q}_{M,N}(i-i_0,k-k_0)^T\mathbf{Q}^{\dagger}(I,K)\tilde{\mathbf{H}}_{i_0,k_0},\end{aligned} \tag{23}$$

where

$$\begin{aligned}&\mathbf{q}_{M,N}(i-i_0,k-k_0)\\ &= \left[q_{0,0}^{i-i_0,k-k_0} \cdots q_{0,N-1}^{i-i_0,k-k_0} \cdots q_{M-1,0}^{i-i_0,k-k_0} \cdots q_{M-1,N-1}^{i-i_0,k-k_0}\right]^T.\end{aligned} \tag{24}$$

Usually, we fix the value of $i-i_0$ and $k-k_0$, that is, we fix the point of estimation inside the window and slide the window to get all the estimations. Then the estimator can be viewed as a two-dimensional filtering process. Arranging the instantaneous estimation inside the window into a matrix form,

$$\tilde{\mathbf{H}} = \begin{bmatrix}\tilde{H}_{-I+i_0,-K+k_0} & \cdots & \tilde{H}_{-I+i,K+k_0}\\ \vdots & & \vdots\\ \tilde{H}_{I+i_0,-K+k_0} & \cdots & \tilde{H}_{I+i_0,K+k_0}\end{bmatrix}. \tag{25}$$

Then the estimation is

$$\hat{H}_{i,k} = \mathbf{q}_N^T(i-i_0)\mathbf{Q}_N^{\dagger T}(I)\tilde{\mathbf{H}}\mathbf{Q}_M^{\dagger}(K)\mathbf{q}_M(k-k_0), \tag{26}$$

where

$$\begin{aligned}\mathbf{q}_N(i) &= \left[i^0 \;\; i^1 \;\; \cdots \;\; i^{N-1}\right]^T,\\ \mathbf{Q}_N(I) &= \left[\mathbf{q}_N(-I) \;\; \cdots \;\; \mathbf{q}_N(0) \;\; \cdots \;\; \mathbf{q}_N(I)\right]^T.\end{aligned} \tag{27}$$

The estimator structure is shown in Figure 4. The coefficients of the frequency domain filter are $\mathbf{Q}_N^{\dagger}(I)\mathbf{q}_N(i-i_0)$ and the coefficients of the time domain filter are $\mathbf{Q}_M^{\dagger}(K)\mathbf{q}_M(k-k_0)$.

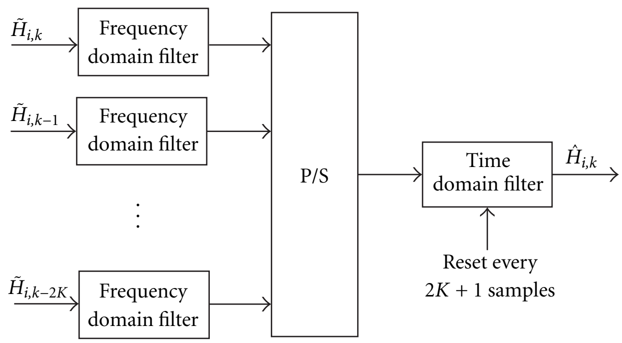

Figure 4: The estimator structure.

4.2. *Recursive algorithms*

The two-dimensional filter can actually be implemented recursively in time and frequency domains, respectively.

Define the basis functions $\mathbf{Q}'_N(I)$ as

$$\mathbf{Q}'_N(I) = \left[\mathbf{q}_N(-I+1) \;\; \cdots \;\; \mathbf{q}_N(0) \;\; \cdots \;\; \mathbf{q}_N(I) \;\; \mathbf{q}_N(I+1)\right]^T. \tag{28}$$

The basis $\mathbf{Q}'_N(I)$ and $\mathbf{Q}_N(I)$ are actually homomorphic to each other, that is, there is an invertible matrix $\mathbf{R}$ such that

$$\mathbf{Q}'_N(I) = \mathbf{Q}_N(I)\mathbf{R}. \tag{29}$$

This means that instead of using $\mathbf{Q}_N(I)$ as basis, we can use $\mathbf{Q}'_N(I)$ as the basis to construct the estimator, that is,

$$\mathbf{q}_N^T(i+1-i_0)\mathbf{Q}'^{\dagger}_N(I) = \mathbf{q}_N^T(i-i_0)\mathbf{Q}_N^{\dagger}(I). \tag{30}$$

Substitute (30) into (26), we can estimate the channel using $\mathbf{Q}'^{\dagger}_n(I)$. Then the core of the recursive algorithm is to calculate $\mathbf{Q}'^{\dagger}_N(I)$ from $\mathbf{Q}_N^{\dagger}(I)$ iteratively.

Let $\mathbf{P}_f = (\mathbf{Q}_N^T(I)\mathbf{Q}_N(I))^{-1}$ and $\mathbf{P}_t = (\mathbf{Q}_M^T(K)\mathbf{Q}_M(K))^{-1}$. At initialization, we estimate model coefficients $\hat{\mathbf{b}}_f(k)$ or $\hat{\mathbf{b}}_t(i)$ regarding to $\mathbf{P}_f$ or $\mathbf{P}_t$ over a window of dimension $2I+1$ or $2K+1$. Similar to the recursive least square (RLS) algorithm, using the *matrix inverse lemma* [16], we can calculate $\mathbf{P}_f^+ = (\mathbf{Q}'^T_N(I+1)\mathbf{Q}'_N(I+1))^{-1}$ or $\mathbf{P}_t^+ = (\mathbf{Q}'^T_M(K+1)\mathbf{Q}'_M(K+1))^{-1}$ and then the corresponding model coefficients $\hat{\mathbf{b}}_f^+$ or $\hat{\mathbf{b}}_t^+$ recursively over the window of dimension $2I+2$ or $2K+2$ during the *updating* process. After that apply the *matrix inverse lemma* again, we can calculate $\mathbf{P}_f^- = (\mathbf{Q}'^T_N(I)\mathbf{Q}'_N(I))^{-1}$ or $\mathbf{P}_t^- = (\mathbf{Q}'^T_M(K)\mathbf{Q}'_M(K))^{-1}$ and the corresponding model coefficients $\hat{\mathbf{b}}_f^-$ or $\hat{\mathbf{b}}_t^-$ from $\mathbf{P}_f^+$ or $\mathbf{P}_t^+$ over the window of dimension $2I+1$ or $2K+1$ during the *downdating* process. Then according to (30), the channel can be estimated as

$$\begin{aligned}\hat{H}_{i+1,k} &= \mathbf{q}_N^T(i+1-i_0)\hat{\mathbf{b}}_f^-,\\ \hat{H}_{i,k+1} &= \mathbf{q}_M^T(k+1-k_0)\hat{\mathbf{b}}_t^-.\end{aligned} \tag{31}$$

As this recursive process going on, the basis function becomes $\mathbf{q}_M^T(k+l-k_0)$ where l is the index of the iteration. The dynamic range of such a basis function may become so large that it will affect the numerical stability of the algorithm. Therefore, regularization using $\mathbf{R}_f$ or $\mathbf{R}_t$ should be used periodically to scale the basis back to $\mathbf{Q}_N(I)$ or $\mathbf{Q}_M(K)$. The frequency domain and time domain recursive algorithms are summarized in Algorithms 1 and 2, respectively. The matrices $\mathbf{K}_f^+$ and $\mathbf{K}_f^-$ or $\mathbf{K}_t^+$ and $\mathbf{K}_t^-$ are the corresponding *gain matrices* in updating and downdating. The two-dimensional filter in the tables is implemented first by frequency domain filtering then by time domain filtering. The order can be switched. In that way, the input in Algorithm 2 are instantaneous estimates while the inputs in Algorithm 1 are the outputs of the time domain filters of Algorithm 2. It is also noted that the order of downdating and updating can be switched, too.

In both tables, $\mathbf{K}_f^+$, $\mathbf{K}_f^-$, $\mathbf{P}_f^+$, $\mathbf{P}_f^-$ and $\mathbf{K}_t^+$, $\mathbf{K}_t^-$, $\mathbf{P}_t^+$, $\mathbf{P}_t^-$ can be

Initialization:
with temporary estimation $\tilde{\mathbf{H}}_k = [\tilde{H}_{-I+i_0,k} \ \cdots \ \tilde{H}_{I+i_0,k}]$, calculate

$$\hat{\mathbf{b}}_f(k) = \mathbf{Q}_N^{\dagger}(I)\tilde{\mathbf{H}}_k,$$

$$\mathbf{P}_f = \left(\mathbf{Q}_N^T(I)\mathbf{Q}_N(I)\right)^{-1}.$$

Updating:
with the new input $\tilde{H}_{I+i_0+1,k}$, calculate

$$\mathbf{K}_f^+ = \mathbf{I} - \frac{\mathbf{P}_f\mathbf{q}_N(I+1)\mathbf{q}_N^T(I+1)}{1+\mathbf{q}_N^T(I+1)\mathbf{P}\mathbf{q}_N(I+1)},$$

$$\mathbf{P}_f^+ = \mathbf{K}_f^+\mathbf{P}_f.$$

$$\Delta\mathbf{b}_f^+ = \mathbf{q}_N(I+1)\tilde{H}_{I+i_0+1,k},$$

$$\mathbf{b}_f^+ = \mathbf{K}_f^+(\hat{\mathbf{b}}_f(k) + \Delta\mathbf{b}_f^+).$$

Downdating:

$$\mathbf{K}_f^- = \mathbf{I} + \frac{\mathbf{P}_f^+\mathbf{q}_N(-I)\mathbf{q}_N^T(-I)}{1-\mathbf{q}_N^T(-I)\mathbf{P}_f^+\mathbf{q}_N(-I)},$$

$$\mathbf{P}_f^- = \mathbf{K}_f^-\mathbf{P}_f^+.$$

$$\Delta\mathbf{b}_f^- = \mathbf{q}_N(-I)\tilde{H}_{-I+i_0,k},$$

$$\mathbf{b}_f^- = \mathbf{K}_f^-(\mathbf{b}_f^+ - \Delta\mathbf{b}_f^-).$$

Regularization:

$$\hat{\mathbf{b}}_f(k) = \mathbf{R}_f\mathbf{b}_f^-,$$

$$\hat{H}_{i+1,k} = \mathbf{q}_N^T(i-i_0)\hat{\mathbf{b}}_f(k).$$

Algorithm 1: Frequency domain recursive algorithm.

Initialization:
with frequency domain filter results
$\hat{\mathbf{H}}_f = [\hat{H}(i+1,-K+k_0) \ \cdots \ \hat{H}(i+1,K+k_0)]$, calculate

$$\hat{\mathbf{b}}_t(i+1) = \mathbf{Q}_M^{\dagger}(K)\hat{\mathbf{H}}_f,$$

$$\mathbf{P}_t = \left(\mathbf{Q}_M^T(K)\mathbf{Q}_M(K)\right)^{-1}.$$

Updating:
with the new input $\hat{H}(i+1,K+k_0+1)$, calculate

$$\mathbf{K}_t^+ = \mathbf{I} - \frac{\mathbf{P}_t\mathbf{q}_M(K+1)\mathbf{q}_M^T(K+1)}{1+\mathbf{q}_M^T(K+1)\mathbf{P}_t\mathbf{q}_M(K+1)},$$

$$\mathbf{P}_t^+ = \mathbf{K}_t^+\mathbf{P}_t.$$

$$\Delta\mathbf{b}_t^+ = \mathbf{q}_M(K+1)\hat{H}(i+1,K+k_0+1),$$

$$\mathbf{b}_t^+ = \mathbf{K}_t^+(\hat{\mathbf{b}}_t(i+1) + \Delta\mathbf{b}_t^+).$$

Downdating:

$$\mathbf{K}_t^- = \mathbf{I} + \frac{\mathbf{P}_t^+\mathbf{q}_M(-K)\mathbf{q}_M^T(-K)}{1-\mathbf{q}_M^T(-K)\mathbf{P}_t^+\mathbf{q}_M(-K)},$$

$$\mathbf{P}_t^- = \mathbf{K}_t^-\mathbf{P}_t^+.$$

$$\Delta\mathbf{b}_t^- = \mathbf{q}_M(-K)\hat{H}(i+1,-K+k_0).$$

$$\mathbf{b}_t^- = \mathbf{K}_t^-(\mathbf{b}_t^+ - \Delta\mathbf{b}_t^-).$$

Regularization:

$$\hat{\mathbf{b}}_t = \mathbf{R}_t\mathbf{b}_t^-,$$

$$\hat{H}_{i+1,k+1} = \mathbf{q}_M^T(k-k_0)\hat{\mathbf{b}}_t(i+1).$$

Algorithm 2: Time domain recursive algorithm.

calculated off-line and do not change if the model order and window dimensions do not change. However, we still put the calculations inside the updating and downdating process in case that the window dimensions may change as what happened in the adaptive algorithm described in Section 6.

The recursive algorithm needs less calculation compared to direct computation of the product of pseudoinverse when the window dimensions are much larger than the model order. Many fast algorithms of recursive least square (RLS) can be used for the practical implementation of such a recursive algorithm [16]. It also provides an easy way to adjust the window dimensions for the implementation of the adaptive algorithm.

5. PERFORMANCE ANALYSIS

Suppose that the channel can be modeled by some basis function, that is, a set of channel responses $\mathbf{H}$ can be projected to a set of basis function $\mathbf{Q}$ and the coefficients of the basis functions are $\mathbf{b}$, that is,

$$\mathbf{H} = \mathbf{Q}\mathbf{b}. \tag{32}$$

The length of $\mathbf{H}$ is L and the length of $\mathbf{b}$ is l. In order to get an accurate channel estimation, we expect that $l \ll L$. This is true if the channel parameters in $\mathbf{H}$ is highly correlated. Given a set of noisy observations,

$$\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{N}, \tag{33}$$

the LS estimation of the coefficients is

$$\hat{\mathbf{b}} = \mathbf{Q}^{\dagger}\tilde{\mathbf{H}}. \tag{34}$$

The channel estimation is then

$$\hat{\mathbf{H}} = \mathbf{Q}\mathbf{Q}^{\dagger}\tilde{\mathbf{H}}. \tag{35}$$

Define the mean-squared estimation error matrix as

$$\varepsilon = \mathrm{E}\left[(\hat{\mathbf{H}} - \mathbf{H})(\hat{\mathbf{H}} - \mathbf{H})^H\right]. \tag{36}$$

We can show that

$$\varepsilon = (\mathbf{I} - \mathbf{Q}\mathbf{Q}^{\dagger})\mathbf{R}_H(\mathbf{I} - \mathbf{Q}\mathbf{Q}^{\dagger}) + \mathbf{Q}\mathbf{Q}^{\dagger}\mathbf{R}_N\mathbf{Q}\mathbf{Q}^{\dagger}, \tag{37}$$

where $\mathbf{R}_H = \mathrm{E}[\mathbf{H}\mathbf{H}^H]$ and $\mathbf{R}_N = \mathrm{E}[\mathbf{N}\mathbf{N}^H] = \sigma^2\mathbf{I}$ if the transmitted signals of all subchannels are all using the same constant envelop modulation and transmit energy of 1.

The estimation error consists of two parts; one is related to the model inaccuracy, that is,

$$\varepsilon_H = (\mathbf{I} - \mathbf{Q}\mathbf{Q}^{\dagger})\mathbf{R}_H(\mathbf{I} - \mathbf{Q}\mathbf{Q}^{\dagger}), \tag{38}$$

and the other is related to the residual noise, that is,

$$\varepsilon_N = \sigma^2 \mathbf{Q}\mathbf{Q}^{\dagger}. \tag{39}$$

Since $\mathbf{R}_H$ is a Toeplitz matrix, it can be decomposed as

$$\mathbf{R}_H = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^H \\ \mathbf{U}_2^H \end{bmatrix}, \tag{40}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with eigenvalues of $\mathbf{R}_H$ on its diagonal. If $\mathbf{Q} = \mathbf{U}_1$, then the model error is zero. This means that the optimal function basis, which we can find in terms of model accuracy, is the eigenbasis $\mathbf{U}_1$. However, it requires the knowledge about the statistics of the channel responses. In some special cases, we can easily find some specific function bases that can diagonalize $\mathbf{R}_H$ without actually knowing $\mathbf{R}_H$. For example, if $\mathbf{H}$ is the channel response for one OFDM block with all the delay paths, τ_i's, at the sampling instance of the OFDM system, then such an optimal function basis is the DFT matrix [6]. However, in most of practical situations, the channel delay profiles do not satisfy this condition. Therefore, using DFT matrix may cause severe leakage problem and incur a large model error.

The average energy of the residual noise over the entire estimation window can be calculated as follows:

$$\bar{e}_N = \frac{\sigma^2}{L} \operatorname{tr}\left[\mathbf{Q}\mathbf{Q}^{\dagger}\right] = \frac{l\sigma^2}{L}. \tag{41}$$

The average mean-squared error over the whole estimation window is actually lower bounded by (41). The lower bound is achieved when $\mathbf{Q} = \mathbf{U}_1$.

Although the average energy of the residual noise maintains the same once the data length and model order is fixed, the estimation error inside the window is often distributed unevenly and differently for different basis functions. For the polynomial model, the estimation error is the least at the center point of the window and larger at the edge. Therefore, we prefer to choose the center of the window to get a better performance. However, along the time domain, we can only choose the end point to get a causal filter.

6. OPTIMAL MODEL PARAMETERS ADAPTATION

With estimation point chosen at the center of the frequency domain window and end point at the time domain window, the estimation error from (23) becomes

$$\epsilon_{I,K} = \mathrm{E}\left[\left|\left|H_{i_0,k_0} - \hat{H}_{i_0,k_0}\right|\right|^2\right] = \epsilon_h + \epsilon_n, \tag{42}$$

where the model error is

$$\begin{aligned} \epsilon_h &= \mathrm{E}\left[\left|\left|H_{i_0,k_0} - \mathbf{q}_{M,N}(0,K)^T \mathbf{Q}_{M,N}^{\dagger}(I,K)\mathbf{H}_{i_0,k_0}\right|\right|^2\right] \\ &= r_H(0,0) - \mathrm{E}\left[H_{i_0,k_0}\mathbf{H}_{i_0,k_0}^T\right]\mathbf{Q}_{M,N}^{\dagger T}(I,K)\mathbf{q}_{M,N}(0,K) \\ &\quad - \mathbf{q}_{M,N}(0,K)^T\mathbf{Q}_{M,N}^{\dagger}(I,K)\,\mathrm{E}\left[\mathbf{H}_{i_0,k_0}H_{i_0,k_0}^*\right] \\ &\quad + \mathbf{q}_{M,N}(0,K)^T\mathbf{Q}_{M,N}^{\dagger}(I,K)\,\mathrm{E}\left[\mathbf{H}_{i_0,k_0}\mathbf{H}_{i_0,k_0}^T\right] \\ &\quad \times \mathbf{Q}_{M,N}^{\dagger T}(I,K)\mathbf{q}_{M,N}(0,K), \end{aligned} \tag{43}$$

(1) *Initialization*: use $I_0 \times K_0$ calculate estimation and $\hat{\epsilon}_0 = \hat{\epsilon}_{I_0,K_0}$.
(2) Use window dimensions $I \times K$ to estimate the kth block and compute the estimated estimation error $\hat{\epsilon}_{I,K}$, $\hat{\epsilon}_{I+1,K}$ and $\hat{\epsilon}_{I,K+1}$.
(3) If $\hat{\epsilon}_{I,K} < \hat{\epsilon}_0$, then $I_0 = I$, $K_0 = K$, $\hat{\epsilon}_0 = \hat{\epsilon}_{I,K}$, and
    (a) if $|\hat{\epsilon}_{I,K} - \hat{\epsilon}_{I+1,K}| < \epsilon_{\text{th}}^f$, then I remains unchanged. Otherwise, if $\hat{\epsilon}_{I,K} > \hat{\epsilon}_{I+1,K}$, then $I = I+1$, if $\hat{\epsilon}_{I,K} < \hat{\epsilon}_{I+1,K}$, then $I = I-1$.
    (b) If $|\hat{\epsilon}_{I,K} - \hat{\epsilon}_{I,K+1}| < \epsilon_{\text{th}}^t$, then K remains unchanged. Otherwise, if $\hat{\epsilon}_{I,K} > \hat{\epsilon}_{I,K+1}$, then $K = K+1$, if $\hat{\epsilon}_{I,K} < \hat{\epsilon}_{I,K+1}$, then $K = K-1$.
    Otherwise, $I = I_0$ and $K = K_0$.
(4) Go to step 2 for block $k+1$.

Algorithm 3: Window dimension adaptive algorithm.

and the residual noise is

$$\epsilon_n = \sigma^2 \mathbf{q}_{M,N}(0,K)^T \mathbf{Q}_{M,N}^{\dagger}(I,K)\mathbf{Q}_{M,N}^{\dagger T}(I,K)\mathbf{q}_{M,N}(0,K). \tag{44}$$

The residual noise is reduced more when the model order $M \times N$ becomes small or the window dimension $I \times K$ becomes large. However, the model error will increase in this case. With fixed polynomial model order M and N, the optimal window dimension is obtained by

$$\min_{I,K} \epsilon_{I,K} = \epsilon_h + \epsilon_n. \tag{45}$$

Usually, there are several local minima in this optimization problem. Considering the computational complexity, we would prefer the one with small $I \times K$.

In order to adaptively adjust the window dimensions we need to know the estimation error. Since the actual channel responses are not known, we have to estimate the estimation error using the instantaneous estimates and the final estimates. Suppose that the noise statistics is known, we can calculate the estimated estimation error as

$$\begin{aligned} \hat{\epsilon}_{I,K} &= \sum_i \sum_k \left|\left|\tilde{H}_{i_0,k_0} - \hat{H}_{i_0,k}\right|\right|^2 - \sigma^2 \\ &\quad + \mathrm{E}\left[N_{i_0,k_0}\mathbf{N}_{i_0,k_0}^H\right]\mathbf{Q}_{M,N}^{\dagger}(I,K)\mathbf{q}_{M,N}(0,K) \\ &\quad + \mathbf{q}_{M,N}(0,K)^T\mathbf{Q}_{M,N}^{\dagger}(I,K)\,\mathrm{E}\left[\mathbf{N}_{i_0,k_0}N_{i_0,k_0}^*\right]. \end{aligned} \tag{46}$$

Using this approximation, the window dimension adaptive algorithm for the optimization of (45) is given in Algorithm 3.

If the recursive algorithm in Algorithms 1 and 2 is adopted, the window adaptation can be implemented easily. We just eliminate one downdating when increasing the window dimension, or eliminate one updating when decreasing window dimension.

One important problem in the adaptive algorithm is to determine the threshold ϵ_{th}^f and ϵ_{th}^t. With large threshold, the algorithm converges faster, but with larger deviation. Especially when the local minima are located closely, the

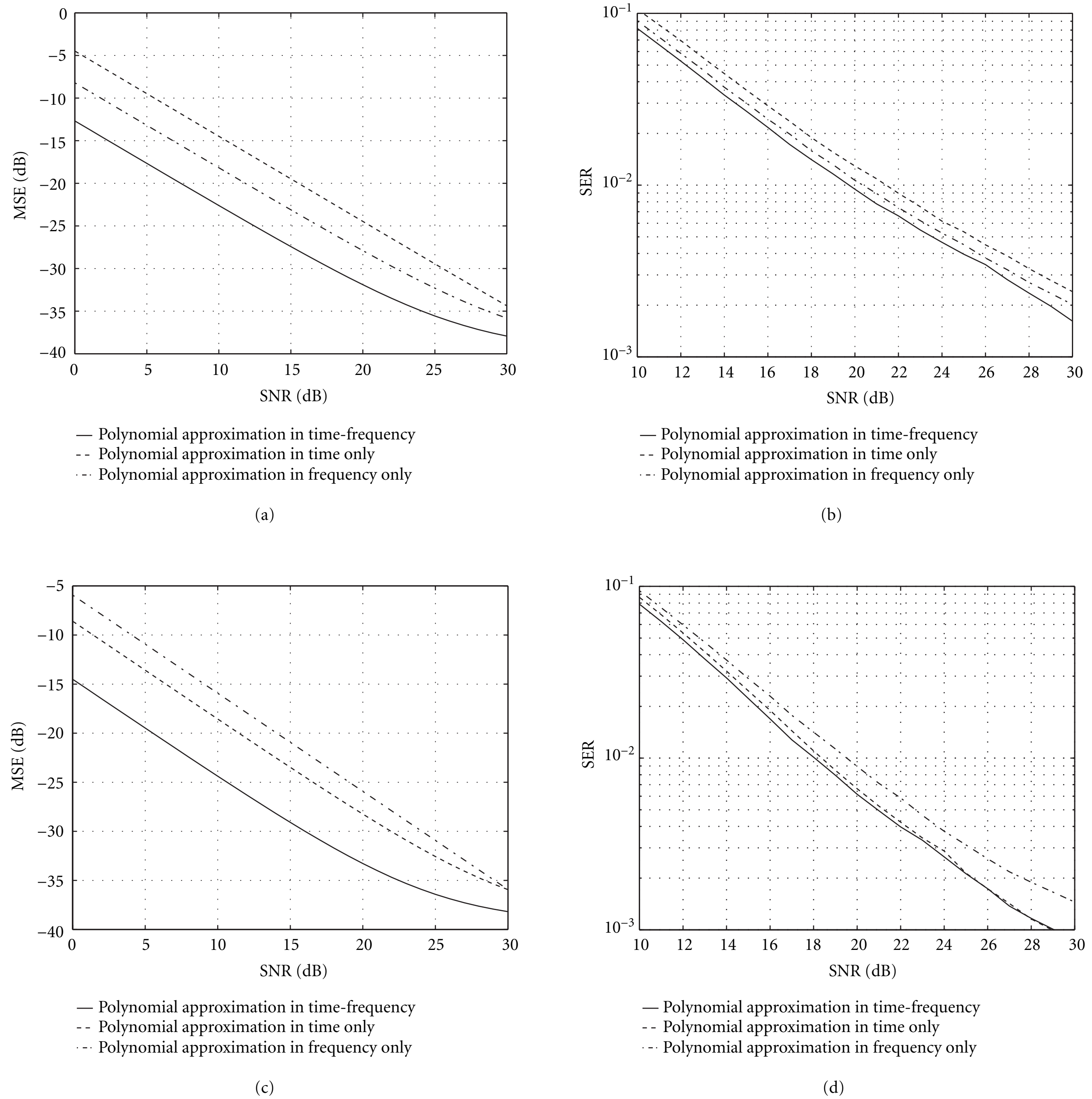

Figure 5: Estimation error and symbol error rate versus SNR (2-ray, $M \times N = 3 \times 3$) (a) MSE, (b) SER ($f_D = 40$ Hz, $T_d = 5$ microseconds, $I \times K = 7 \times 10$), (c) MSE, and (d) SER ($f_D = 20$ Hz, $T_d = 10$ microseconds, $I \times K = 4 \times 30$).

large threshold may result in unstable convergence. Hence, it would be preferred to use smaller thresholds here.

7. SIMULATION RESULTS

The OFDM system used in the simulations is the system introduced in Section 2. QPSK modulation is used throughout all subchannels. Figure 5 shows the mean-squared estimation error and the symbol error rate (SER) comparison of the algorithm based on the approximations in both time and frequency domains with those based on approximation either in time or frequency domain. Figures 5a and 5b show the case of a 2-ray channel with delay spread of 5 microseconds and Doppler frequency of 40 Hz, while Figures 5c and 5d show the case of another 2-ray channel with delay spread of 10 microseconds and Doppler frequency of 20 Hz. In both cases, $f_D T_d$ remains the same. We can see that the performance of using both time and frequency domain expansions

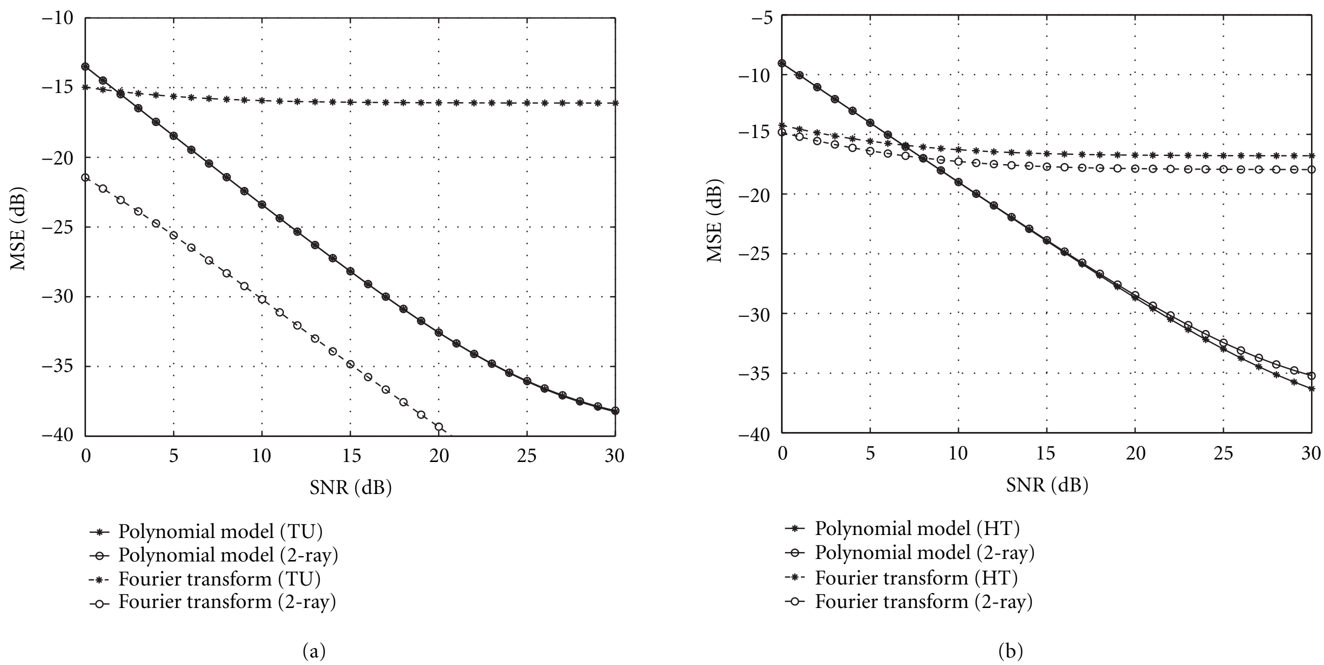

FIGURE 6: Estimation error versus SNR ($M \times N = 3 \times 3$, $f_D = 40$ Hz), (a) $I \times K = 5 \times 15$ and (b) $I \times K = 2 \times 15$.

is better than that of using only frequency domain expansion or using only time domain expansion in both cases. However, in the first case, the delay spread is smaller while the Doppler is larger, then the channel responses have more correlation in the frequency domain than in the time domain. Therefore, we use larger frequency domain window to exploit the frequency domain correlation. In the second case, the delay spread is larger while the Doppler frequency is smaller, then the channel responses have more correlation in the time domain and we use a larger time domain window to exploit it. It is shown that we have to use different time and frequency estimator to best exploit the channel correlations for different channels. Using only time or frequency domain scheme is not enough.

Figure 6 shows the estimation error under different delay profiles with Doppler frequency of 40 Hz. Figure 6a shows the estimation error with TU delay profile and 2-ray delay profile of $T_d = 5$ microseconds, which is the maximal delay spread of TU while Figure 6b shows the estimation error with HT delay profile and 2-ray delay profile of $T_d = 17.2$ microseconds which is the maximal delay spread of HT. We also compared the results using the Fourier-transform-based method of [6]. We can see that for TU or HT, the proposed algorithm performs much better than the Fourier-transform-based method. However, for 2-ray channel with $T_d = 5$ microseconds, the Fourier-transform-based method performs the best. The reason is that $T_d = 5$ microseconds is an integer multiplication of the sampling period of the OFDM system, which is $t_s = 1/800$ KHz $= 1.25$ microseconds. The impulse response of this 2-ray channel has energy only at the sampling instance

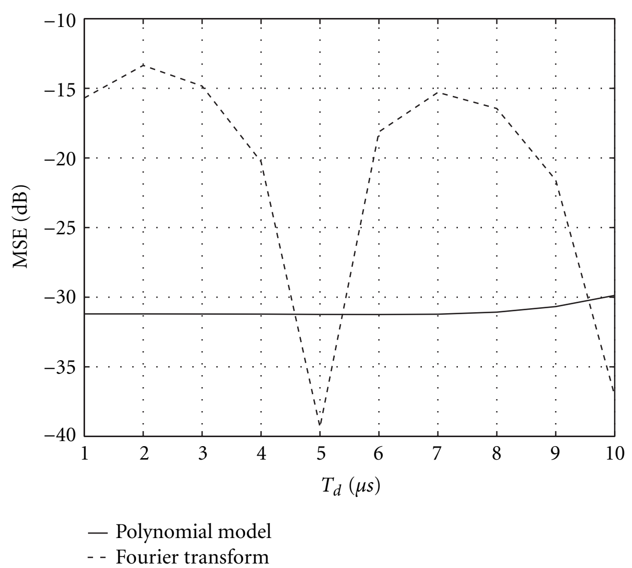

FIGURE 7: Estimation error versus delay spread (SNR = 20 dB, $f_D = 40$ Hz, 2-ray, $M \times N = 3 \times 3$, $I \times K = 5 \times 15$).

of the OFDM system, hence there is no leakage or model error using Fourier transform, which is used as frequency domain estimator in [6]. In this case, the Fourier-transform-based method actually provides a minimum mean-squared error estimator. Unfortunately, in the practice, such a case is quite unlikely especially for the time-varying channel. It is show that for TU or HT delay profiles and 2-ray with $T_d = 17.2$ microseconds, there is great amount of leakage

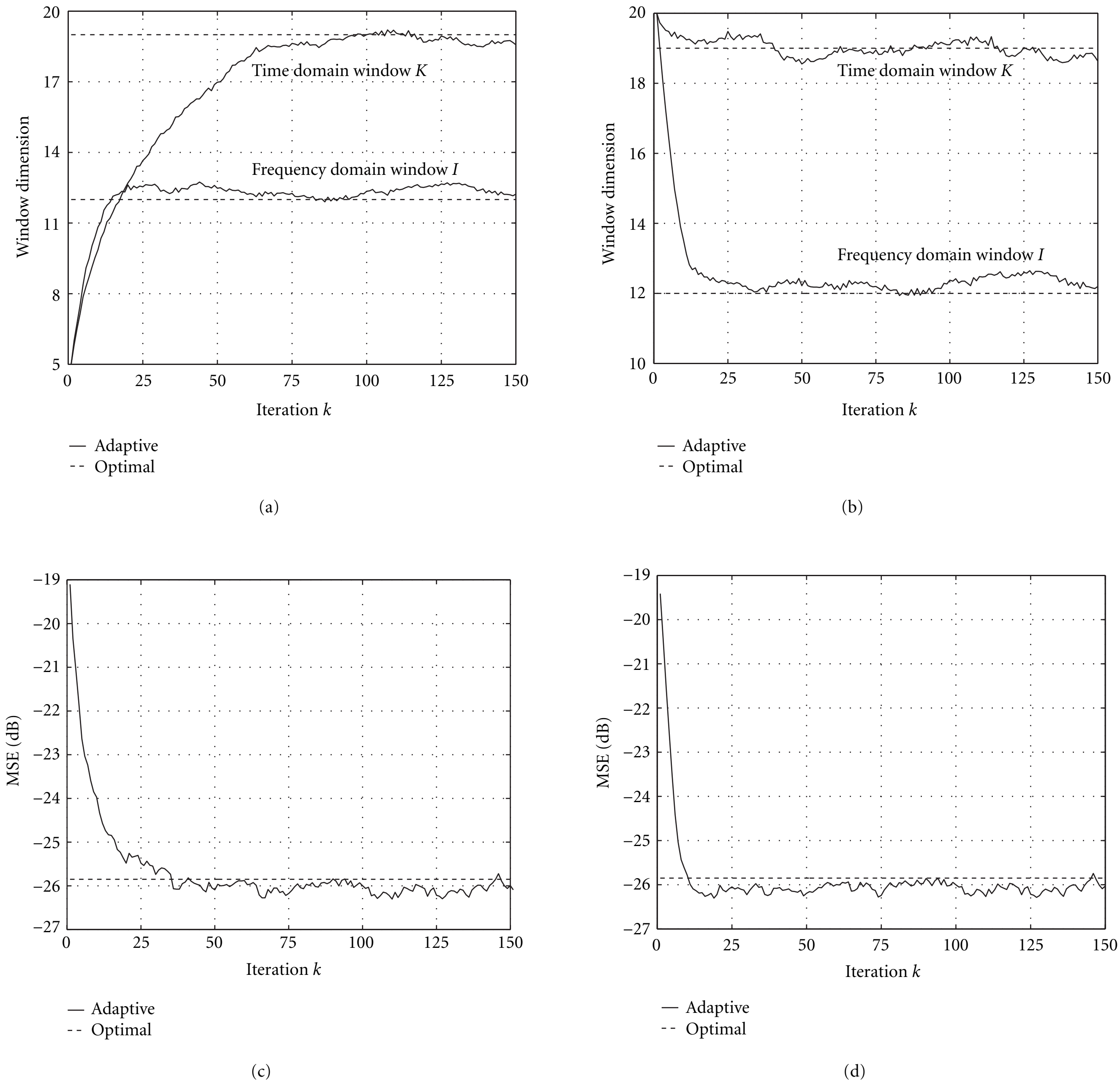

Figure 8: Window dimensions adaptation (TU, SNR = 10 dB, f_D = 40 Hz) (a) Window dimensions (starting from 5×5), (b) Window dimensions (starting from 20×20), (c) Estimation error (starting from 5×5), and (d) Estimation error (starting from 20×20).

using Fourier-transform-based on the sampling frequency of the OFDM system. The leakage greatly degrades the performance of the Fourier-transform-based method. In contrast, the polynomial-model-based method performs consistently for the channels with same maximal delay spread and hence is more robust to the channel statistics. This is because the model errors are bounded by the same bound for the channels with the same T_d and f_D as stated in Section 3.2. Therefore, the performance of the polynomial-model-based channel estimation is not sensitive to the specific correlation functions of the channels with the same Doppler frequency and maximum delay spread.

Figure 7 shows the mean-squared estimation error at SNR of 20 dB with different delay spread of a 2-ray channel. It further demonstrates the robustness of the polynomial-model-based method compared to the Fourier-transform-based method. The Fourier-transform-based method performs better only when the delay spread is at the sampling instance of the system. For most of the cases, it performs poorly. However, for the polynomial-model-based method, it performs consistently and outperforms the Fourier-transform-based method most of the time.

Figure 8 shows the window dimension adaptation. The window dimension variation is shown in Figures 8a and 8b. The estimation error is shown in Figures 8c and 8d. Two cases with different initial conditions are simulated, which are shown in Figures 8a and 8c and Figures 8b and 8d, respectively. In Figures 8a and 8c, the window dimension is 5×5 at the beginning, while in Figures 8b and 8d, it is 20×20. In both cases, after about 100 iterations, the algorithm converges to a window dimension of 12×10 and an estimation error under -26 dB. However, as mentioned in Section 6, smaller window dimensions are preferred for the sake of the computation complexity. With this adaptation algorithm, the polynomial-model-based method is not only robust to the specific correlation of the channel variation and dispersion, but also robust to T_d and f_D and can follow the variation of the statistics of the channel. Moreover, in the previous simulation, fixed window dimensions are used, by applying this window dimension adaptation algorithm, the performance in Figure 6 can be further improved.

8. CONCLUSIONS

In this work, we proposed a channel estimation algorithm for the OFDM system with fading multipath channels, which is suitable for the applications in 3G wireless communications. The algorithm is based on the time-frequency polynomial model that exploits the correlation of the channel responses in both time and frequency domains. The channel response is approximated by a small number of time-frequency polynomial basis functions and estimated by first estimating the coefficients of the bases. The residual noise is significantly reduced in this way, compared to the results when approximation is only done either in time or frequency domain, and the estimator design is more flexible. Therefore, the approach is more robust to the channel statistics and system parameters than the existing Fourier-transform-based method. It does not require the delay profiles to be integer multiples of the system sampling period. Moreover, the algorithm can be implemented recursively and can adjust the model parameters adaptively to the delay and fading characteristics.

REFERENCES

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.

[2] L. J. Cimini Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Communications*, vol. 33, no. 7, pp. 665–675, 1985.

[3] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Communications*, vol. 43, no. 2, pp. 773–775, 1995.

[4] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Baörjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Communications*, vol. 46, no. 7, pp. 931–939, 1998.

[5] V. Mignone and A. Morello, "CD3-OFDM: a novel demodulation scheme for fixed and mobile receivers," *IEEE Trans. Communications*, vol. 44, no. 9, pp. 1141–1151, 1996.

[6] Y. (G.) Li, L. J. Cimini Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Communications*, vol. 46, no. 7, pp. 902–915, 1998.

[7] E. W. Cheney, *Introduction to Approximation Theory*, McGraw-Hill, New York, NY, USA, 1966.

[8] H. N. Mhaskar, *Introduction to the Theory of Weighted Polynomial Approximation*, World Scientific Publishing, Singapore, 1996.

[9] D. K. Borah and B. D. Hart, "A robust receiver structure for time-varying, frequency-flat Rayleigh fading channels," *IEEE Trans. Communications*, vol. 47, no. 3, pp. 360–364, 1999.

[10] D. K. Borah and B. D. Hart, "Frequency-selective fading channel estimation with a polynomial time-varying channel model," *IEEE Trans. Communications*, vol. 47, no. 6, pp. 862–873, 1999.

[11] M. Luise, R. Reggiannini, and G. M. Vietta, "Blind equalization/detection for OFDM signals over frequency-selective channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1568–1578, 1998.

[12] W. C. Jakes, *Microwave Mobile Communications*, Wiley, New York, NY, USA, 1974.

[13] Y. (G.) Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 461–471, 1999.

[14] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Communications Systems*, vol. 11, no. 4, pp. 360–393, 1963.

[15] X. Wang and K. J. R. Liu, "OFDM channel estimation based on time-frequency polynomial model of fading multipath channel," in *VTC*, Fall 2001.

[16] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.

Xiaowen Wang received her B.S. degree from the Department of Electronics Engineering, Tsinghua University, Beijing, China in 1993, and the M.S. and Ph.D. degrees from the Department of Electrical and Computer Engineering, University of Maryland, College Park, Md, in 1999 and 2000, respectively. From 1993 to 1996, Dr. Wang was a Teaching Assistant with Tsinghua University, Beijing, China. From 1996 to 2000, she was a Research Assistant with University of Maryland, College Park, Md. Since 2000, she has been with the Wireless Systems Research Department, Agere Systems (formerly Bell Labs, Lucent Technologies, Microelectronics). Her research interests include adaptive digital signal processing, wireless communications and networking. Dr. Wang was ranked first among the class of department of Electronics Engineering for her B.S. degree from Tsinghua University in 1993, and was the recipient of the Graduate School Fellowship from University of Maryland.

K. J. Ray Liu received his B.S. degree from the National Taiwan University, and the Ph.D. degree from UCLA, both in electrical engineering. He is Professor of Electrical and Computer Engineering Department of University of Maryland, College Park. His research interests span broad aspects of signal processing architectures; multimedia signal processing, wireless communications and networking, information security, and bioinformatics in which he has published over 230 refereed papers, of which over 70 are in archival journals. Dr. Liu is the recipient of numerous awards including the 1994 National Science Foundation Young Investigator, the IEEE Signal Processing Society's 1993 Senior Award, IEEE 50th Vehicular Technology Conference Best Paper Award, Amsterdam, 1999. He also received the George Corcoran Award in 1994 for outstanding contributions to electrical engineering education and the Outstanding Systems Engineering Faculty Award in 1996 in recognition of outstanding contributions in interdisciplinary research, both from the University of Maryland. Dr. Liu is the Editor-in-Chief of EURASIP Journal on Applied Signal Processing, and has been an Associate Editor of IEEE Transactions on Signal Processing, a Guest Editor of special issues on Multimedia Signal Processing of Proceedings of the IEEE, a Guest Editor of special issue on Signal Processing for Wireless Communications of IEEE Journal of Selected Areas in Communications, a Guest Editor of special issue on Multimedia Communications over Networks of IEEE Signal Processing Magazine, a Guest Editor of special issue on Multimedia over IP of IEEE Trans. on Multimedia, and an editor of Journal of VLSI Signal Processing Systems.