Long Range Prediction and Reduced Feedback for Mobile Radio Adaptive OFDM Systems

Alexandra Duel-Hallen, Hans Hallen, and Tung-Sheng Yang

***Abstract*— Adaptive orthogonal frequency division multiplexing (AOFDM) modulation is a promising technique for achieving high data rates required for wireless multimedia services. To accomplish efficient adaptive channel loading, the channel state information (CSI) needs to be fed back to the transmitter. Since the fading channel varies rapidly for fast vehicle speeds, long range fading prediction (LRP) is required for mobile radio AOFDM to insure reliable adaptation. We use past channel observations to predict future CSI and perform adaptive bit and power allocation for the OFDM system. We derive the minimum mean-square-error (MMSE) long-range channel prediction that utilizes the time and frequency domain correlation functions of the Rayleigh fading channel. Since the channel statistics are usually unknown, robust prediction methods that do not require the knowledge of the correlation functions are developed. Statistical model of the prediction error is created and used in the design of reliable adaptive modulation. In addition, several methods that significantly reduce the feedback load for mobile radio AOFDM systems are developed and compared. We use a standard sum-of-sinusoids model and our realistic physical model to validate perfrormance of proposed methods. Simulation results demonstrate reliable performance and robustness of the proposed techniques, thus validating feasibility of AOFDM for rapidly varying mobile radio channels.**

***Index Terms*— Adaptive modulation, fading channel prediction, multipath fading, orthogonal frequency division multiplexing (OFDM), physical channel modeling.**

I. Introduction

CONTINUED increase in demand for all types of wireless services such as voice, data, and multimedia is fueling the need for higher capacity and data rate. Orthogonal frequency division multiplexing (OFDM) [1], [2] has been proposed for use in high-speed wireless data applications due to its relatively simple receiver structure compared with single carrier transmission in frequency selective fading channels. In OFDM, the frequency band is divided into narrow subcarriers, and data bits are multiplexed onto these subcarriers. These systems eliminate the need for an equalizer by greatly reducing the intersymbol interference (ISI), and have been extremely popular in Digital Subscriber Line (DSL), digital audio and television broadcasting (DAB and DVB) in Europe and wireless Internet access.

Adaptive OFDM (AOFDM) system, similarly to adaptive modulation for single carrier flat fading channels [3-6], [14], [15], involves optimizing the modulation level and the transmit power over the entire frequency band to maximize the spectral efficiency. In a frequency selective fading channel, some subcarriers experience a deep fade while others are subject to channel gain. The spectral efficiency can be improved by allocating more bits to those subcarriers with favorable channel conditions than to those in a deep fade, as motivated by the "*water filling*" distribution [7]. Several practical integer-bit and power allocation algorithms have been addressed in [8-11] that perform the optimum or near-optimum loading of bit and power in an OFDM frame.

Since the channel is rapidly time variant for fast vehicle speeds, there is a mismatch between the channel quality that is estimated by the receiver and fed back to the transmitter, and that is actually experienced during the transmission of the following OFDM frame. This degrades the achievable performance gain of AOFDM, especially in rapidly time variant fading. Even in an open-loop system where the CSI is obtained by channel estimation at the transmitter as in time division duplex (TDD) operation [5], current CSI is not sufficient since future channel conditions need to be known to adapt transmission parameters. To realize the potential of AOFDM, the channel variations have to be reliably predicted at least several milliseconds ahead.

Recently, several techniques were developed to predict future behavior of the mobile channel. In [30], the multivariate adaptive regression splines (MARS) model was used to capture the dynamics for predicting parameters of wideband fading channels several millisecond ahead for fast vehicle speed. The sub-space based [24] root-MUSIC method [31] and ESPRIT type algorithm [32] were employed to estimate the power spectrum that constitutes the fading process. Then these sinusoids were extrapolated to predict future samples. These methods were tested using synthetic and measured data, and it was concluded that reliable prediction is feasible at least one wavelength into the future. An adaptive long-range prediction (LRP) method for flat fading channel was also proposed in [12-15]. This algorithm employs an autoregressive (AR) model to characterize the fading channel and computes the minimum mean-square-error (MMSE) estimate of a future fading coefficient based on a number of past observations. The advantage of this algorithm relative to conventional methods is due to its low sampling rate (on the order of twice the max-

Manuscript received April 14, 2004; revised March 15, 2005 and October 3, 2005; accepted October 3, 2005. The associate editor coordinating the review of this paper and approving it for publication was C. Tellambura. This research was supported by NSF grant CCR-0312294 and ARO grant DAAD 19-01-1-0638.

A. Duel-Hallen is with the Dept. of Electrical and Computer Engineering, North Carolina State University, P. O. Box 7911, Raleigh, NC 27695-7911 USA (e-mail: sasha@eos.ncsu.edu).

H. Hallen with the Dept. of Physics, North Carolina State University, P. O. Box 8202, Raleigh, NC 27695-8202 USA (e-mail: Hans_Hallen@ncsu.edu).

T.-S. Yang was with the Electrical and Computer Engineering Dept. at North Carolina State University. He is now with Mobile Devices Technology, Taiwan ROC (e-mail: tsyang_99@yahoo.com).

Digital Object Identifier 10.1109/TWC.2006.xxxxx

imum Doppler shift and much lower than the data rate), which results in longer memory span and prediction further into the future for a fixed filter length. The low sampling rate also results in reduced feedback rate. The LRP can be implemented adaptively and thus is less complex and more robust than other fading prediction techniques. More recently, the LRP was extended to frequency selective channels. In [18], [19], LRP and adaptive modulation using CSI of another carrier was addressed, and in [27], LRP for frequency hopping (FH) systems was investigated. In this work and [38], we develop the long-range prediction algorithms for OFDM systems. An ideal MMSE method that utilizes previous observations in time and frequency domain, and robust adaptive LRP algorithms are developed and compared. The LRP is utilized in adaptive bit and power allocation for the OFDM system. Statistical model of the prediction error is created and used in the design of reliable adaptive modulation based on the method proposed in [6]. Adaptive modulation was also combined with single carrier channel prediction in [14], [15], [33], [34]. Adaptive OFDM with imperfect channel state information was analyzed in [35-37].

The fading channel is characterized as superposition of several reflected components. The accuracy of the LRP algorithm is determined by the rate of change of amplitude, frequency and phase associated with each reflector [12], [15-17]. However, the standard Jakes model or a stationary random process description does not capture the variation of these parameters. To validate the LRP, a realistic physical channel modeling based on the method of images was proposed in [15-17]. It was demonstrated in [15-17] that this physical model generates datasets that closely resemble measured data, and results of the LRP for the physical model and measured data are similar. In addition, this model can generate different scenarios to classify typical and challenging cases for testing the algorithm. These scenarios are more difficult to identify with the measured data. Thus, the physical model allows to test robustness and to determine practical constraints of the proposed adaptive transmission methods. In this paper, we employ this physical model to test performance of the LRP for AOFDM.

The transmitter can obtain the knowledge of the CSI to properly adapt the modulation parameters for each subcarrier from different sources. If the communication between the two stations is bi-directional and the channel can be considered reciprocal, as, for example, in TDD systems, then each station can estimate the channel quality on the basis of the received symbols and adapt the parameters to this estimation. This is called open-loop adaptation [5]. If the channel is not reciprocal, the receiver has to estimate channel quality from feedback resulting in closed-loop adaptation. The feedback load consumes power and bandwidth, and the fed back CSI needs to be quantized resulting in degraded performance. (Note that for many adaptive transmission applications, e.g., selective transmitter diversity or fixed power adaptive modulation, it is not necessary to feed back the actual fading coefficient. It is sufficient to send to the transmitter just the antenna selection or modulation index bits derived from the estimates of predicted values at the receiver. However, feedback of complex fading coefficients is required for some adaptive transmission applications and/or if the prediction is performed at the transmitter [15].) For AOFDM, the CSI is required for all subcarriers, resulting in high feedback load that increases for fast vehicle speeds due to high feedback rate. Hence feedback load should be minimized while providing sufficient information for the transmitter to predict the future CSI accurately. In this paper, we investigate reduction of the feedback load for closed loop systems by using the correlation induced by the multipath fading between the subcarriers.

The remainder of this paper is organized as follows. Section II introduces the channel model and statistics used to perform and test our prediction algorithms. In Section III, we first briefly describe the adaptive OFDM system, and then present the theoretical MMSE long-range prediction and robust prediction methods that do not require the knowledge of the channel statistics for the AOFDM system. In Section IV, robust adaptive bit and power loading for mismatched channel information is investigated and the reduced feedback techniques are explored. Computer simulation results are presented to demonstrate the effectiveness of the proposed system.

II. Propagation and Channel Model Statistics

A. Channel Statistics

The equivalent lowpass complex fading coefficients at K subcarriers, $f^1 < f^2 < \ldots < f^K$, where $|f^i - f^j| \ll$ the carrier frequency f_c, can be closely approximated as [21]:

$$c\left(f^i,t\right)=\sum_{n=1}^{N}A\left(n\right)\exp\left\{j\left(2\pi f_d\left(n\right)t+\phi_i\left(n\right)\right\}\right., \quad i=1,2\ldots K \tag{1}$$

In (1), for the n^{th} path, $A(n)$ is the (real) amplitude and $f_d(n) = f_c\frac{v}{c}cos(\theta(n)) = f_{dm}cos(\theta(n))$ is the Doppler shift, where v, c, f_{dm}, and $\theta(n)$ is the speed of the mobile, the speed of light, the maximum Doppler shift, and the incident angle of the path to the direction of the mobile, respectively. The phase difference for the n^{th} path, $\phi_i(n)-\phi_j(n) = 2\pi\Delta f\tau(n)$ where $\Delta f = f^j - f^i$ is the frequency separation, and $\tau(n)$ is the excess propagation delay. Let $\{\underline{A},\underline{\theta},\underline{\tau},\underline{\phi_1}\}$ denote the set $\{A\left(n\right),\theta\left(n\right),\tau\left(n\right),\phi_1\left(n\right),n=1\ldots N\}$ that parameterizes this channel model. While there are many reflections that contribute to the fading signal, the number of dominant paths is modest in practice (a dominant, or significant, reflected component has power within ~10dB of the strongest component) [15-17]. The $c(f^i,t)$ are distributed approximately as a zero mean complex Gaussian random variables. Therefore, the amplitudes $|c(f^i,t)|$ are Rayleigh distributed.

For a deterministic channel where $\{\underline{A},\underline{\theta},\underline{\tau}\}$ are known, *the temporal average correlation function* (TACF) [24] can be calculated from (1) as (2), which is found at the top of the next page. Note that TACF does not depend on the parameters $\left\{\underline{\phi}_1\right\}$. Assume $\{\phi_1(1)\ldots\phi_1(N)\}$ are mutually independent random variables uniformly distributed on $[0,2\pi]$. It can be readily shown from (1) that $E_{\left\{\underline{\phi}_1\right\}}[c(f^i,t)c^*(f^j,t+\Delta t)|\{\underline{A},\theta,\underline{\pi}\}]$ = TAFC in (2), where $E_{\left\{\underline{\phi}_1\right\}}$ denotes the statistical expectation over $\{\underline{\phi}_1\}$.

$$R_T(\Delta t, \Delta_f) = \lim_{T \to \infty} \frac{1}{2T - |\Delta t|} \int_{-T+|\Delta t|/2}^{T-|\Delta t|/2} c\left(f^i, t - \Delta t/2\right) c^*\left(f^j, t + \Delta t/2\right) dt = \sum_{n=1}^{N} A(n)^2 \exp\{-j2\pi f_d(n)\Delta t\} \exp\{j2\pi \Delta f \tau(n)\} \tag{2}$$

For $c(t)$ characterized as wide sense stationary uncorrelated scattering (WSSUS) [7], the *ensemble average correlation function* (EACF) for two fading signals with frequencies f^i and f^j with the time difference Δt and the frequency separation $\Delta f = f^j - f^i$ is defined as $R_E(\Delta t, \Delta f) = E[c(f^i, t)c^*(f^j, t + \Delta t)]$. Equivalently, $R_E(\Delta t, \Delta f)$ is the statistical average of the TACF over $\{\underline{A}, \underline{\theta}, \underline{\tau}\}$, i.e., $R_E(\Delta t, \Delta f) = E_{\{\underline{A},\underline{\theta},\underline{\tau}\}}\{E_{\{\underline{\phi}_1\}}[c(f^i, t)c^*(f^j, t + \Delta t)|\{\underline{A}, \underline{\theta}, \underline{\tau}\}]\}$. It can be factored into the time-domain correlation function $R_t(\Delta t)$ and the frequency domain correlation function $R_f(\Delta f)$ as [23], [20]:

$$R_E(\Delta t, \Delta f) = \Omega R_t(\Delta t) R_f(\Delta f) \tag{3}$$

where $\Omega = E\left[\left|c\left(f^i, t\right)\right|^2\right]$ is the average power of the fading signals. (We normalize Ω to 1 throughout the paper.) Assume $\theta(n)$ is uniformly distributed around 2π, and the propagation delay $\tau(n)$ is exponentially distributed [21] with the probability density function (*pdf*)

$$p(\tau) = \frac{1}{\sigma} \exp\{-\tau/\sigma\} \tag{4}$$

where σ is the *rms* delay spread [22]. Then $R_t(\Delta t) = J_0(2\pi f_{dm}\Delta t)$ is the zero order Bessel function [21] and $R_f(\Delta f) = \frac{1}{1+(2\pi\Delta f\sigma)^2} + j\frac{2\pi\Delta f\sigma}{1+(2\pi\Delta f\sigma)^2}$. We define $f_{dm}\Delta t$ and $\Delta f\sigma$ as the *normalized time difference* (NTD) and the *normalized frequency separation* (NFS), respectively. Note that (3) is expressed in terms of these normalized quantities.

The Jakes model [21] is widely used to model the fading channel. In this method, N equal strength multipath components in (1) are equidistant on the unit circle, i.e., $\theta(n) = 2\pi n/N$, and $A(n)$ are given by the same constant value for all $n = 1, 2 \ldots N$. This model results in fixed TACF $R_T(\Delta t, 0)$ since the incident angles $\theta(n)$ and the amplitudes are fixed in (1). The function $R_T(\Delta t, 0)$ approaches $R_t(\Delta t) = J_0(2\pi f_{dm}\Delta t)$ as N becomes large. (The performance analysis of the LRP for the Jakes model is presented in [13]). In this paper, we employ a standard modification of the Jakes model by modeling the incident angles $\theta(n)$ as independent and uniformly distributed on $[0, 2\pi]$ instead of using fixed $\theta(n) = 2\pi n/N$. Moreover, we do not make use of approximations for large N since spectral analysis of measured fading data strongly supports the conjecture that the complex baseband fading process is dominated by a moderate number of sinusoids [15-17]. We refer to this model as the *random phase model* (RPM). This model provides a more realistic environment for testing the LRP, while maintaining low complexity. The coefficients at different frequencies have the same parameters $\{\underline{A}, \underline{\theta}\}$, and propagation delay distribution is generated as in (4).

B. Physical Model

In the Jakes and the RPM models discussed above, the parameters associated with the reflectors (the amplitudes, the Doppler shifts and the phases) are fixed once they are chosen. Thus, the correlation function (2) does not vary in time. However, in real mobile radio environments, the correlation function is time-variant and is affected by many factors such as the number and the locations of the reflectors, vehicle speed, carrier frequency, distance between the transmitter and the receiver, etc. The LRP predicts the channel far ahead, and requires large observation interval and memory span [15]. Therefore, the performance of this algorithm is affected by the variation in time of the parameters associated with the reflectors. In practice, this variation has to be taken into account in the estimation of the correlation function. Thus, realistic non-stationary modeling is necessary.

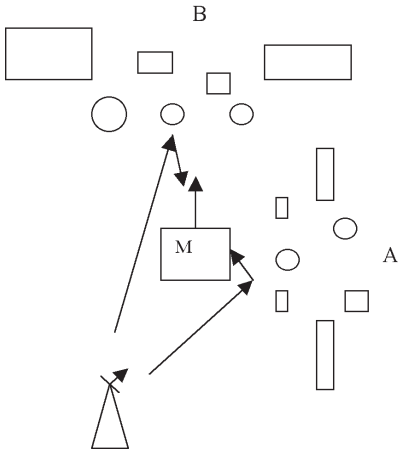

Fig. 1. Geometry of the physical modeling.

We have proposed a realistic physical model based upon the method of images combined with diffraction in [15], [16], [17]. This model can provide physical insights into the nature of the signal fading that affects the performance of the LRP algorithm. The geometry for generating the model data set used in this paper is shown in Fig. 1. The reflecting objects are arranged approximately to the side of (group A) and in front of (group B) the mobile. The fading amplitude is shown in Fig. 2. The carrier frequency is 900MHz, the maximum Doppler shift is 100Hz, and the sampling rate is 500Hz. The CSI is dominated by the reflectors that are closer. This is group A for the first ~600 samples, and group B thereafter. Taking into account the direction of the mobile, away from group A and towards group B, we expect that the dominant paths change from group A reflections to group B reflections, and the Doppler shifts of the dominant paths will rapidly change in both magnitude and sign in the vicinity of sample 600. This is evident in Fig. 2 as a faster fading rate when the Doppler shifts are larger, for points >600, when group B reflections dominate. We use this transition interval to test the robustness of the LRP to parameter variation in Sections III and IV.

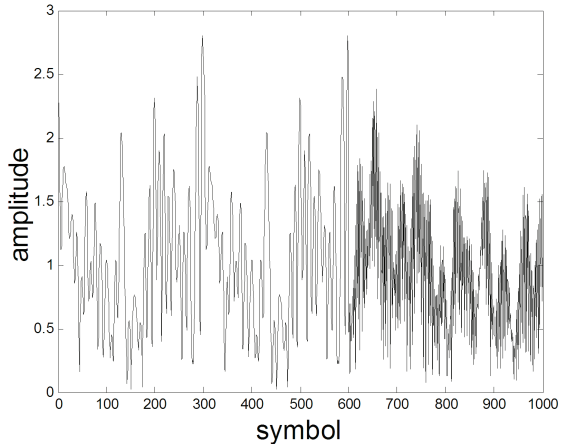

Fig. 2. Physical model data set.

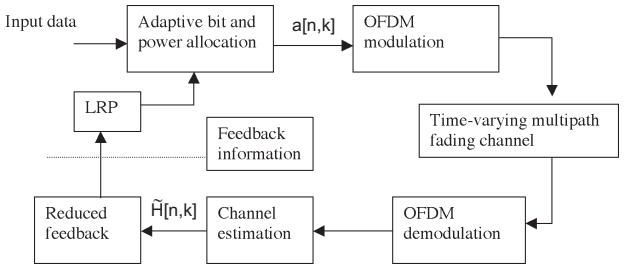

Fig. 3. Block diagram of an adaptive OFDM system.

III. System Model and Long Range Prediction

A. Adaptive OFDM System Model

Consider an OFDM signal with K subcarriers, symbol (block) duration T_s, and adjacent subcarrier (tone) spacing Δf_s. Assume the channel bandwidth of each subcarrier is much smaller than the coherence bandwidth and the channel state information does not change within one OFDM symbol duration T_s, but varies from symbol to symbol. The equivalent complex channel gain $H_s[n,k]$ at n^{th} symbol block and k^{th} subcarrier can be modeled as the samples of the time-varying frequency selective channel in (1) with the time domain and frequency domain sampling interval T_s and Δf_s. From (2, 3), the temporal and ensemble channel correlation functions for the OFDM symbols with block difference Δn and tone spacing Δk can be expressed as $R_T(\Delta n T_s, \Delta k \Delta f s)$ and $R_E(\Delta n T_s, \Delta k \Delta f_s)$, respectively.

The uncoded AOFDM system aided by the LRP and reduced feedback considered in this paper is depicted in Fig. 3. The input data is allocated to the subcarriers according to the CSI fed back from the receiver. The LRP is employed to enhance the CSI accuracy. Let $a[n,k]$ denote the complex baseband symbols at n^{th} block and k^{th} tone. The received signal after OFDM demodulation can be expressed:

$$X\left[n,k\right] = H_s\left[n,k\right] a\left[n,k\right] + w\left[n,k\right] \tag{5}$$

where $w[n,k]$ is complex additive white Gaussian noise with variance $E[|w[n,k]|^2] = N_0$. Then frequency domain coherent channel estimation of the complex symbols associated with each of the K subcarriers is employed. A 2-D MMSE channel estimator was proposed in [23]. Let

$$\widetilde{H}_s\left[n,k\right] = H_s\left[n,k\right] + \widetilde{w}\left[n,k\right] \tag{6}$$

denote an accurate MMSE estimate of the CSI, where $\widetilde{w}\left[n,k\right]$ is the estimation error modeled as white Gaussian noise with power spectrum $\widetilde{N}_0$. We define the observation SNR as $E[|H_s[n,k]|^2]/\widetilde{N}_0$. Due to the correlated subcarriers, the estimated CSI $\widetilde{H}\left[n,k\right], k = 1 \ldots K$, can be reduced and fed back to the long-range predictor at the transmitter at low rate. Alternatively, the predictor can be placed at the receiver between the channel estimation and the reduced feedback blocks depending on the implementation issues such as complexity, performance and costs. The reduced feedback methods will be discussed in Section IV. In this section, we assume that estimates of all subcarriers $\widetilde{H}\left[n,k\right], k = 1 \ldots K$ are available for the following long-range prediction algorithm.

B. MMSE Long Range Prediction

We derive the linear MMSE-based channel predictor for a multicarrier fading channel (1) where $\{\underline{A}, \underline{\theta}, \underline{\tau}\}$ are fixed and the components of $\{\underline{\phi}_1\}$ are mutually independent random variables uniformly distributed on $[0, 2\pi]$. One important parameter for the LRP is the sampling rate. For narrow band single carrier systems, the sampling rate of the LRP is much lower than the symbol rate [15]. While the symbol interval in OFDM systems is longer, it is still beneficial to choose the sampling rate of the LRP lower than the symbol rate. Let $\widetilde{H}\left[n,k\right]$ denote the estimated CSI (as in (6)) with a sample interval T_p (an integer multiple of the OFDM symbol interval T_s). The channel predictor for the CSI at the k^{th} tone and the n^{th} sample based on the p previously observed samples at K subcarriers can be constructed by:

$$\widehat{H}[n,k] = \sum_{j=1}^{p} \sum_{m=1}^{K} d^*(j,m) \widetilde{H}[n-j,m] \tag{7}$$

Provided that the correlation function (2) is known, the optimal filter coefficients $d_o(j,m)$ that minimize the conditional MSE $E_{\{\underline{\phi}_1\}}\left[\left|\left(H\left[n,k\right] - \widehat{H}\left[n,k\right]\right)\right|^2 |\{\underline{A}, \underline{\theta}, \underline{\tau}\}\right]$ can be obtained by using the orthogonality principle [24]. The resulting MMSE J_{min} is given by

$$J_{min} = 1 - tr\left[\underline{\mathbf{D}}^H \underline{\mathbf{G}}\right] \tag{8}$$

where tr[•] is the trace of a matrix, the subscript H is *Hermitian* transpose and the matrix

$$\underline{\mathbf{D}} = \begin{bmatrix} d_0\left(1,1\right) & d_0\left(1,2\right) & \cdots & d_0\left(1,K\right) \\ d_0\left(2,1\right) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ d_0\left(p,1\right) & \cdot & \cdot & d_0\left(p,K\right) \end{bmatrix} \text{ and}$$

$$\underline{\mathbf{G}} = \begin{bmatrix} r\left(1,k-1\right) & r\left(1,k-2\right) & \cdots & r\left(1,k-K\right) \\ r\left(2,k-1\right) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ r\left(p,k-1\right) & \cdot & \cdot & r\left(p,k-K\right) \end{bmatrix}$$

where $r(\Delta n, \Delta k) = R_T(\Delta n T_p, \Delta k \Delta f_s)$. This result serves as a theoretical foundation for our prediction problem and will be used in the performance analysis. The MMSE J_{min} depends on the scattering configuration $\{\underline{A}, \underline{\theta}, \underline{\tau}\}$ and number of waves N of the fading process. In general, the performance of our prediction method degrades as N increases [13, 20]. As N becomes large, the TACF (2) for the Jakes model and for the RPM approaches EACF (3) [21]. Therefore, J_{min} is upper-bounded by (8) with $r(\Delta n, \Delta k) \approx R_t(\Delta n T_p) R_f(\Delta k \Delta f_s)$. Moreover, the MMSE prediction given past observations is achieved by the proposed linear MMSE algorithm due to the assumption of joint Gaussian distribution of the observations and future samples [25].

In the linear prediction algorithm (7), the optimum MMSE is achieved by observing previous symbols of multiple subcarriers. However, this method is very complex in practice. Moreover, we have found that if the SNR of the observed feedback samples is high, the improvement in the prediction accuracy when non-adjacent subcarriers' observations are used relative to utilizing just past samples of desired and adjacent subcarriers is negligible [20]. In fact, for the asymptotic case when the TAFC approaches the EACF factorization (3), and when the CSI is noiseless, it is sufficient to use just the past samples of the desired subcarrier to achieve the optimal MMSE performance [20]. Thus, we propose to simplify the algorithm by using only previously observed samples at subcarrier k to predict the CSI $H[n,k]$:

$$\widehat{H}[n,k] = \sum_{j=1}^{p} d_j^*(n)\, \widehat{H}(n-j,k) \qquad k = 1, 2, \ldots K \quad (9)$$

Adjacent subcarriers can be easily incorporated to reduce the noise level at the cost of the system complexity. For the channel model (1), the same coefficient vector $\underline{d}(n) = [d_1 d_2 \ldots d_p]^T$ in (9) can be employed for each subcarrier, since the MMSE solution is determined by TACF in (2) with $\Delta f = 0$. (More specifically, it was shown in [12] that for single carrier prediction, the coefficients d_j depend only on the Doppler shifts for that carrier. We assume the total bandwidth is sufficiently small to model the Doppler shifts as fixed in (1), although in practice the carrier dependency becomes more significant for large frequency separations.) Since the filter coefficient vector $\underline{d}(n)$ are assumed tone-invariant, significantly reduced computational complexity and greatly improved tracking ability can be achieved for the adaptive prediction methods discussed in the following sections because all feedback observations can be used jointly to update the coefficients. We call this method *simplified multiple carriers prediction* (SMCP). Note that this tone invariability can be generalized to the case when observation of several adjacent carriers are used provided that the same number of adjacent carriers is employed on each side of the desired carrier [20]. This method extends to adaptive transmitter antenna diversity systems since the channels for all antennas have the same Doppler shifts [28].

The optimum MMSE channel prediction above relies on the knowledge of the time and frequency domain correlation functions (2). However, these correlation functions depend on the particular environment and usually are unknown. In addition, the coefficients $\underline{d}(n)$ in (9) needs to be computed adaptively as the Doppler shifts in (1) vary with time. In the following section, we employ the adaptive Least Mean Square (LMS) and Recursive Least Squares (RLS) algorithms, which do not require the knowledge of the correlation functions of the channel, to update the prediction filter coefficients for the OFDM system. The error between the desired response and the predicted CSI at subcarrier k is:

$$e[n,k] = H[n,k] - \sum_{j=1}^{p} d_j^*[n]\, \widetilde{H}[n-j,k], \qquad k = 1 \ldots K. \quad (10)$$

The average mean square error (AMSE) over all subcarriers is

$$AMSE = J(n) = \frac{1}{K} \sum_{k=1}^{K} |e[n,k]|^2 \quad (11)$$

This AMSE is used for updating the coefficients of the LMS and RLS SMCP algorithms. Note that $e[n,k]$ is the MSE of a single carrier prediction method (9), and equation (11) is the average MSE for all subcarriers. The subcarriers are parameterized by the same $\{\underline{A}, \underline{\theta}, \tau\}$ but different phases $\{\underline{\phi}_i\}$, where $\{\phi_i\} = \{\phi_i(1) \ldots \phi_i(N), i = 1 \ldots K\}$. As the number of subcarriers K increases, this phase vector can be modeled as a random vector with components that are uniformly distributed on the unit circle, and the average (11) can be approximated as $J(n) \approx E_{\{\underline{\phi}_k\}}[e[n,k]^2 | \{\underline{A}, \underline{\theta}, \underline{\tau}\}]$, the expectation (over the random phase) of the MSE of a single carrier predictor. Thus, it is lower bounded by the ideal single carrier MMSE J_{smin}

$$J_{smin} = J_{min} \text{ for } K = 1, \quad (12)$$

where J_{min} is defined in (7, 8). Note that using AMSE in SMCP, we adapt the coefficient vector $\underline{d}(n)$ jointly using the errors for all subcarriers. As discussed below, this improves accuracy and convergence relative to single carrier adaptive prediction [13], [15], [17], [20]. We also observed that the prediction algorithm is more robust to noise in the feedback signals compared to the single carrier prediction [14], [15] for both the LMS and RLS algorithms if the adjacent subcarriers are employed for prediction.

Assume the first p (filter length) samples are available for all subcarriers. Define the desired samples at time $p+n$ as $\underline{\widetilde{H}}_d[n] = [\widetilde{H}[p+n,1], \widetilde{H}[p+n,2], \ldots, \widetilde{H}[p+n,K]]^T$ and the input $p \times p$ data matrix, which is presented at the top of the next page. This notation will be used in the following robust prediction algorithms.

C. LMS and RLS Algorithms for AOFDM

The LMS algorithm uses the cost function $J(n)$ in (11). The equations for updating the predictor coefficients, shown in Table 1, are derived by calculating the gradient vector $\nabla J(n)$. From [24], the LMS learning curve depends on the step size μ and the eigenvalue spread of the correlation matrix, which can be determined from the TAFC (2), of the input process. The selection of step size μ is a trade-off between the convergence rate and the excess MSE:

$$J_{ex}(n) = J(n) - J_{smin} \quad (13)$$

$$\underline{\widetilde{H}}[n] = \begin{bmatrix} \widetilde{H}(p+n-1,1) & \widetilde{H}(p+n-1,2) & \cdots & \widetilde{H}(p+n-1,K) \\ \widetilde{H}(p+n-2,1) & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \widetilde{H}(n,1) & \cdot & \cdot & \widetilde{H}(n,K) \end{bmatrix}$$

TABLE I
LMS ALGORITHM FOR OFDM CHANNEL PREDICTION

$\underline{d}(0) = \underline{0}$

For each time instance, n=1,2…, compute

$\underline{\xi}(n) = \underline{\tilde{H}}_d(n) - \underline{d}(n-1)^H \underline{\tilde{\boldsymbol{H}}}(n)$

$\underline{d}(n) = \underline{d}(n-1) + \mu \frac{1}{K} \underline{\tilde{\boldsymbol{H}}}(n) \underline{\xi}^*(n)$

TABLE II
RLS ALGORITHM FOR OFDM CHANNEL PREDICTION

Initialize the algorithm by setting

$\underline{P}(0) = (\underline{\tilde{\boldsymbol{H}}}(1)\underline{\tilde{\boldsymbol{H}}}^H(1) + \delta \underline{I})^{-1}$

$\underline{d}(0) = \underline{0}$

For each time instance, n=1,2…, compute

$$\underline{k}(n) = \frac{\lambda^{-1}\underline{P}(n-1)\underline{\tilde{\boldsymbol{H}}}(n)}{1+\lambda^{-1}\underline{\tilde{\boldsymbol{H}}}^H(n)\underline{P}(n-1)\underline{\tilde{\boldsymbol{H}}}(n)}$$

$\underline{\xi}(n) = \underline{\tilde{H}}_d(n) - \underline{d}(n-1)^H \underline{\tilde{\boldsymbol{H}}}(n)$

$\underline{d}(n) = \underline{d}(n) + \underline{k}(n)\underline{\xi}^*(n)$

$\underline{P}(n) = \lambda^{-1}\underline{P}(n-1) - \lambda^{-1}\underline{k}(n)\underline{\tilde{\boldsymbol{H}}}^H(n)\underline{P}(n-1)$

where $J(n)$ and J_{smin} are defined in (11) and (12), respectively.

With the RLS algorithm, the predictor coefficients are calculated so that they minimize the error $\varepsilon[n] = \sum_{i=1}^{n}\sum_{k=1}^{K}\lambda^{n-i}|e[i,k]|$ where λ with $0 < \lambda \leq 1$ is the forgetting factor that accounts for possible non-stationary of the input CSI. The resulting update equation for the predictor coefficient vector $\underline{d}[n]$ is shown in Table 2. The parameter δ in the initialization procedure should be chosen small compared to the variance of the data samples. We use $\delta = 0.01$ in our simulation. For the RLS, the excess mean square error $J_{ex}(n)$ for SMCP is derived similarly to that for the single carrier case [24]. Its learning curve, unlike LMS algorithm, is independent of the eigenvalue spread of the input process and decays almost linearly with nK (the convergence rate is approximately K times faster than for single carrier prediction). For $\lambda = 1, J_{ex}(n)$ converges to zero, and for λ close to 1 and large K, the $J_{ex}(\infty) \approx J_{smin}\frac{(1-\lambda)p}{2K}$. Hence for large K, $J_{ex}(\infty) \approx 0$. (While $J_{ex}(\infty) \geq 0$ in our example, for small values of K and proper choice of λ (close to 1 but not 1), it is possible for $J_{ex}(\infty)$ to be smaller than 0 [20] due to the nonstationarity of the deterministic channel model (1) [29] and the fast tracking property of the RLS algorithm.) Thus, SMCP improves the convergence rate and the steady state MSE for the RLS relative to the single carrier prediction [12-15].

D. Numerical Simulations

We use the RPM and the physical model to validate the performance of the LRP for the OFDM system. To test the performance of our prediction algorithm on the fading channel modeled by the RPM, $N = 34$ is chosen and multiple deterministic channel realizations are generated by using independent angles $\{\underline{\theta}\}$ and propagation delays $\{\underline{\tau}\}$. We apply the LRP for each independent realization and calculate the average AMSE (11). Furthermore, the J_{smin} (12) averaged over these independent realizations is presented for comparison. Note that when we use the RPM, each channel realization has different TACF. Thus, we can test the tracking ability and the prediction accuracy (AMSE) of our prediction algorithm in different reflecting environments. This task cannot be accomplished using the Jakes model, where the incident angles are fixed. The prediction filter length p in (9) is 50. The maximum Doppler shift of 100 Hz is used in both models. The *rms* delay spreads are approximately $1\mu s$ in both channel models. To construct an OFDM symbol, assume that the entire channel bandwidth, 800kHz, is divided into 128 subcarriers. The symbol duration is $160\mu s$. An additional $5\mu s$ guard interval is used to provide protection from ISI due to channel multipath delay spread. Thus the total block length is $165\mu s$ and the subcarrier symbol rate is approximately 6KHz. For each subcarrier, the fading signal is sampled at the low rate of 466Hz for the LRP (the prediction range is 1/466Hz $\approx$ 2ms). In this paper, we assume reliable channel estimation and high effective SNR (80 dB) of the observed CSI. While the actual SNR of the observed samples is usually much lower, noise reduction techniques can be employed to decrease the estimation error greatly, and additive noise correlation in (6) that might result from noise reduction does not affect the accuracy of the prediction [13], [15], [23], [31], [33].

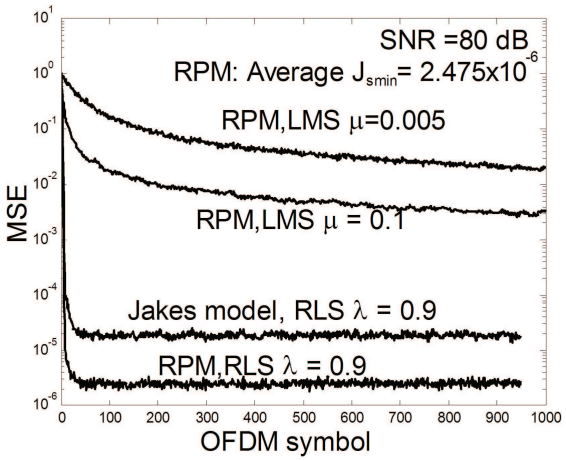

Fig. 4. Performance of different adaptive prediction methods for the random phase model.

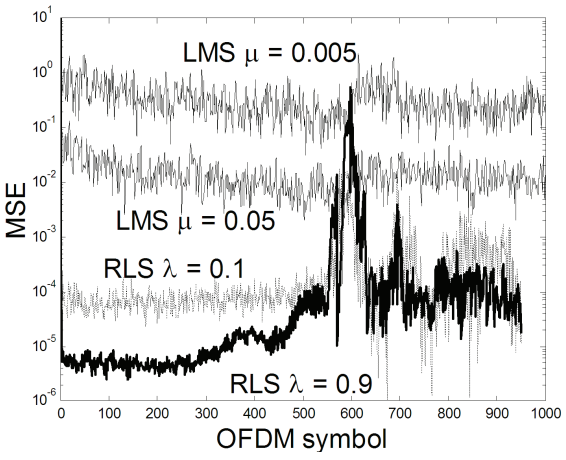

Fig. 5. Performance of different adaptive prediction methods for the physical model.

Moreover, our investigation in [20] shows that degradation due to lower effective SNR values (e.g., 50 dB) is negligible. Therefore, accurate channel estimation assumed in this paper is realistic and is not a limiting factor in the performance of adaptive prediction. Interpolation is utilized to predict channel coefficients at the subcarrier symbol rate [13-15].

Fig. 4 demonstrates the average AMSE (11) over all the channel realizations for the SMCP method for the RPM model. When these results are compared with the single carrier prediction, we find that the excess mean square error for the LMS algorithm $J_{ex}(n)$ (13) of these two approaches is approximately the same given the same step size μ [20]. The MSE curve shown for $\mu = 0.005$ corresponds to both methods. However, the single carrier algorithm diverges for large μ, while for the SMCP (9), μ can be chosen as large as 0.1 without divergence, thus improving the convergence rate. As the NFS increases, larger step size μ can be chosen, resulting in faster convergence [20]. While the RLS has higher computational complexity than the LMS algorithm, its learning curve and the excess MSE $J_{ex}(n)$ (13) are significantly improved relative to the LMS. It is observed that the RLS algorithm converges rapidly with almost no excess MSE for $\lambda = 0.9$, whereas the LMS algorithm converges more slowly with significant excess MSE relative to the RLS algorithm. We also demonstrate the RLS algorithm for the Jakes model. We observe that for $\lambda = 0.9$, the AMSE of the more realistic RPM is better than that of the Jakes model.

In Fig. 5, the SMCP is explored for the physical model. It is demonstrated that during the transition period (from samples 500 to 700 in Fig. 2), the RLS with the forgetting factor $\lambda = 0.1$ has better tracking ability than that with $\lambda = 0.9$. Hence it is more robust to the non-stationary environment. The tracking results for the LMS algorithm are much poorer. The actual impact of the choice of adaptive tracking method should be measured by the performance of adaptive transmission (the MSE is not always indicative of this performance.) This impact is analyzed in Section IV.

IV. Adaptive OFDM and Reduced Feedback

A. Robust Adaptive Bit and Power Loading

In this paper, we employ channel loading optimization under the *bit rate maximization* (BRM) criterion, where the goal is to allocate the limited energy among the subcarriers to maximize the overall bit rate subject to a target bit error rate constraint [11]. A simplified loading method similar to [10] (see [20] for the detailed description) is compared with the optimal Hughes-Hartogs algorithm [8] in the presence of imperfect CSI that results from prediction errors.

For each subcarrier we employ rectangular $M(i)$-QAM modulation [26] where $M(1) = 0$, $M(i) = 2^{i-1}$, $i = 2 \ldots 6$. Let $\widehat{c}$ denote the CSI obtained from the linear prediction algorithm (9) and c the actual complex gain at a certain subcarrier. Hence $\widehat{c}$ and c are jointly complex Gaussian and their amplitudes $\widehat{a}$ and α are both Rayleigh distributed. For subcarrier k, let $P(= E[|a[n,k]|^2](5))$ denote the transmitted signal power of the complex $M(i)$-QAM symbol that is determined by allocation algorithm. (Note the sum of the allocated powers for all subcarriers does not exceed the total power constraint P_{total}). Assume each subcarrier has the same noise power N_0 (see (5)). The $SNR_{\gamma M(i)}(= P/N_0)$ required to employ $M(i)$-QAM modulation given the predicted channel gain $\widehat{a}$ at the k^{th} subcarrier can be found by numerical search to meet the bit error rate constraint BER_c:

$$BER_c = \int_0^\infty BER_{M(i)}\left(\gamma_{M(i)} x^2\right) p_{\alpha|\widehat{\alpha}}(x)\, dx \qquad (14)$$

where $BER_{M(i)}$, calculated from [26], is the bit error rate for the M-QAM modulation on the AWGN channel, and $p_{\alpha|\widehat{\alpha}}$, the conditional probability density function of α given $\widehat{\alpha}$, is given by [6], [20]:

$$p(\alpha|\widehat{\alpha}) = \frac{2a}{(1-\rho)\,\Omega} I_0\left(\frac{2\sqrt{\rho}\alpha\widehat{\alpha}}{(1-\rho)\sqrt{\Omega\widehat{\Omega}}}\right) \exp\left(-\frac{1}{1-\rho}\left(\frac{\alpha^2}{\Omega} + \frac{\rho\widehat{\alpha}^2}{\Omega}\right)\right) \qquad (15)$$

where the parameter ρ is the correlation coefficient between

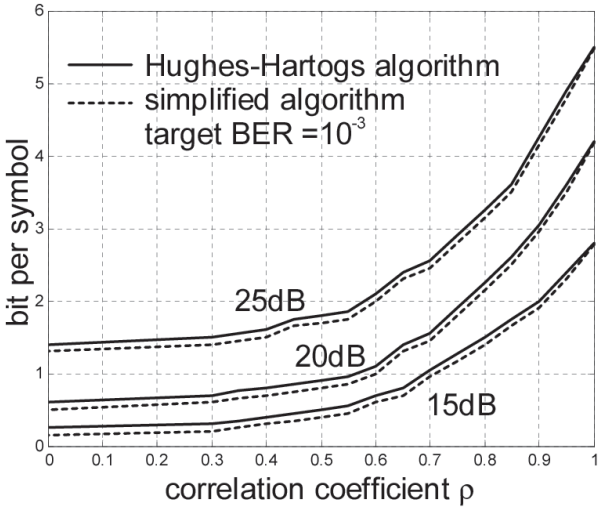

Fig. 6. BPS vs, ρ for different SNR for adaptive OFDM system, SNR = P_{total}/KN_0.

α^2 and $\widehat{\alpha}^2$:

$$\rho = \frac{Cov\left(\alpha^2, \widehat{\alpha}^2\right)}{\sqrt{Var\left(\alpha^2\right) Var\left(\widehat{\alpha}^2\right)}} \tag{16}$$

and $\Omega = E\left\{\alpha^2\right\} = 1$, $\widehat{\Omega} = E\left\{\widehat{\alpha}^2\right\}$ and I_0 is the 0^{th} order modified Bessel function. Once the $\gamma_{M(i)}$ are calculated for each modulation level and each subcarrier, they are used to implement the Hughes-Hartogs and the simplified algorithm in the presence of imperfect CSI. The only difference in the implementation (relative to the perfect CSI case) is that the SNR $\gamma_{M(i)}$ in (14) is used in place of the ideal SNR required to achieve the BER with $M(i)$-QAM [26].

From (14) and (15), the required power to employ M-QAM modulation is dependent on the parameters ρ and $\widehat{\Omega}$. Nevertheless, it can be shown that when the estimated power $\widehat{\Omega}$ is scaled to 1, the performance of adaptive modulation is not affected, so throughout the paper, we employ this normalization for simplicity, and the performance depends only on the parameter ρ. The average bits per symbol (BPS) vs. the correlation coefficient ρ vs. the average SNR constraint given by $P_{total}/(KN_0)$ is shown in Fig. 6, where we assume each subcarrier has the same prediction accuracy ρ. The BER constraint for each subcarrier is 10^{-3}. The correlation $\rho = 1$ corresponds to perfect prediction, while $\rho = 0$ represents the worst case when the BPS of the adaptive modulation converges to that of the non-adaptive M-QAM for given SNR and bit error rate constraint BER$_c$. It is observed that the simplified algorithm is near-optimal when ρ is close to 1 and has performance loss less than 0.1 BPS for $\rho << 1$ compared with the optimal Hughes-Hartogs algorithm.

To reduce signaling overhead complexity and to accommodate the hardware constraints, we investigated the performance of the LRP and adaptive modulation where the modulation level remains constant over ten or even hundreds of symbols, and block loading, where neighboring subcarriers use the same modulation level in [20]. Furthermore, the system can be simplified by assigning fixed power for each modulation level as in discrete rate discrete power adaptive modulation [3].

B. Feedback Load Reduction

In addition to adaptive channel loading, many other adaptive transmission techniques can be implemented in a wireless OFDM system. These include adaptive coding, adaptive transmitter antenna diversity and interference suppression at the transmitter. In a closed-loop system the transmitter obtains the CSI for this adaptive transmission from the receiver via a feedback channel. The prediction can be performed either at the transmitter or at the receiver, depending on the complexity and feedback requirements. In some adaptive transmission applications, it might be feasible to predict at the receiver and feed back the CSI parameters derived from the predictions (e.g., the modulation level and power for adaptive modulation). In other adaptive transmission techniques (e.g., adaptive antenna diversity), predicted channel gains are often required at the transmitter, and thus channel gains have to be fed back, independent of whether the prediction is accomplished at the transmitter or receiver. This CSI is required for all subcarriers. It is desirable to minimize this feedback load since it consumes resources that would otherwise be used for data. At the same time, the feedback signal should carry enough information so the transmitter can perform reliable adaptive transmission.

In this section, we explore several methods for reducing the feedback of the OFDM signal vectors $\underline{\widetilde{H}} = [\widetilde{H}(n,1) \ldots \ \widetilde{H}(n,K)]$ (the estimates in (6) sampled at low rate) while insuring accurate reconstruction at the transmitter. Since $\underline{\widetilde{H}}$ is modeled as Gaussian, the estimates of the reconstructed signals are formed as linear combinations of the signals that are fed back. The performance is measured by the correlation coefficient ρ (16) between the reconstructed signals and actual CSI for each subcarrier and is dependent on the feedback density (FD) given by:

$$FD = \frac{\textit{Number of fed back symbols}}{\textit{Total number of subcarriers}} \tag{17}$$

We also define the normalized feedback density (NFD) as:

$$NFD = \frac{FD}{\textit{normalized subcarrier frequency separation}} \tag{18}$$

The first reduced feedback method utilizes Karhunen-Loeve (K-L) low rank modeling [24]. The K-L method requires the knowledge of the eigen-vectors of the correlation matrix of the feedback signal vector at both the transmitter and the receiver. For the fading channel characterized as WSSUS with the frequency domain correlation function $R_f(\Delta_f)$ (3), the $K \times K$ correlation matrix defined as $\underline{\mathbf{R}} = E[\underline{\widetilde{H}}\underline{\widetilde{H}}^H]$ is computed from $R_f(\Delta_f)$. Let $q_1, q_2 \ldots q_K$ be the eigenvectors associated with K eigenvalues $(\lambda_1 > \lambda_2 > \ldots \lambda_K)$ of the matrix $\underline{\mathbf{R}}$. The vector $\underline{\widetilde{H}}$ can be expressed as a linear combination of these eigenvectors as

$$\underline{\widetilde{H}} = \sum_{i=1}^{K} \nu_i \underline{q_i} \tag{19}$$

The coefficients ν_i of the expansion are zero-mean, uncorrelated random variables defined by the inner product $\nu_i = \underline{q_i}^H \underline{\widetilde{H}}$. From (19), we can reduce the feedback by sending only ν_i for $i = 1 \ldots m$ and to approximately reconstruct $\underline{\widetilde{H}}$ at the transmitter by $\underline{\widetilde{H}} = \sum_{i=1}^{m} \nu_i \underline{q_i}$, $m < K$. Hence the feedback density (17) is m/K. The reconstruction error vector

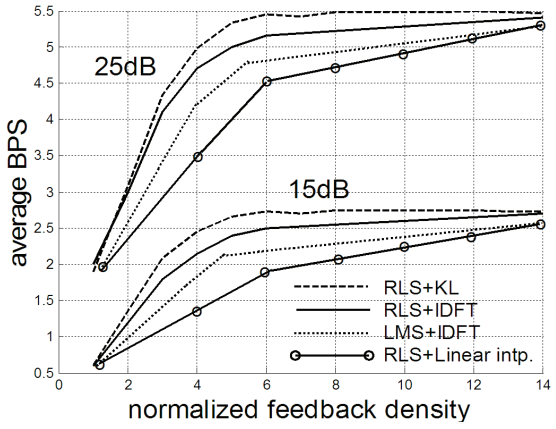

Fig. 7. Performance comparison for different reduced feedback methods.

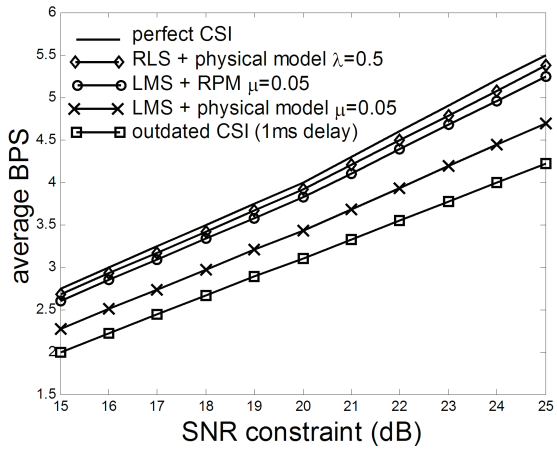

Fig. 8. Comparison of average BPS performance for adaptive OFDM aided by different prediction methods for the RPM and physical model.

is defined by $\underline{e} = \underline{\widehat{H}} - \underline{\widetilde{H}}$. It can be shown that the mean square error is $E[\underline{e}^H \underline{e}] = \sum_{i=m+1}^{K} \lambda_i$, [24], Thus, accurate reconstruction is achieved if the eigenvalues $\lambda_{m+1}, \ldots, \lambda_K$ are very small. The number of significant eigenvalues depends on the rms delay spread, frequency separation, and the number of the subcarriers and can be shown to be small for typical OFDM channels.

Since computation of the basis of the $K - L$ low rank modeling [24] requires the knowledge of the channel correlation function, we propose to utilize the discrete Fourier basis and transform the CSI $\underline{\widetilde{H}}$ using the inverse discrete Fourier transform (IDFT). This choice is meaningful since the IDFT corresponds to the channel impulse response. The K-point IDFT of the CSI $\underline{\widetilde{H}}$ is given by $I(m) = \frac{1}{K}\sum_{k=1}^{K} \widetilde{H}(n,k)\exp\{j2\pi(k-1)/(K(m-1))\}, m = 1 \ldots K$. The samples $I(m)$ are relatively small for $m > \tau_{max}\Delta f K$, where τ_{max} is the maximum excess delay, Δf is the subcarrier frequency separation and K is the total number of subcarriers. This suggests that $FD > \tau_{max}\Delta f$ is required to obtain good performance. In OFDM channels, K is chosen much larger than $\tau_{max}\Delta f K$ (i.e. $\Delta f << 1/\tau_{max}$) to avoid intersymbol interference (ISI). Thus, the transformed signal $I(m)$ can be truncated, fed back to the transmitter and reconstructed by the DFT.

Alternatively, we can directly feed a subset of the CSI samples $\underline{\widetilde{H}}$ back to the transmitter without any transformation. The feedback signals are sampled uniformly over the entire frequency band. The original signal can be reconstructed by interpolation or using the MMSE criterion if the correlation functions are known. This method results in different accuracy for each subcarrier and increases the complexity of the bit and power allocation. We call this method *direct reduced feedback with linear interpolation*.

For all reduced feedback methods, the performance depends on the rms delay spread and the subcarrier frequency separation. Furthermore, it will affect the LRP accuracy and hence the performance of the adaptive loading algorithm. Therefore, there is a tradeoff between the feedback load and the performance of the AOFDM.

C. Numerical Simulation for AOFDM and Reduced Feedback

We use the RPM and the physical model to validate the performance of our AOFDM system aided by the LRP with reduced feedback. The target BER for the adaptive OFDM system is 10^{-3}. The system parameters are described in Section III. The performance of the prediction algorithm using reduced feedback methods for different SNR constraint is shown in Fig. 7 for the RPM. The signals are reduced by several methods and fed back to the transmitter. The transmitter predicts the next OFDM channel coefficient vector, estimates the prediction accuracy ρ, and performs adaptive bit and power allocation. It is shown that for the KL method, performance is near optimal when the NFD (18) is larger than 6. This implies from (18) that only 4 symbols need to be fed back for a 128-subcarrier OFDM system with normalized subcarrier frequency separation 0.005. For the practical IDFT method, the performance loss is less than 0.5 BPS at NFD = 6 as opposed to 1 BPS loss for the direct reduced feedback method with linear interpolation "RLS+Linear intp" curve). The choice of the feedback density provides a trade-off between the feedback load and the prediction accuracy, and hence the transmission rate.

The average BPS of the AOFDM for different prediction algorithms for the RPM and physical channel models is plotted in Fig. 8. Perfect feedback is assumed. Note that the spectral efficiency will be slightly less than the average BPS due to the guard interval used to eliminate the ISI. Comparison reveals that the RLS has better performance than the LMS algorithm for the RPM and non-stationary physical model. The performance of the RLS algorithm for the RPM is near-optimal (not shown), whereas the loss is less than 0.5 dB for the physical model compared to the perfect knowledge of CSI. Note that although we employ high effective observation SNR = 80 dB for our simulation, we observed the achievable BPS is approximately the same as long as the effective observation SNR $>$ 40 dB. The performance of the AOFDM using the outdated CSI samples (1 ms delay) without prediction for the RPM is also shown in Fig. 8. Calculation of thresholds for this case was studied in [6]. We found that even very small delay causes significant loss of the bit rate for fast vehicle

speeds when accurate LRP is not utilized. For example, the delay of 1ms for f_{dm} =100Hz corresponds to the correlation coefficient $\rho = 0.9$ in (16). This results in the bit rate loss about 1.5 bits/symbol for the target BER = 10^{-3} and the SNR per symbol = 25 dB assuming stationary Rayleigh fading channel, while prediction results in near-optimal BPS for non-stationary channels.

V. Conclusions

A mobile radio AOFDM system aided by the long-range prediction and reduced feedback was investigated. A realistic physical model and a stationary random phase model were employed to validate the prediction performance. The simulation results demonstrated that accurate long range prediction is required to achieve the potential of adaptive OFDM system for fast vehicle speeds and realistic delays. Specifically, the RLS LRP that uses combined observations of all carriers was shown to enable adaptive loading for the physical model and practical OFDM parameters. Finally, several methods were developed to reduce the feedback load, and it was shown that the IDFT method offers significant feedback load reduction while maintaining near-optimal spectral efficiency.

References

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," IEEE Commun. Mag., pp. 5–14, May 1990.

[2] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency-division multiplexing using the discrete Fourier transform," IEEE Trans. Commun., vol. 19, no. 5, pp. 628–634, Oct. 1971.

[3] A. J. Goldsmith and S. G. Chua, "Variable-rate variable-power MQAM for fading channels," IEEE Trans. Commun., vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[4] A. J. Goldsmith and S. G. Chua, "Adaptive coded modulation for fading channels," IEEE Trans. Commun., vol. 46, no. 5, pp. 595–601, May 1998.

[5] T. Ue, S. Sampei, N. Morinaga, and K. Hamaguchi, "Symbol rate and modulation level-controlled adaptive modulation/TDMA/TDD system for high-bit-rate wireless data transmission," IEEE Trans. Veh. Technol., vol. 47, No. 4, pp. 1134–1147, Nov. 1998.

[6] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated channel estimates," IEEE Trans. Commun., vol. 47, no. 6, pp. 845–855, June 1999.

[7] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[8] D. Hughes-Hartogs. Ensemble modem structure for imperfect transmission media. U S Patent 4,833,706, issued July 1987.

[9] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," IEEE Trans. Commun., vol. 48, no. 1, pp. 23–27, Jan. 2000.

[10] A. Czylwik, "Adaptive OFDM for wideband radio channels," in *Proc. IEEE GLOBECOM*, Nov. 1996, pp. 713–718.

[11] J. Campello, "Practical bit loading for DMT," in *Proc. IEEE ICC*, June 1999, vol. 2, pp. 801–805.

[12] T. Eyceoz, A. Duel-Hallen, and H. Hallen, "Deterministic channel modeling and long range prediction of fast fading mobile radio channels," IEEE Commun. Lett., vol. 2, no. 9, pp. 254–256, Sep. 1998.

[13] T. Eyceoz, S. Hu, and A. Duel-Hallen, "Performance analysis of long range prediction for fast fading channels," in *Proc. 33rd Conf. on Inform. Sciences and Systems*, Mar. 1999, vol. 2, pp. 656–661.

[14] S. Hu, A. Duel-Hallen, and H. Hallen, "Long range prediction makes adaptive modulation feasible for realistic mobile radio channels," in *Proc. 34rd Annual Conf. on Inform. Sciences and Systems*, Mar. 2000, vol. 1, pp. WP4-7–WP4-13.

[15] A. Duel-Hallen, S. Hu, and H. Hallen, "Long range prediction of fading signals: Enabling adaptive transmission for mobile radio channels," IEEE Signal Processing Mag., vol. 17, no. 3, pp. 62–75, May 2000.

[16] H. Hallen, S. Hu, M. Lei, and A. Duel-Hallen, "A physical model for wireless channels to understand and test long range prediction of flat fading," in *Proc. Wireless 2001*, July 2001.

[17] H. Hallen, A. Duel-Hallen, S. Hu, T. S. Yang, and M. Lei, "A physical model for wireless channels to provide insights for long range prediction," in *Proc. IEEE MILCOM*, Oct. 2002, vol. 1, pp. 627–631.

[18] T. S. Yang, A. Duel-Hallen, and H. Hallen, "Long range fading prediction to enable adaptive transmission at another carrier," in *Proc. IEEE SPAWC*, June 2003, pp. 195–199.

[19] Tung-Sheng Yang, A. Duel-Hallen, and H. Hallen, "Reliable adaptive modulation aided by observations of another fading channel," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 605–611, Apr. 2004.

[20] Tung-Sheng Yang, "Performance Analysis of Adaptive Transmission Aided by Long Range Channel Prediction for Realistic Single-and Multi-Carrier Mobile Radio Channels." PhD thesis, North Carolina State University, 2004.

[21] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[22] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2000.

[23] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid diversive fading channel," *IEEE Trans Commun.*, vol. 46, pp. 902–915, Apr. 1998.

[24] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[25] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[26] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, pp. 1074–1080, July 2002.

[27] M. Lei and A. Duel-Hallen, "Long range channel prediction and adaptive transmission for frequency hopping communications," in *Proc. 41st Annual Allerton Conference on Communications, Control, and Computing*, Oct. 2003, pp. 1-10.

[28] A. Arredondo, K. R. Dandekar, and Guanghan Xu, "Vector channel modeling and prediction for the improvement of downlink received power," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1121–1129, July 2002.

[29] M. F. Pop and N. C. Beaulieu, "Limitations of sum-of-sinusoids fading channel simulators," *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 699–708, Apr. 2001.

[30] T. Ekman and G. Kubin, "Nonlinear prediction of mobile radio channels: Measurements and MARS modeling designs," in *Proc. IEEE Internatinoal Conference on Acoustics, Speech and Signal Processing*, Mar. 1999, pp. 2667–2670.

[31] J. K. Hwang and J. H. Winters, "Sinusoidal modeling and prediction of fast fading processes," in *Proc. IEEE GLOBECOM*, Nov. 1998, pp. 892–897.

[32] J. B. Anderson, J. Jensen, S. Holdt Jensen, and F. Frederiksen, "Prediction of future fading based on past measurements," in *Proc. IEEE VTC*, Sep. 1999, vol. 1, pp. 151–155.

[33] S. Falahati, A. Svensson, T. Ekman, and M. Stenard, "Adaptive modulation systems for predicted wireless channels," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 307–316, Feb. 2004.

[34] G. E. Oien, H. Holm, and K. J. Hole, "Impact of channel prediction on adaptive coded modulation performance in Rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 758–769, May 2004.

[35] Z. Song, K. Zhang, and Y. Guan, "Statistical adaptive modulation for QAM-OFDM systems," in *Proc. IEEE GLOBECOM*, Nov. 2002, vol. 1, pp. 706–710.

[36] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO OFDM based on partial channel state information," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 202–213, Jan. 2004.

[37] S. Ye, R. Blum, and L. Cimini, "Adaptive modulation for variable rate OFDM systems with imperfect channel information," in *Proc. IEEE VTC,* May 2002, vol. 2, pp. 767–771.

[38] T. S. Yang, A. Duel-Hallen, and H. Hallen, "Enabling adaptive OFDM for mobile radio channels," *Proc. IEEE MILCOM*, Oct. 2004, vol. 2, pp. 704–710.

Alexandra Duel-Hallen received B.S. degree in mathematics from Case Western Reserve University in 1982, M.S. degree in Computer, Information and Control Engineering from the University of Michigan in 1983, and Ph.D. in Electrical Engineering from Cornell University in 1987. During 1987-1990 she was a Visiting Assistant Professor at the School of Electrical Engineering, Cornell University, Ithaca, NY. In 1990-1992, she was with the Mathematical Sciences Research Center, AT&T Bell Laboratories, Murray Hill, NJ. She is a Professor at the Department of Electrical and Computer Engineering at North Carolina State University, Raleigh, NC, which she joined in January 1993. From 1990 to1996, Dr. Duel-Hallen was Editor for Communication Theory for the IEEE TRANSACTIONS ON COMMUNICATIONS. Dr. Duel-Hallen's current research interests are in wireless communications. She was selected to appear as Highly Cited Researcher on ISIHighlyCited.Com in Computer Science category. Her paper was selected for the *IEEE Communications Society 50th Anniversary Journal Collection* as one of 41 key papers in physical and link layer areas, 1952-2002.

Hans Hallen received his B.S. degree in Engineering Physics from Cornell University in 1984, and his M.S. and Ph.D. degrees in applied physics from Cornell University in 1986 and 1991, respectively. At Cornell, he fabricated and tested sub-micron-scale Josephson junctions, then conducted photoemission studies of atomic ordering near interfaces of plasma-oxidized silicon. He then studied nano-scale modifications of metal surfaces with a scanning tunneling microscope, determining the mechanism for atomic movement. During 1991-1993, he was with the Physical Research Laboratory, AT&T Bell Laboratories, Murray Hill, NJ where he developed the first scanning Hall probe microscope, and used it to study high temperature superconductivity and vortex proagation in small structures. He joined the North Carolina State University Physics Department in 1993, and is currently an Associate Professor. He led the group that produced the first near-field Raman images, and identified new physics in nanoscale optical spectroscopy. He studied carrier dynamics in nanoscale silicon. His results identifying electron induced motion of atoms in conductors such as Au and YBCO have led to a novel view of transport of few eV electrons in metals. He has active projects in nanoscale characterization with scanning proximal probe microscopes utilizing optical, electrical, and photoemission probes. He is investigating oriented molecular deposition project, scanning nano-transport microscopy, surface modification for functionality, and has active research efforts in microfluidics, biophysics, lidar, and wireless communications.

Tung-Sheng Yang received the B.S. degree in electrical engineering from the National Tsing Hua University, Hsinchu, Taiwan, in 1998 and the M.S. and Ph.D. degrees in electrical engineering from North Carolina State University, Raleigh, in 2000 and 2004, respectively. From 2000 to 2004, he was a Research Assistant at North Carolina State University. In 2005, he joined the Mobile Devices Technology in Taiwan and is currently a member of the technical staff. His research interests include transmitter and receiver design for fading mobile radio channels, estimation and prediction of fading channels, and adaptive transmission for wireless mobile systems.