Adaptive OFDM Systems With Imperfect Channel State Information

Sigen Ye, *Member, IEEE*, Rick S. Blum, *Fellow, IEEE*, and Leonard J. Cimini, Jr., *Fellow, IEEE*

***Abstract*— Adaptive modulation has been shown to have significant benefits for high-speed wireless data transmission when orthogonal frequency division multiplexing (OFDM) is employed. However, accurate channel state information (CSI) is required at the transmitter to achieve the benefits. Imperfect CSI arises from noisy channel estimates, which may also be outdated due to a delay in getting the CSI to the transmitter. In this paper, we study adaptive OFDM with imperfect CSI for the uncoded variable bit rate case, where a target bit error rate is set. A loading algorithm based on the statistics of the real channel is proposed. Performance results in terms of the average spectral efficiency are provided for adaptive OFDM systems when there is noisy channel estimation or CSI delay. The use of multiple estimates is then proposed to improve the performance. It is shown that multiple estimates from different frequencies or times can enhance the performance significantly, which enables the system to tolerate larger errors in channel estimation or longer delay in CSI.**

***Index Terms*— Adaptive modulation, imperfect channel state information, OFDM, variable bit rate.**

I. Introduction

ORTHOGONAL frequency division multiplexing (OFDM) is an attractive technique for combating the effects of delay spread in high-speed wireless data transmission [1]. By transmitting multiple data streams in parallel over different low-rate subchannels, intersymbol interference (ISI) can be reduced significantly. If each subchannel is narrow enough such that the multipath fading can be characterized as flat, the need for equalization can be eliminated. The popularity of OFDM is evident by its use in recent standards for digital audio and video broadcasting (DAB/DVB) [2], [3], asymmetric digital subscriber line (ADSL) [4], and wireless local area network (WLAN) [5] applications.

In conventional OFDM, the same modulation scheme is employed on all subchannels. However, the bit error rate can be dominated by a small set of severely faded subchannels. One way to improve the performance is to use adaptive modulation, where the modulation used on each subchannel is dependent on the multipath channel response for that subchannel. Adaptive modulation takes advantage of the frequency selectivity and time variation by adapting the transmitted signal to match the multipath channel, which is sometimes called "adaptive loading". Both the power and the data rate in each subchannel can be adapted. The benefits have been described in previous works [6]-[12]. Moreover, adaptive OFDM has already been used in ADSL [4]. To achieve the performance advantages of adaptive modulation, however, accurate receiver channel state information (CSI) is required at the transmitter. In wireless communications, because the channel is noisy and time-varying, the estimated channel response may be noisy and outdated. Some researchers [11]-[19] have studied the performance when imperfect or partial CSI is available, and some of them have proposed possible solutions for the imperfect CSI case. Adaptive transmission based on imperfect or partial CSI for multiple-antenna OFDM systems has also been studied, for example, in [20]-[23].

Adaptive OFDM can be employed for two different kinds of services: constant bit rate (CBR) and variable bit rate (VBR). Typically an average bit error rate (BER) requirement is set for the VBR services. In [11], the performance degradation from both Doppler and channel estimation error, in terms of throughput and BER, was investigated when a subband approach was employed. The impact of imperfect CSI was also studied in [15] and [16] for the CBR case. A low-complexity ordered subcarrier selection algorithm was proposed in [12], and the robustness of this algorithm against Doppler was studied. In [13], an algorithm underloading all subchannels is proposed to combat the BER degradation which results from the channel errors in coherent detection at the receiver. In [14], channel prediction is employed to mitigate the impact of outdated CSI for the CBR case. A statistical adaptive modulation scheme based on the long-term channel statistics (partial CSI) was proposed in [18]. In [19], optimal power loading algorithms based on average and outage capacity criteria were pursued when partial CSI was available at the transmitter.

In this paper, we study adaptive OFDM for the uncoded VBR case. The goal is to investigate how CSI errors due to noise and delay impact the system performance and how we can improve the performance. We quantify the performance degradation due to CSI errors and propose the use of multiple estimates to mitigate the effect of CSI errors. Section II introduces the wireless channel model and the OFDM system

Manuscript received January 5, 2005; revised August 13, 2005; accepted December 9, 2005. The associate editor coordinating the review of this paper and approving it for publication was J. Andrews. This material is based on research supported by the Air Force Research Laboratory under agreement No. F49620-03-1-0214, and by the National Science Foundation under Grant No. CCR-0112501.

Sigen Ye was with the Electrical and Computer Engineering Department, Lehigh University, Bethlehem, PA 18015 USA. She is now with Lucent Technologies, Whippany, NJ 07981 USA (e-mail: sye@lucent.com).

Rick S. Blum is with the Electrical and Computer Engineering Department, Lehigh University, Bethlehem, PA 18015 USA (e-mail: rblum@ece.lehigh.edu).

Leonard J. Cimini, Jr. is with the Electrical and Computer Engineering Department, University of Delaware, Newark, DE 19716 USA (e-mail: cimini@ece.udel.edu).

Digital Object Identifier 10.1109/TWC.2006.05004

model, together with the system parameters used in the simulations. The average spectral efficiency of non-adaptive and adaptive OFDM with perfect channel information is derived in Section III. Section IV presents a new loading algorithm and discusses the performance of adaptive OFDM with imperfect CSI. The effects of both channel estimation error and outdated CSI are considered. In Section V, the use of multiple estimates is proposed to improve the performance, and simulation results are provided. Finally, conclusions are given in Section VI.

Throughout this paper, a continuous constellation size is assumed (that is, the number of bits per symbol can range from 0 to ∞), which provides easily interpreted results and gives an upper bound on achievable performance. Power adaptation is not considered because it increases implementation complexity and has been shown to have a very limited effect on performance [14] [24]. For notation, we use $(\cdot)^T$ and $(\cdot)^\dagger$ to denote transpose and transpose conjugate, respectively.

II. Channel Model and System Parameters

We assume the wide-sense stationary uncorrelated scattering (WSSUS) model for the frequency-selective mobile radio channel. This has been widely used in the literature as a physically reasonable approximation (for example, see [25]-[27]), and the material presented will be based on this model. Mathematically, the baseband impulse response of the channel can be described by

$$h(t,\tau) = \sum_{\ell=1}^{L} \gamma_\ell(t)\delta(\tau - \tau_\ell), \tag{1}$$

where L is the total number of paths, τ_ℓ is the delay of the ℓth path, and $\gamma_\ell(t)$ is the corresponding complex amplitude. Moreover, the $\gamma_\ell(t)$'s are wide-sense stationary (WSS) complex Gaussian processes and independent for different paths with $E\{|\gamma_\ell(t)|^2\} = q_\ell^2$. The channel is normalized such that $\sum_{\ell=1}^{L} q_\ell^2 = 1$. It is further assumed that the $\gamma_\ell(t)$'s have the same normalized correlation function $r_t(\Delta t)$, which means that

$$E\{\gamma_\ell(t+\Delta t)\gamma_\ell^*(t)\} = q_\ell^2 r_t(\Delta t). \tag{2}$$

For the simulations in this paper, the time-varying fading channel is generated based on Jakes' model [28], in which the correlation function is given by

$$r_t(\Delta t) = J_0(2\pi f_D \Delta t). \tag{3}$$

Here $J_0(\cdot)$ is the zeroth-order Bessel function, and f_D is the maximum Doppler frequency.

Let B be the total bandwidth of the system and K the total number of subchannels. This results in a subchannel spacing of $\Delta f = B/K$ and an OFDM block length of $T = 1/\Delta f = K/B$, assuming that the guard interval (cyclic prefix) is so small that it can be neglected. Then the sampling interval is given by $T_s = T/K = 1/B$. We assume that each subchannel is narrow enough so that it experiences flat fading. Let $H[n,k]$ denote the frequency response of the kth tone in the nth OFDM block, given by

$$H[n,k] = \sum_{\ell=1}^{L} \gamma_\ell(nT) W_K^{kp_\ell}, \tag{4}$$

where $W_K = \exp\{-j2\pi/K\}$ and $p_\ell = \tau_\ell/T_s$. Then the following properties regarding $H[n,k]$ follow directly.

Property 1: $H[n,k]$ is a complex Gaussian random variable with zero mean and unit variance for any n and k.

Property 2: The correlation function of the frequency response at different times and frequencies is given by

$$\begin{aligned} &E\{H[n+\Delta n, k+\Delta k]H^*[n,k]\} \\ &= E\left\{\sum_{\ell=1}^{L}\sum_{m=1}^{L} \gamma_\ell((n+\Delta n)T)\gamma_m^*(nT) W_K^{(k+\Delta k)p_\ell} W_K^{-kp_m}\right\} \\ &= r_t(\Delta n \cdot T) \sum_{\ell=1}^{L} q_\ell^2 W_K^{\Delta k \cdot p_\ell}. \end{aligned} \tag{5}$$

Note that the correlation function reduces to $r_t(\Delta n \cdot T)$ for $\Delta k = 0$, and reduces to $\sum_{\ell=1}^{L} q_\ell^2 W_K^{\Delta k \cdot p_\ell}$ for $\Delta n = 0$. In other words, the correlation function is separated as the multiplication of the time-domain correlation (depending on Doppler) and the frequency domain correlation (depending on the multi-path delay profile).

Assuming the system parameters are chosen such that no intercarrier or interblock interference exists, the received signal is

$$Z[n,k] = H[n,k]\, S[n,k] + W[n,k], \tag{6}$$

where $S[n,k]$ is the transmitted signal, and $W[n,k]$ is additive white Gaussian noise, which is independent for different n and k. Let E_s be the symbol energy at the transmitter, and $N_0/2$ be the variance of the real/imaginary part of the Gaussian noise $W[n,k]$.

Two different sets of values for the system parameters will be used in the simulation results that are presented in the following sections:

- Indoor system: The parameters are similar to those for IEEE 802.11a wireless local area networks (WLANs) [5] and suitable for indoor environments. A total bandwidth of 20 MHz is divided into 64 subchannels, resulting in a block length of 3.2 μs. The carrier frequency is 5 GHz. An exponential delay profile is assumed with rms and maximum delay spreads of 0.2 μs and 0.8 μs, respectively.
- Outdoor system: The parameters correspond to an extension of IEEE 802.11a to wide-area outdoor environments [29]. A total bandwidth of 5 MHz is divided into 1024 subchannels, resulting in a block length of 204.8 μs. The carrier frequency is 2 GHz. An exponential delay profile is assumed with rms and maximum delay spreads of 5 μs and 20 μs, respectively.

The average spectral efficiency is used as the performance measure. Although we assume the overhead due to the guard interval is ignored, the guard interval would reduce the spectral efficiency by about 10-20% in practical OFDM systems. This assumption, however, does not affect the comparison between different OFDM systems.

III. Spectral Efficiency of Adaptive OFDM With Perfect CSI

There are two kinds of adaptive OFDM systems: in the first, the modulation varies over both the subchannels and time;

in the second, the modulation varies over time but all the subchannels employ the same modulation at a given time. We restrict our attention to the first kind only.

Assume QAM is employed for each subchannel, and $\beta[n,k]$ bits/symbol are sent for the kth tone in the nth block. We investigate the spectral efficiency under a constrained average BER requirement. According to [24], given the channel frequency response $H[n,k]$, the instantaneous BER for the kth tone in the nth block can be approximated by

$$P_e[n,k] \approx c_1 \exp\left\{ -\frac{c_2 \frac{E_s}{N_0} |H[n,k]|^2}{2^{\beta[n,k]} - 1} \right\}, \qquad (7)$$

where $c_1 = 0.2$, $c_2 = 1.6$.

A. Non-Adaptive Modulation

Consider first the case of non-adaptive modulation, where $\beta[n,k] = \beta$ is a constant for all n and k. Since $H[n,k]$ is a complex Gaussian random variable and all $H[n,k]$'s have identical distributions (Property 1), the overall average BER becomes

$$\bar{P}_e = E_{H[n,k]}\left\{P_e[n,k]\right\} \approx \frac{c_1}{\frac{c_2 E_s/N_0}{2^\beta - 1} + 1}. \qquad (8)$$

Assume P_{tar} is the target average BER. Then, by inverting (8) with $\bar{P}_e = P_{tar}$, the maximum number of bits that can be transmitted given the average BER constraint is

$$\beta = \log_2\left[\frac{c_2 \frac{E_s}{N_0}}{\frac{c_1}{P_{tar}} - 1} + 1\right]. \qquad (9)$$

The spectral efficiency (number of bits per second per Hz) is equal to β, under the assumption that the symbol interval is the reciprocal of the subchannel bandwidth. The spectral efficiency of non-adaptive OFDM for $P_{tar} = 10^{-3}$ is plotted in Fig. 1 (dashed line).

B. Ideal Adaptive OFDM

For adaptive OFDM, different modulation schemes are used for different subchannels and the assignments vary over time. In this subsection, we assume that perfect knowledge of the receiver channel information is available at the transmitter. One way to choose the modulation schemes to achieve the target BER is to set the instantaneous BER $P_e[n,k]$ equal to P_{tar}. While this is just one way to achieve the constraint on average BER, in a previous study [24], fixing the instantaneous BER was shown to have negligible loss compared to the optimum approach for single-carrier transmission. Using this approach, the number of bits transmitted in each subchannel can be derived by inverting (7) to obtain

$$\beta_{ideal}[n,k] = \log_2\left[\frac{c_2 \frac{E_s}{N_0} |H[n,k]|^2}{\ln \frac{c_1}{P_{tar}}} + 1\right]. \qquad (10)$$

Therefore, the average spectral efficiency R_{ideal} is

$$\begin{aligned} R_{ideal} &= E_{H[n,1],\cdots,H[n,K]}\left\{\frac{1}{K}\sum_{k=1}^{K}\beta_{ideal}[n,k]\right\} \qquad (11)\\ &= E_{H[n,k]}\{\beta_{ideal}[n,k]\}. \qquad (12)\end{aligned}$$

The last equality holds because of the identical statistics of $H[n,k]$ for all k (Property 1). Therefore, the average spectral efficiency R_{ideal} does not depend on the system parameters or the specific power delay profile of the WSSUS mobile radio channel. However, it is important to remember that the guard interval is ignored in R_{ideal}. Since the guard interval length depends on both the power delay profile and the system parameters, so does the real average spectral efficiency. And this comment also applies to the results presented later.

The average spectral efficiency of adaptive OFDM for a target BER of 10^{-3}, obtained through a Monte Carlo simulation, is shown in Fig. 1 (solid line labeled "ideal"). The results indicate that a significant improvement in spectral efficiency, or equivalently bit rate, is potentially possible by matching the modulation scheme in each subchannel to the corresponding multipath channel response. For example, the average spectral efficiency is increased from 0.87 b/s/Hz to 4.26 b/s/Hz when $E_s/N_0 = 20$ dB. On the other hand, we also observe a 14-dB power gain achieved by using adaptive OFDM when the average spectral efficiency is 2 b/s/Hz.

IV. Spectral Efficiency of Adaptive OFDM With Imperfect CSI

In practice, it is impossible to obtain perfect channel information due to noisy channel estimation and the unavoidable delay between when channel estimation is performed and when the estimation result is used for actual transmission. We will consider imperfect CSI in this section and study the performance degradation in spectral efficiency resulting from these CSI errors. We assume that the receiver has perfect CSI for coherent detection so that the effect of CSI errors on adaptive loading at the transmitter can be separated from that on coherent detection at the receiver.

A. New Loading Algorithm

If perfect CSI is available, we fix the instantaneous BER and determine the number of bits transmitted using (10). However, with errors in CSI, this loading method will degrade the performance and will not satisfy the BER requirement. The proposed approach uses statistical information about the CSI errors to exactly maintain the required average BER level and gives the resulting performance in terms of spectral efficiency.

Suppose the estimated channel gain $H'[n,k]$ is the only known information about the current CSI for the kth tone in the nth block, and $\beta[n,k]$ is computed based on this value of $H'[n,k]$. Since the instantaneous BER, $P_e[n,k]$, depends on the value of the true channel $H[n,k]$, which is assumed unknown, it is not possible to fix $P_e[n,k]$ to be the target value. However, we can define the average BER given $H'[n,k]$ for the kth tone in the nth block as

$$\begin{aligned} \bar{P}_e[n,k] &= E_{H[n,k]\,|\,H'[n,k]}\left\{P_e[n,k]\right\} \qquad (13)\\ &= E_{|H[n,k]|\,|\,H'[n,k]}\left\{P_e[n,k]\right\}, \qquad (14)\end{aligned}$$

where the expectation is evaluated over $H[n,k]$ or $|H[n,k]|$. The second equality holds because $P_e[n,k]$ depends on $H[n,k]$ only through $|H[n,k]|$. Under certain circumstances, as we will discuss in the following subsections, the conditional probability density function (pdf) of $H[n,k]$ or $|H[n,k]|$

given $H'[n,k]$ can be estimated. Assuming this conditional pdf is known, $\bar{P}_e[n,k]$ can be calculated, and it becomes a function of $H'[n,k]$. The proposed algorithm employs the $\beta[n,k]$ which sets $\bar{P}_e[n,k]$ to P_{tar} for the known $H'[n,k]$, thus satisfying the final average BER requirement. Intuitively, this loading algorithm tends to underload each subchannel to account for the statistical properties of the errors in CSI. The system performance is then measured by the average spectral efficiency, which is

$$R_{imp} = E_{H'[n,1],\cdots,H'[n,K]} \left\{ \frac{1}{K} \sum_{k=1}^{K} \beta[n,k] \right\}. \tag{15}$$

If the $H'[n,k]$'s have the same distribution, this reduces to

$$R_{imp} = E_{H'[n,k]} \left\{ \beta[n,k] \right\}. \tag{16}$$

In this case, the performance does not depend on the system parameters or the power delay profile of the WSSUS channel when the guard interval is ignored, as in Section III-B.

Consider the special case when $H[n,k]$ given $H'[n,k]$ is complex Gaussian with mean s and variance σ^2. It follows that $r = |H[n,k]|$ conditioned on $H'[n,k]$ is Ricean distributed. Using (7) and the conditional pdf of $|H[n,k]|$ to calculate the expectation in (14), the average BER becomes

$$\begin{aligned} &\bar{P}_e[n,k] \\ \approx\ & \int_0^\infty c_1 \exp\left\{ -\frac{c_2 \frac{E_s}{N_0} r^2}{2^{\beta[n,k]}-1} \right\} \\ &\cdot \frac{2r}{\sigma^2} \exp\left\{ -\frac{r^2+|s|^2}{\sigma^2} \right\} I_0\left(\frac{2r|s|}{\sigma^2} \right) dr \qquad (17) \\ =\ & c_1 \frac{2^{\beta[n,k]}-1}{a+(2^{\beta[n,k]}-1)} \exp\left\{ -\frac{b}{a+(2^{\beta[n,k]}-1)} \right\}, \qquad (18) \end{aligned}$$

where

$$a = c_2 \sigma^2 \frac{E_s}{N_0} \tag{19}$$

and

$$b = c_2 |s|^2 \frac{E_s}{N_0}. \tag{20}$$

Here $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind. By taking the derivative of (18) with respect to $\beta[n,k]$, we have

$$\begin{aligned} &\frac{d\,\bar{P}_e[n,k]}{d\,\beta[n,k]} \\ =\ & \frac{c_1 \ln 2 \cdot 2^{\beta[n,k]}}{\left[a+(2^{\beta[n,k]}-1)\right]^2} \left[a + \frac{b(2^{\beta[n,k]}-1)}{a+(2^{\beta[n,k]}-1)} \right] \\ &\cdot \exp\left\{ -\frac{b}{a+(2^{\beta[n,k]}-1)} \right\}, \qquad (21) \\ >\ & 0 \qquad \text{for} \quad \beta[n,k] > 0. \qquad (22) \end{aligned}$$

Therefore, $\bar{P}_e[n,k]$ is a monotonically increasing function of $\beta[n,k]$, with $\bar{P}_e[n,k] = 0$ for $\beta[n,k] = 0$. This means that there is a unique solution for $\beta[n,k]$ when we use the proposed loading algorithm. Although it is not easy to find a closed-form solution for $\beta[n,k]$, a numerical approach can easily be used to find the solution. For example, in our simulations, we use the function *fzero* provided by Matlab to solve this non-linear equation. The two different sources of imperfect CSI, noisy estimation and delay, will be considered separately in the following subsections.

B. Noisy Channel Estimation

Channel estimation in OFDM systems has been studied in many papers (for example, see [30]-[34]). The channel estimation error is often measured by the mean square error (MSE), which can be defined as

$$\text{MSE} = E\left\{ \frac{1}{K} \sum_{k=1}^{K} |H'[n,k] - H[n,k]|^2 \right\}. \tag{23}$$

Generally speaking, the statistics of the channel estimation error is very complicated, and it highly depends on the estimation approach used and the system details. To avoid such complications, we choose to characterize the estimation error as additive Gaussian noise. The estimated channel response is $H'[n,k] = H[n,k] + e[n,k]$, where the estimation error $e[n,k]$, independent from $H[n,k]$, is a complex Gaussian random variable with zero mean and a variance, σ_e^2, equal to the MSE in (23). This simplified model, which has also been used in the literature (for example, [35][36]), enables us to roughly evaluate the impact of estimation errors.

Using the Gaussian error assumption, it can be shown that $H[n,k]$ given $H'[n,k]$ is complex Gaussian with mean

$$s_1 = \frac{1}{1+\sigma_e^2} H'[n,k] \tag{24}$$

and variance

$$\sigma_1^2 = \frac{\sigma_e^2}{1+\sigma_e^2}. \tag{25}$$

Replacing s and σ^2 in (19) and (20) with s_1 and σ_1^2, respectively, yields the average BER $\bar{P}_e[n,k]$ for the case under consideration. Note that for ideal CSI ($\sigma_e = 0$), this $\bar{P}_e[n,k]$ reduces to (7), as expected.

Using the proposed loading algorithm, the average spectral efficiency obtained from Monte Carlo simulations is shown in Fig. 1 for different levels of error in the channel estimate with $P_{tar} = 10^{-3}$. Since all the $H'[n,k]$'s have the same distribution, these results do not depend on the system parameters or the power delay profile with the given assumptions according to (16). Performance loss is moderate when the MSE in the channel estimate is less than -15 dB. We conclude that noisy channel estimation will not be a significant problem in the application of adaptive OFDM, as long as we have a reasonably good channel estimator to achieve an MSE of -15 dB [31] [33]. Note that $H[n,k]$ is normalized as explained in Section II. Therefore, the MSE here indicates how channel estimation error compares to the true channel energy.

C. Delay in CSI

Now, consider the component of the CSI error caused by the channel information delay, while assuming perfect channel estimation. In this case, the channel estimate is given by $H'[n,k] = H[n-\Delta n,k]$, where $\tau_D \triangleq \Delta n \cdot T$ is the delay time between the channel estimation and the actual transmission. According to (5) and (3), the correlation coefficient between $H[n,k]$ and $H'[n,k]$ is $\rho = r_t(\tau_D) = J_0(2\pi f_D \tau_D)$. In

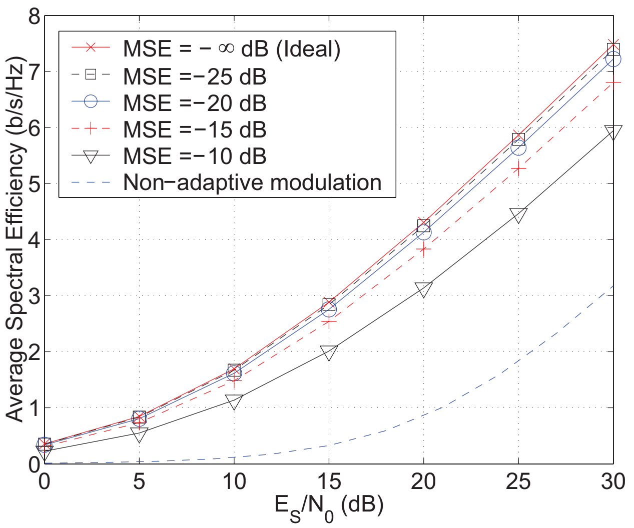


Fig. 1. Average spectral efficiency for adaptive OFDM with Gaussian noise in CSI for $P_{tar} = 10^{-3}$.

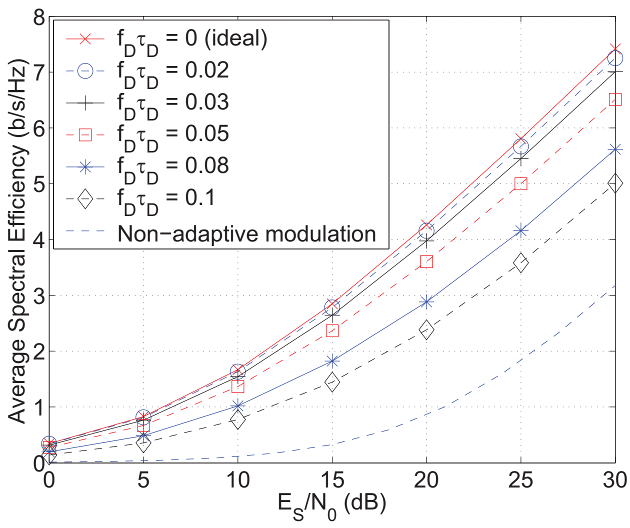


Fig. 3. Average spectral efficiency for adaptive OFDM with CSI delay for $P_{tar} = 10^{-3}$.

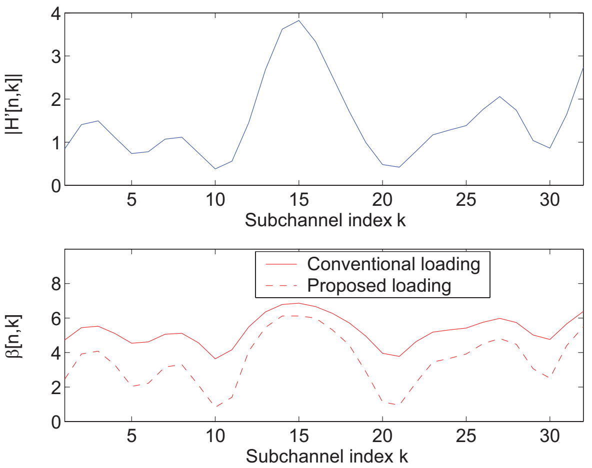


Fig. 2. Snapshots of the estimated channel and the results of the bit allocation algorithm.

practical systems, if we can estimate the maximum Doppler frequency [37], then this correlation coefficient can be calculated. Considering the fact that both $H[n,k]$ and $H'[n,k]$ are complex Gaussian with zero mean and unit variance, $H[n,k]$ given $H'[n,k]$ is complex Gaussian with mean

$$s_2 = \rho H'[n,k] \tag{26}$$

and variance

$$\sigma_2^2 = 1 - \rho^2. \tag{27}$$

The average BER $\bar{P}_e[n,k]$ is given by (18) with $s = s_2$ and $\sigma^2 = \sigma_2^2$. This $\bar{P}_e[n,k]$ reduces to (7) when $f_D\tau_D = 0$ (ideal CSI). Note that a similar idea has been described in [38] for the single-carrier case.

Snapshots of the estimated channel frequency response and the results of the loading algorithm are shown in Fig. 2 for the case $f_D\tau_D = 0.1$ and $E_s/N_0 = 20$ dB. The conventional loading results using (10) are also shown for reference. It turns out that the new loading algorithm always allocates less bits for each subchannel than the conventional one, called "underloading" in [13]. However, here underloading is used to combat the effects of CSI errors on adaptive loading at the transmitter rather than the effects of those errors on coherent detection at the receiver as in [13]. It is interesting to note that the most severe underloading occurs for subchannels in deep fades. To see why this is so, note that the variance of the conditional pdf of $H[n,k]$ does not depend on $H'[n,k]$, while the mean does. Thus the uncertainty in $H[n,k]$ relative to the size of the actual value of $H[n,k]$ increases as the mean becomes smaller, which makes a more severe underloading reasonable. This type of underloading is different from that in [13], where the underloading was more severe for subchannels with higher gains.

In Fig. 3, the average spectral efficiency is shown for several different values of the Doppler-delay product, $f_D\tau_D$, with $P_{tar} = 10^{-3}$. As explained in (16), these results do not depend on the system parameters or the power delay profile under the given assumptions. The performance degradation is very small for small values of the Doppler-delay product, specifically for $f_D\tau_D \leq 0.03$. For indoor environments, as in WLANs with 802.11a parameters, if a user is moving at the speed of 10 miles per hour, even when the delay is quite large, for example, $\tau_D = 100T$, we have $f_D\tau_D \approx 0.024$. Therefore, the performance loss will be very small. However, for outdoor environments, due to the higher Doppler rates and the longer block length, $f_D\tau_D$ could be larger than 0.1, which will cause significant degradation. For example, a delay of 3 OFDM blocks in the outdoor system with a user velocity of 60 miles per hour results in an $f_D\tau_D$ of 0.12, and around 50% reduction in average spectral efficiency when $E_s/N_0 = 20$ dB.

V. Multiple Estimates

We have been using a single estimate for the real channel $H[n,k]$. However, in some scenarios, multiple estimates can be available. These multiple estimates can come from either different frequencies or different times or both. So the natural questions are: what is an efficient way to employ multiple estimates? If we use them, can we obtain more accurate information of $H[n,k]$ and improve the performance?

To answer these questions, we begin with the distribution of a random variable given multiple estimates. Consider a random variable H and a $M \times 1$ random vector $\mathbf{H}_e$. Assume that both H and $\mathbf{H}_e$ are Gaussian distributed with zero mean, and the covariance matrix is given by

$$E\left\{\left(\begin{array}{c} H \\ \mathbf{H}_e \end{array}\right)\left(H^* \ \mathbf{H}_e^\dagger\right)\right\} = \left[\begin{array}{cc} 1 & \mathbf{a}^\dagger \\ \mathbf{a} & \mathbf{B} \end{array}\right]. \quad (28)$$

Then it can be shown that H given $\mathbf{H}_e$ is Gaussian distributed with mean

$$s_H = \mathbf{a}^\dagger \mathbf{B}^{-1} \mathbf{H}_e \quad (29)$$

and variance

$$\sigma_H^2 = 1 - \mathbf{a}^\dagger \mathbf{B}^{-1} \mathbf{a}. \quad (30)$$

Suppose we have multiple estimates and we know the joint distribution of the real channel and the multiple estimates, then the conditional pdf of the real channel can be obtained as just described. With this conditional distribution, (18) holds directly with $s = s_H$ and $\sigma^2 = \sigma_H^2$, and the loading algorithm proposed in Section IV-A can be easily extended to the case with multiple estimates, while still satisfying the average BER constraint. In the following subsections, we will use this loading algorithm to investigate the benefit of multiple estimates.

A. Multiple Estimates From Different Frequencies

So far for $H[n,k]$, only the estimate in the same subchannel $H'[n,k]$ has been used. However, the estimates of the frequency responses in all subchannels, that is, $H'[n,1], \cdots, H'[n,K]$, are available in many scenarios. For example, in the IEEE 802.11a standard, training symbols are sent out before the actual data transmission, which enables channel estimation in all subchannels. So all these estimates, instead of $H'[n,k]$ alone, can be employed to form the distribution of $H[n,k]$. We will consider the case with channel estimation error and the case with CSI delay separately, and see whether the performance can be improved by using all of the available information.

1) Noisy Channel Estimation: As in Section IV-B, we assume $H'[n,k] = H[n,k] + e[n,k]$, where $e[n,k]$ is complex Gaussian with zero mean and variance σ_e^2. Moreover, the $e[n,k]$'s are assumed to be independent from each other. We first focus on the case when $H'[n,k]$ and $H'[n,i]$ $(i \neq k)$ are employed to form the distribution of $H[n,k]$, where

$$H'[n,k] = H[n,k] + e[n,k], \quad (31)$$

$$H'[n,i] = H[n,i] + e[n,i]. \quad (32)$$

Here $e[n,k]$ and $e[n,i]$ are independent. Therefore, the covariances between $H[n,k]$, $H'[n,k]$ and $H'[n,i]$ are given by

$$E\{H[n,k]H'[n,k]^*\} = E\{H[n,k]H[n,k]^*\} = 1, \quad (33)$$

$$E\{H[n,k]H'[n,i]^*\} = E\{H[n,k]H[n,i]^*\} = \sum_{\ell=1}^{L} q_\ell^2 W_K^{(k-i)p_\ell} \triangleq \eta, \quad (34)$$

$$E\{H'[n,k]H'[n,i]^*\} = E\{H[n,k]H[n,i]^*\} = \eta. \quad (35)$$

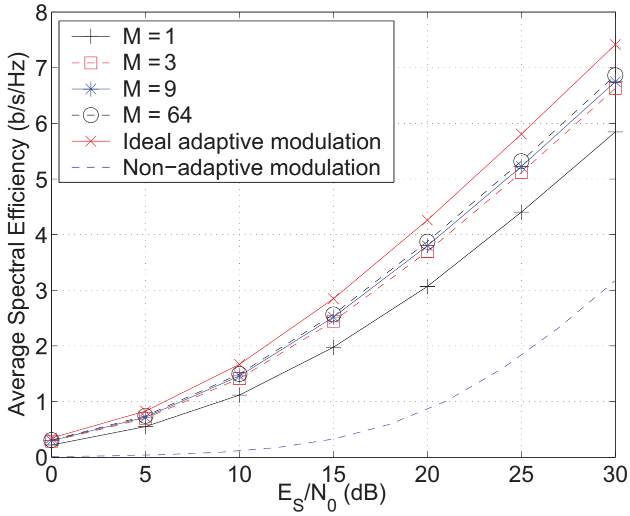


Fig. 4. Average spectral efficiency for the indoor adaptive OFDM system with CSI estimation error ($\sigma_e^2 = -10$ dB, M estimates) for $P_{tar} = 10^{-3}$.

Moreover, the variance of $H'[n,k]$ and $H'[n,i]$ is $1 + \sigma_e^2$. Define $\mathbf{v}_1 = [H[n,k], H'[n,k], H'[n,i]]^T$, then

$$E\{\mathbf{v}_1 \mathbf{v}_1^\dagger\} = \left[\begin{array}{ccc} 1 & 1 & \eta \\ 1 & 1+\sigma_e^2 & \eta \\ \eta^* & \eta^* & 1+\sigma_e^2 \end{array}\right]. \quad (36)$$

According to (29) and (30), the pdf of $H[n,k]$ given $H'[n,k]$ and $H'[n,i]$ is a Gaussian distribution with mean

$$s_3 = \frac{(1+\sigma_e^2 - |\eta|^2)H'[n,k] + \eta\sigma_e^2 H'[n,i]}{(1+\sigma_e^2)^2 - |\eta|^2} \quad (37)$$

and variance

$$\sigma_3^2 = \frac{\sigma_e^2(1+\sigma_e^2 - |\eta|^2)}{(1+\sigma_e^2)^2 - |\eta|^2}. \quad (38)$$

The uncertainty of $H[n,k]$ directly affects the performance, and it is determined by the variance σ_3^2. Consider two extreme cases: $\eta = 0$ and $|\eta| = 1$. For $\eta = 0$, as expected, (37) and (38) reduce to (24) and (25), respectively. Because $H[n,k]$ and $H'[n,i]$ are independent, the knowledge of $H'[n,i]$ does not provide any additional information about $H[n,k]$. On the other hand, if $|\eta| = 1$, we have $s_3 = \frac{1}{2+\sigma_e^2}(H'[n,k] + H'[n,i])$ and $\sigma_3^2 = \frac{\sigma_e^2}{2+\sigma_e^2}$. In this case σ_3^2 is reduced almost by half when σ_e is small, which means the variance is about 3 dB smaller. For $0 < \eta < 1$, the actual value of η determines the gain achieved by using these two estimates.

The derivation can be easily extended to the cases with more than two estimates. We simulate the performance of both the indoor and outdoor systems introduced in Section II, assuming M estimates are available. Specifically, to obtain the distribution of $H[n,k]$, we use the estimated channel in the M subchannels closest to the kth subchannel, that is, $H'[n, \mathrm{mod}(k - J + i - 1, K) + 1], i = 0, \cdots, M-1$, where $J = \lfloor (M-1)/2 \rfloor$. For example, for $M = 3$, $H'[n,k-1], H'[n,k], H'[n,k+1]$ are used to estimate $H[n,k]$. In Fig. 4, the average spectral efficiency for the indoor system is shown for different values of M with $\sigma_e^2 = -10$ dB and $P_{tar} = 10^{-3}$. On the one hand, using the information in all subchannels ($M = 64$) greatly enhances the performance compared to

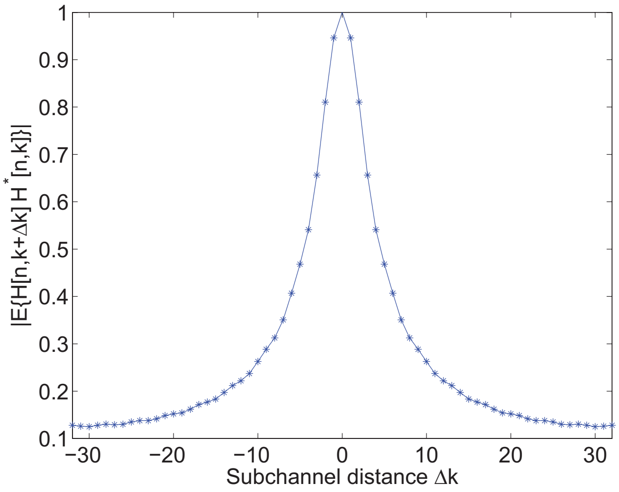


Fig. 5. Magnitude of the covariance between two subchannels with distance Δk for the indoor system.

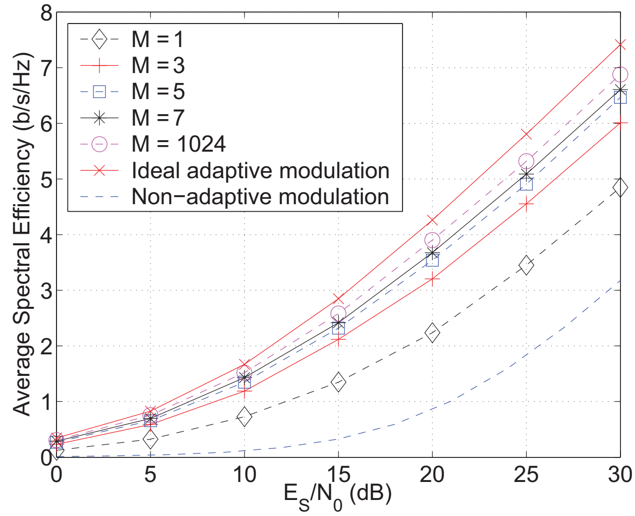


Fig. 6. Average spectral efficiency for the outdoor adaptive OFDM system with CSI estimation error ($\sigma_e^2 = -6$ dB, M estimates) for $P_{tar} = 10^{-3}$.

$M = 1$. In particular, $\sigma_e^2 = -10$ dB and $M = 64$ gives similar performance as $\sigma_e^2 = -15$ dB and $M = 1$. Therefore, by using these estimates, we create an extra 5 dB margin in the MSE of the estimation error, thus making the channel estimation much easier. On the other hand, by using only a few subchannels ($M = 3$), we can achieve a performance very close to $M = 64$. So, in practice, it is enough to use only a few adjacent subchannels to obtain the majority of the gain, while reducing the complexity. This makes perfect sense if we look at the covariance between different subchannels. As shown in Fig. 5, the magnitude of the covariance decreases rapidly as the distance between the two subchannels increases. Similar performance curves for the outdoor system are given in Fig. 6. Here we see that $\sigma_e^2 = -6$ dB and $M = 1024$ provides similar performance as $\sigma_e^2 = -15$ dB and $M = 1$, which gives 9 dB difference in MSE. Also, $M = 7$ performs quite closely to $M = 1024$. We note that the achieved performance gain is directly related to the covariance between different subchannels, which is mainly decided by the delay spread and the subchannel spacing.

2) Delay in CSI: For the case with CSI delay, we have a completely different conclusion. In particular, we have the following theorem, which states that knowing the information in other subchannels does not provide additional knowledge about the desired channel information, thus giving no performance enhancement.

Theorem 1: Assuming the channel model introduced in Section II, we have

$$\begin{aligned} & I(H[n,k]; H[n-\Delta n,1], \cdots, H[n-\Delta n,K]) \\ = \; & I(H[n,k]; H[n-\Delta n,k]). \end{aligned} \tag{39}$$

Here $I(\cdot\,;\cdot)$ denotes the mutual information.

So even when $H[n-\Delta n,1], \cdots, H[n-\Delta n,K]$ are known, $H[n-\Delta n,k]$ alone contains all the information about $H[n,k]$. And there is no need to use the information other than $H[n-\Delta n,k]$.

Proof: For illustration purpose, we prove the theorem for $k = 1$. The case for any other k can be proved in the same way.

Firstly, according to Property 2, for any $i \neq k$, we have

$$\begin{aligned} E\{H[n,k]H^*[n-\Delta n,k]\} &= r_t(\Delta n \cdot T) \\ &= \rho, \end{aligned} \tag{40}$$

$$\begin{aligned} E\{H[n-\Delta n,k]H^*[n-\Delta n,i]\} &= \sum_{\ell=1}^{L} q_\ell^2 W_K^{(k-i)p_\ell} \\ &= \eta, \end{aligned} \tag{41}$$

$$E\{H[n,k]H^*[n-\Delta n,i]\} = \rho\eta. \tag{42}$$

Define $\mathbf{v}_2 = [H[n-\Delta n,1], \cdots, H[n-\Delta n,K]]^T$, with the covariance matrix denoted by

$$\boldsymbol{\Sigma}_2 = E\{\mathbf{v}_2\mathbf{v}_2^\dagger\} = \begin{bmatrix} 1 & \mathbf{c}^\dagger \\ \mathbf{c} & \mathbf{D} \end{bmatrix}. \tag{43}$$

Then for $\mathbf{v}_3 = [H[n,1], H[n-\Delta n,1], \cdots, H[n-\Delta n,K]]^T$, the covariance matrix is given by

$$\boldsymbol{\Sigma}_3 = E\{\mathbf{v}_3\mathbf{v}_3^\dagger\} = \begin{bmatrix} 1 & \rho & \rho\mathbf{c}^\dagger \\ \rho^* & 1 & \mathbf{c}^\dagger \\ \rho^*\mathbf{c} & \mathbf{c} & \mathbf{D} \end{bmatrix} \triangleq \begin{bmatrix} 1 & \mathbf{b}^\dagger \\ \mathbf{b} & \boldsymbol{\Sigma}_2 \end{bmatrix}. \tag{44}$$

The mutual information between complex Gaussian random vectors $\mathbf{x}$ and $\mathbf{y}$ is known [39] as

$$I(\mathbf{x};\mathbf{y}) = \log_2 \frac{\det \boldsymbol{\Sigma}_\mathbf{x} \cdot \det \boldsymbol{\Sigma}_\mathbf{y}}{\det \boldsymbol{\Sigma}_\mathbf{xy}}, \tag{45}$$

where $\boldsymbol{\Sigma}_\mathbf{x}$, $\boldsymbol{\Sigma}_\mathbf{y}$ and $\boldsymbol{\Sigma}_\mathbf{xy}$ are the covariance matrices of $\mathbf{x}$, $\mathbf{y}$ and the stacked vector of $\mathbf{x}$ and $\mathbf{y}$, respectively. Therefore, it follows that

$$I(H[n,1]; H[n-\Delta n,1]) = \log_2 \frac{1}{\det \begin{bmatrix} 1 & \rho \\ \rho^* & 1 \end{bmatrix}} \tag{46}$$

$$= \log_2 \frac{1}{1-|\rho|^2}, \tag{47}$$

and

$$\begin{aligned} & I(H[n,1]; H[n-\Delta n,1], \cdots, H[n-\Delta n,K]) \\ & = \log_2 \frac{\det \boldsymbol{\Sigma}_2}{\det \boldsymbol{\Sigma}_3}. \end{aligned} \tag{48}$$

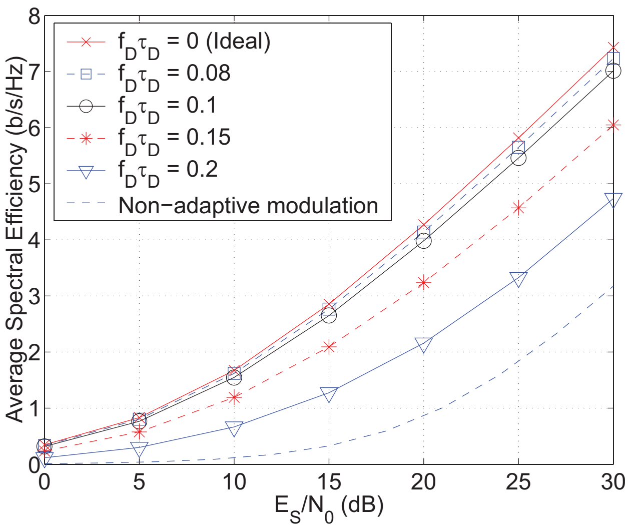


Fig. 7. Average spectral efficiency for adaptive OFDM with CSI delay with two outdated estimates for $P_{tar} = 10^{-3}$.

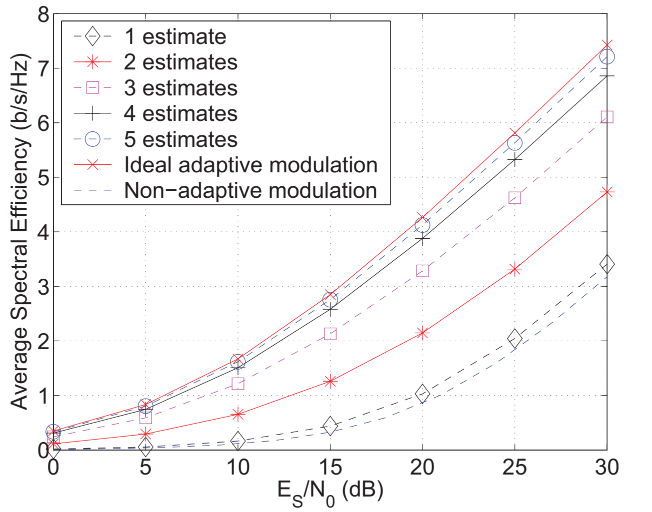


Fig. 8. Average spectral efficiency for adaptive OFDM with CSI delay with multiple outdated estimates for $P_{tar} = 10^{-3}$ and $f_D\tau_D = 0.2$.

Using the identities on the inverse and determinant of a partitioned matrix [40], we have (49) and (50), shown at the bottom of the page. Then (48) becomes

$$\begin{aligned} & I(H[n,1]; H[n-\Delta n,1], \cdots, H[n-\Delta n,K]) \\ = \; & \log_2 \frac{\det \mathbf{\Sigma}_2}{\det \mathbf{\Sigma}_2 (1 - \mathbf{b}^\dagger \mathbf{\Sigma}_2^{-1}\mathbf{b})} && (51) \\ = \; & \log_2 \frac{1}{1-|\rho|^2} && (52) \\ = \; & I(H[n,1]; H[n-\Delta n,1]). && (53) \end{aligned}$$

This completes the proof. ■

Note that Theorem 1 is derived based on the key property that the time-domain and frequency-domain correlation are decoupled and separated, as stated in Property 2.

B. Multiple Estimates From Different Times

Multiple estimates from different times can also be available in some scenarios. For example, if the receiver gets multiple OFDM training symbols before the actual data, or the receiver has multiple packets from the same transmitter in a short time period, the receiver can have multiple channel estimates at different times. In this subsection, we will focus on the case with CSI delay.

Previously, we have used the outdated information $H[n-\Delta n, k]$ as the estimate. If we assume we store previous channel estimates, produced prior to $H[n-\Delta n, k]$, these estimates could potentially be used in combination to reduce the uncertainty in $H[n,k]$. For example, assume we have one more previous channel estimate $H''[n,k] = H[n-2\Delta n, k]$. Now $\beta[n,k]$ is determined based on the values of both $H'[n,k]$ and $H''[n,k]$. Define $\rho_1 = J_0(2\pi f_D \tau_D)$ and $\rho_2 = J_0(2\pi f_D 2\tau_D)$. It can be shown that the conditional pdf of $H[n,k]$ given $H'[n,k]$ and $H''[n,k]$ is complex Gaussian with mean

$$s_4 = \frac{\rho_1(1-\rho_2)H'[n,k] - (\rho_1^2 - \rho_2)H''[n,k]}{1-\rho_1^2} \tag{54}$$

and variance

$$\sigma_4^2 = \frac{(1-\rho_2)(1+\rho_2-2\rho_1^2)}{1-\rho_1^2}. \tag{55}$$

The average spectral efficiency obtained by using the proposed loading algorithm based on (18) with $s = s_4$ and $\sigma^2 = \sigma_4^2$ is shown in Fig. 7. The performance loss is very small when $f_D\tau_D = 0.1$, which is a great improvement over the case when only a single estimate, $H[n-\Delta n, k]$, is known (Fig. 3).

Results for more than two outdated estimates can be derived in a similar way. Assuming M estimates, $H[n-\Delta n, k], H[n-2\Delta n, k], \cdots, H[n-M\Delta n, k]$, are available, then the correlation coefficient is given by

$$E\{H[n-i\Delta n, k]H^*[n-j\Delta n, k]\} = J_0(2\pi f_D (j-i)\tau_D). \tag{56}$$

Given this, we can derive the mean and the variance of $H[n,k]$ conditioned on these M estimates according to (29) and (30). Then, using the proposed loading algorithm, the average spectral efficiency can be calculated. The results are shown in Fig. 8 for $f_D\tau_D = 0.2$ when 1, 2, 3, 4, and 5 estimates are available. It is obvious that multiple estimates can offer a significant advantage. For $f_D\tau_D = 0.2$, the adaptive system performs similarly to the non-adaptive system when only a single estimate is available, but performs quite close to the ideal adaptive system when 4 or 5 estimates are available.

$$\begin{aligned} \mathbf{b}^\dagger \mathbf{\Sigma}_2^{-1}\mathbf{b} &= \rho \, [1 \;\; \mathbf{c}^\dagger] \begin{bmatrix} (1-\mathbf{c}^\dagger\mathbf{D}^{-1}\mathbf{c})^{-1} & -(1-\mathbf{c}^\dagger\mathbf{D}^{-1}\mathbf{c})^{-1}\mathbf{c}^\dagger\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{c}(1-\mathbf{c}^\dagger\mathbf{D}^{-1}\mathbf{c})^{-1} & \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{c}(1-\mathbf{c}^\dagger\mathbf{D}^{-1}\mathbf{c})^{-1}\mathbf{c}^\dagger\mathbf{D}^{-1} \end{bmatrix} \rho^* \begin{bmatrix} 1 \\ \mathbf{c} \end{bmatrix} && (49) \\ &= |\rho|^2 \, [1 \;\; 0] \begin{bmatrix} 1 \\ \mathbf{c} \end{bmatrix} = |\rho|^2, && (50) \end{aligned}$$

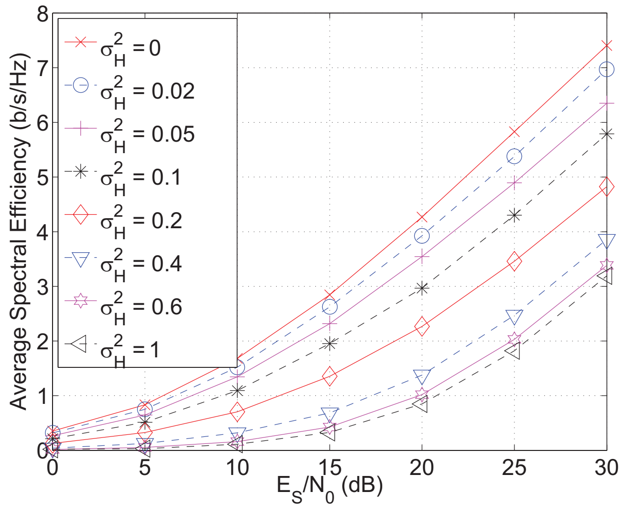


Fig. 9. Average spectral efficiency for different σ_H^2 values for $P_{tar} = 10^{-3}$.

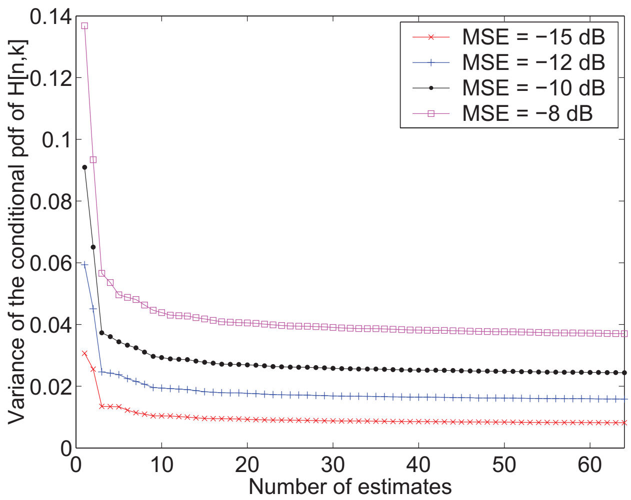


Fig. 10. Variance of the conditional pdf of $H[n,k]$ given different number of estimates for the indoor system with Gaussian noise in CSI.

The use of multiple outdated estimates also enables the system to tolerate longer delay (that is, longer delay time or higher Doppler). When only one estimate is available, $f_D\tau_D$ up to 0.03 gives performance close to the ideal system. When 2 estimates are available, we obtain very good performance for $f_D\tau_D$ up to 0.1, which means the system can tolerate 3 to 4 times longer delay after adding one additional piece of outdated information. Further, when 5 estimates are available, $f_D\tau_D$ up to 0.2 gives very good performance. Similarly, as in (16), we can see that the results in Fig. 7 and Fig. 8 do not depend on the system parameters or the power delay profile under the given assumptions.

It is worth noting that, although the idea of using multiple outdated estimates might seem similar to channel prediction [14], here, we are using the extra data to better characterize the statistical properties of the error in the estimates instead of trying to produce a single predicted value for the channel gain.

C. The Role of σ_H^2

Throughout the discussion of multiple estimates, the variance of the conditional pdf of $H[n,k]$, which determines the uncertainty of the real channel based on the given estimates, plays a very important role. In fact, the average spectral efficiency is uniquely determined by this variance. To show this, we consider the most general case when a vector $\mathbf{H}_e$ composed of multiple estimates is used to form the distribution of $H[n,k]$, with the covariance matrix following (28). Then the mean s_H and the variance σ_H^2 of the conditional pdf of $H[n,k]$ is given by (29) and (30). We know that after inverting (18) with $s = s_H$ and $\sigma^2 = \sigma_H^2$, $\beta[n,k]$ can be represented as a function of s_H and σ_H^2, namely $g(s_H, \sigma_H^2)$. Given the covariance matrix in (28), s_H is a function of $\mathbf{H}_e$ according to (29), while σ_H^2 is a deterministic value. Since $\mathbf{H}_e$ is a random vector, s_H can also be considered as a random variable. Therefore, the average spectral efficiency is

$$\begin{aligned} R &= E_{\mathbf{H}_e}\{\beta[n,k]\} = E_{\mathbf{H}_e}\{g(s_H, \sigma_H^2)\} && (57)\\ &= E_{s_H}\{g(s_H, \sigma_H^2)\}. && (58)\end{aligned}$$

The third equality follows because the variance σ_H^2 does not depend on $\mathbf{H}_e$, and the function $g(\cdot)$ depends on $\mathbf{H}_e$ only through s_H. Further, based on (29), we have $E\{s_H\} = 0$ because $E\{\mathbf{H}_e\} = 0$, and the variance of s_H is

$$\begin{aligned} \sigma_{s_H}^2 &= \mathbf{a}^\dagger \mathbf{B}^{-1} E\{\mathbf{H}_e \mathbf{H}_e^\dagger\} \mathbf{B}^{-1}\mathbf{a} && (59)\\ &= \mathbf{a}^\dagger \mathbf{B}^{-1}\mathbf{a} && (60)\\ &= 1 - \sigma_H^2. && (61)\end{aligned}$$

Therefore, for any fixed σ_H^2, the distribution of s_H is uniquely determined, as is the average spectral efficiency according to (58). This means that, no matter how many estimates are available and whether these estimates come from different frequencies or times, the system performance can be fully characterized by the variance σ_H^2.

Moreover, (58) provides another way to calculate the average spectral efficiency given the variance σ_H^2. To be more specific, for any fixed σ_H^2, we can first generate s_H according to the complex Gaussian distribution with zero mean and variance $1-\sigma_H^2$, then evaluate the expectation in (58) numerically. The average spectral efficiency calculated in this way is plotted for different σ_H^2 values in Fig. 9. We note that $\sigma_H^2 = 0$ and $\sigma_H^2 = 1$ correspond to the ideal adaptive system and non-adaptive system, respectively. It is seen that the adaptive system has very small performance loss compared to the ideal system if $\sigma_H^2 \le 0.02$. The degradation is moderate as long as $\sigma_H^2 \le 0.05$. On the other hand, if $\sigma_H^2 \ge 0.4$, the advantage over the non-adaptive system diminishes.

Since the variance σ_H^2 uniquely determines the average spectral efficiency, knowing σ_H^2 is enough to determine the performance according to Fig. 9, without doing the actual simulations. As one example, the variance σ_H^2 is plotted in Fig. 10 for the indoor system with estimation error in CSI when multiple estimates from different subchannels are available. Another example is given in Fig. 11 for the case with CSI delay when multiple estimates from different times are available. Combining these figures with Fig. 9, we can easily figure out the system performance. For the second case, it is seen that, when only 1 estimate is available, $f_D\tau_D$ needs

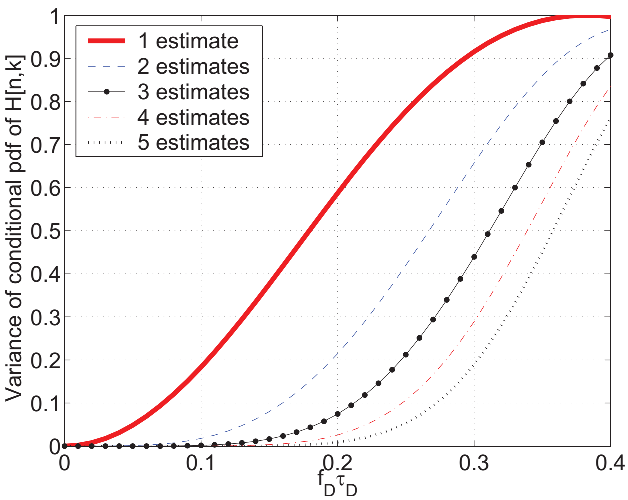


Fig. 11. Variance of the conditional pdf of $H[n,k]$ given different number of outdated estimates.

to be smaller than 0.03 to achieve negligible loss. But when 5 estimates are available, $f_D\tau_D$ can be as large as 0.2 to give similar performance. All these agree with the previous observations.

Note that the analysis in this section can also be extended to the case when multiple estimates from both different frequencies and different times are available. Since it is quite straightforward, no further details will be given here.

VI. Conclusions

In this paper, the large potential gain of adaptive OFDM was investigated when ideal channel information is known. The performance degradation due to CSI errors, caused by noisy channel estimation and delay, was studied using the proposed loading algorithm. Simulation results show that the degradation is moderate for the level of error which can be achieved by a good channel estimator, but is severe for CSI delay when $f_D\tau_D \geq 0.1$, which can occur in outdoor environments.

To improve the performance, we proposed to use multiple estimates, and the loading algorithm is extended to the case when multiple estimates are available. These multiple estimates can come from either different frequencies or times or both. Simulation results showed that the average spectral efficiency can be improved significantly by using multiple estimates, which also means that the system can tolerate larger channel estimation error or longer delay to achieve reasonable performance.

References

[1] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Norwood, MA: Artech House, 2000.

[2] ETSI, *Digital Audio Broadcasting (DAB)*, 2nd ed., May 1997, ETS 300 401.

[3] ——, *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television*, Aug. 1997, EN 300 744 V1.1.2.

[4] ANSI, *ANSI/T1E1.4/94-007, Asymmetric Digital Subscriber Line (ADSL) Metallic Interface*, Aug. 1997.

[5] IEEE 802.11, "Draft supplement to standard for telecommunications and information exchange between systems - LAN/MAN specific requirement — part 11: wireless MAC and PHY specifications: High speed physical layer in the 5 GHz band," *P802.11a/D6.0*, May 1999.

[6] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119–124, Feb. 1989.

[7] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *Proc. IEEE*, vol. 88, pp. 611–640, May 2000.

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

[9] S. T. Chung and A. J. Goldsmith, "Adaptive multicarrier modulation for wireless systems," in *Proc. Asilomar Conf. on Sig., Sys. and Comp.*, Oct. 2000, pp. 1603–1607.

[10] P. Chow, J. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 772–775, Feb.-Mar.-Apr. 1995.

[11] L. Hanzo, W. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation: Principles and Applications for Personal Communications, WLANs and Broadcasting*, 2nd ed. New York: John Wiley & Sons, 2000.

[12] D. Dardari, "Ordered subcarrier selection algorithm for OFDM-based high-speed WLANs," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1452–1458, Sep. 2004.

[13] A. Leke and J. M. Cioffi, "Multicarrier systems with imperfect channel knowledge," in *Proc. PIMRC*, Sep. 1998, pp. 549–553.

[14] M. R. Souryal and R. L. Pickholtz, "Adaptive modulation with imperfect channel information in OFDM," in *Proc. ICC*, June 2001, pp. 1861–1865.

[15] Q. Su and S. Schwartz, "Effects of imperfect channel information on adaptive loading gain of OFDM," in *Proc. IEEE Veh. Technol. Conf.-Fall*, Oct. 2001, vol. 1, pp. 475–478.

[16] Q. Su, L. J. Cimini, Jr., and R. S. Blum, "On the problem of channel mismatch in constant-bit-rate adaptive modulation for OFDM," in *Proc. IEEE Veh. Technol. Conf.*, May 2002, vol. 2, pp. 585–589.

[17] S. Ye, R. S. Blum, and L. J. Cimini, Jr., "Adaptive modulation for variable-rate OFDM systems with imperfect channel information," in *Proc. IEEE Veh. Technol. Conf.*, May 2002, vol. 2, pp. 767–771.

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

[19] Y. Yao and G. B. Giannakis, "Rate-maximizing power allocation in OFDM based on partial channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1073–1083, May 2005.

[20] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO OFDM based on partial channel state information," *IEEE Trans. Signal Processing*, vol. 52, pp. 202–213, Jan. 2004.

[21] G. Barriac and U. Madhow, "Space-time communication for OFDM with implicit channel feedback," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3111–3129, Dec. 2004.

[22] Z. Hu, G. Zhu, Y. Xia, and G. Liu, "Multiuser subcarrier and bit allocation for MIMO-OFDM systems with perfect and partial channel information," in *Proc. WCNC*, Mar. 2004, vol. 2, pp. 1188–1193.

[23] Y.-H. Pan, K. B. Letaief, and Z. Cao, "Space-time coded adaptive transmit antenna arrays for OFDM wireless systems utilizing channel side information," in *Proc. IEEE Veh. Technol. Conf.*, Apr. 2003, vol. 2, pp. 1127–1131.

[24] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: a unified view," *IEEE Trans. Commun.*, vol. 49, pp. 1561–1571, Sep. 2001.

[25] R. Steele, *Mobile Radio Communications*. New York: IEEE Press, 1992.

[26] P. A. Bello, "Characterization of randomly time-variant linear channels," *IEEE Trans. Commun. Syst.*, vol. 11, pp. 360–393, Dec. 1963.

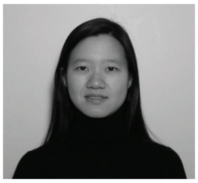
[27] D. Cox, "Delay doppler characteristics of multipath propagation at 910 mhz in a suburban mobile radio environment," *IEEE Trans. Antennas Propagat.*, vol. 20, pp. 625–635, Sep. 1972.

[28] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: IEEE Press, 1994.

[29] B. McNair and L. J. Cimini, Jr., "OFDM for high data rate, high-mobility, wide-area wireless communications," in *Proc. Sarnoff Symposium*, Mar. 2001.

[30] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, July 1998.

[31] Y. Li, L. J. Cimini, Jr., and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 902–915, July 1998.
[32] O. Simeone, Y. Bar-Ness, and U. Spagnolini, "Pilot-based channel estimation for OFDM systems by tracking the delay-subspace," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 315–325, Jan. 2004.
[33] K.-Y. Han, S.-W. Lee, J.-S. Lim, and K.-M. Sung, "Channel estimation for OFDM with fast fading channels by modified Kalman filter," *IEEE Trans. Consumer Electron.*, vol. 50, pp. 443–449, May 2004.
[34] C. R. N. Athaudage and A. D. S. Jayalath, "Enhanced MMSE channel estimation using timing error statistics for wireless OFDM systems," *IEEE Trans. Broadcast.*, vol. 50, pp. 369–376, Dec. 2004.
[35] R. Narasimhan, "Performance of diversity schemes for OFDM systems with frequency offset, phase noise, and channel estimation errors," *IEEE Trans. Commun.*, vol. 50, pp. 1561–1565, Oct. 2002.
[36] H. Cheon and D. Hong, "Effect of channel estimation error in OFDM-based WLAN," *IEEE Commun. Lett.*, vol. 6, pp. 190–192, May 2002.
[37] J. M. Holtzman and A. Sampath, "Adaptive averaging methodology for handoffs in cellular systems," *IEEE Trans. Veh. Technol.*, vol. 44, pp. 59–66, Feb. 1995.
[38] D. L. Goeckel, "Adaptive coding for fading channels using outdated channel estimates," *IEEE Trans. Commun.*, vol. 47, pp. 844–855, June 1999.
[39] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.
[40] H. Lütkepohl, *Handbook of Matrices*. New York: John Wiley & Sons, 1996.



Sigen Ye (S'00-M'06) received her B.E. and M.S. degrees in Electrical Engineering from Tsinghua University, Beijing, China, in 1996 and 1999, respectively, and received her Ph.D. degree in Electrical Engineering from Lehigh University in 2005.

She is currently a Member of Technical Staff in the System Engineering Department at Lucent Technologies, Whippany, NJ. She is working on the performance analysis, system design and optimization for 3G wireless communication systems.

Rick S. Blum (S'83-M'84-SM'94-F'05) received a B.S. in Electrical Engineering from the Pennsylvania State University in 1984 and his M.S. and Ph.D in Electrical Engineering from the University of Pennsylvania in 1987 and 1991.

From 1984 to 1991 he was a member of technical staff at General Electric Aerospace in Valley Forge, Pennsylvania and he graduated from GE's Advanced Course in Engineering. Since 1991, he has been with the Electrical Engineering and Computer Science Department at Lehigh University in Bethlehem, Pennsylvania where he is currently a Professor and holds a Class of 1961 Professorship. His research interests include signal detection and estimation and related topics in the areas of signal processing and communications. He is currently an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and for IEEE COMMUNICATIONS LETTERS. He is also a member of the Signal Processing for Communications Technical Committee of the IEEE Signal Processing Society.

Dr. Blum is a member of Eta Kappa Nu and Sigma Xi, and holds a patent for a parallel signa and image processor architecture. He was awarded an ONR Young Investigator Award in 1997 and an NSF Research Initiation Award in 1992.

Leonard J. Cimini, Jr. (S'77-M'78-SM'89-F'00) received a Ph.D. in Elec. Eng. from the University of Pennsylvania in 1982, and worked at AT&T, first in Bell Labs and then AT&T Labs, for twenty years, where his research concentrated on wireless communications. Dr. Cimini has been very active within the IEEE and was the founding Editor-in-Chief of the IEEE JSAC: WIRELESS COMMUNICATIONS SERIES. Currently, among other duties, he is the Chair of the Emerging Technologies Committee of ComSoc and a member of the Board of Governors. He was elected a Fellow of the IEEE in 2000, and has been a Professor in the ECE Department at the University of Delaware since 2002.