Channel Estimation and Optimal Power Allocation for a Multiple-Antenna OFDM System

Tai-Lai Tung

GlobespanVirata, Inc., 5171 California Avenue, Suite 200, Irvine, CA 92612, USA
Email: ttung@globespanvirata.com

Kung Yao

Electrical Engineering Department, 68-113 Engineering IV Building, University of California,
Los Angeles, CA 90095-1594, USA
Email: yao@ee.ucla.edu

We propose combining channel estimation and optimal power allocation approaches for a multiple-antenna orthogonal frequency division multiplexing (OFDM) system in high-speed transmission applications. We develop a least-square channel estimation approach, derive the performance bound of the estimator, and investigate the optimal training sequences for initial channel acquisition. Based on the channel estimates, the optimal power allocation solution which maximizes the bandwidth efficiency is derived under power and quality of service (QoS) (symbol error rate) constraints. It is shown that combining channel tracking and adaptive power allocation can dramatically enhance the outage capacity of an OFDM multiple-antenna system when severe fading occurs.

Keywords and phrases: OFDM, multiple-antenna, channel estimation, optimal power allocation, capacity, optimal training sequences, MSE bounds.

1. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) technique has recently gained increasing interest as an effective alternative to conventional single tone modulation methods for high-rate communications systems [1, 2, 3]. The main advantage of OFDM is allowing transmission over a frequency-selective fading channel using a low cost receiver. On the other hand, current research has shown that intelligent use of antenna arrays at both receiver and transmitter side can significantly increase the bandwidth efficiency as well as improving the transmission reliability [4]. More recent work then show that combining OFDM technique with multiple-antenna architecture can provide high-rate high performance transmission [5, 6, 7, 8].

In this paper, we first propose least-square (LS) channel estimation approach for OFDM multiple-antennas systems. Based on the channel estimates, the bandwidth efficiency is then maximized by performing optimal power allocation over multiple antennas and subbands where some practical constraints such as specified quality of service (QoS) requirements in terms of symbol error probability and limited transmit power constraints are considered. It is shown

that combining the optimal power allocation method with a closed-loop channel state information updating mechanism can dramatically increase the system outage capacity in time-varying environment.

The organization of the paper is as follows. In Section 2, we first propose least-squares channel estimator. To analyze the performance of the channel estimator, the mean square error (MSE) bound is derived in Section 3. The optimal training sequences which can achieve minimum MSE channel acquisition are designed in Section 4. The tracking and BER performance in time-varying channels are shown in Section 5. In Sections 6 and 7, we introduce the capacity of the OFDM multiple-antenna system, and show that the capacity of the system can be achieved by performing joint optimal power allocation in spatial and frequency domains. In Section 8, we show the bandwidth efficiency of the system using adaptive QAM modulation where the power and symbol error rate are constrained. The numerical results of the optimal power and bit allocation are shown in Section 9. The improvement of system outage capacity in time-varying environment through optimal power allocation with channel tracking are shown in Section 10. Finally, Section 11 contains the conclusions.

2. CHANNEL ESTIMATION

In an OFDM system, the frequency response at the kth tone corresponding to the channel from the pth transmit antenna to the qth receive antenna is given by ([5, 7])

$$H_k^{(qp)} = \sum_{\ell=0}^{L-1} h_\ell^{(qp)} W_K^{k\ell}, \tag{1}$$

where $h_\ell^{(qp)}$ is the ℓth channel impulse of

$$\mathbf{h}^{(qp)} = \left[h_0^{(qp)}, \ldots, h_{L-1}^{(qp)} \right]^H, \tag{2}$$

which is the channel impulse response from the pth transmit antenna to the qth receive antenna, $W_K = e^{-j(2\pi/K)}$, and K is the total number of spectral channels (tones) in the OFDM system, T_s is the sampling interval of the system, and $1/T_s$ is the entire channel bandwidth of the OFDM system. $T_f = (K + v)T_s$ is the OFDM frame length where v is the number of cyclic prefix symbols. We usually have $K \gg v$ for bandwidth efficiency and let $v \geq L - 1$ to avoid interblock interference; therefore, $T_f \gg LT_s$, which also induces flat frequency response in each subchannel. Given N_t transmit antenna and N_r receive antenna, the received signal in the kth subband at the qth receive antenna can be expressed as

$$y_k^{(q)} = \sum_{p=1}^{N_t} H_k^{(qp)} x_k^{(p)} + \zeta_k^{(q)}, \tag{3}$$

for $q = 1, \ldots, N_r$, $k = 0, \ldots, K - 1$, where $x_k^{(p)}$ is the transmit signal at the pth transmit antenna, and $\zeta_k^{(q)}$ is the noise at the qth receive antenna. For the sake of conciseness, the index q for different antennas is omitted from $H_k^{(qp)}$, $h_\ell^{(qp)}$, $y_k^{(q)}$, $\zeta_k^{(q)}$. Based on (3), we propose a channel estimation approach by forming a matrix equation with all the unknown channel parameters in one vector, and estimate that vector by the least-square (LS) method. Given K subchannels, the received vector can be expressed as

$$\mathbf{y} = \sum_{p=1}^{N_t} \mathbf{X}^{(p)} \mathbf{g}^{(p)} + \zeta, \tag{4}$$

where $\mathbf{y} = [y_0, \ldots, y_{K-1}]^H$, $\zeta = [\zeta_0, \ldots, \zeta_{K-1}]^H$, and $\mathbf{X}^{(p)}$ is a diagonal matrix with the kth diagonal element as $x_k^{(p)}$, and $\mathbf{g}^{(p)}$ is a $K \times 1$ vector in which its kth element is $H_k^{(p)}$. Define

$$\mathbf{F} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_K^1 & \cdots & W_K^{(L-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_K^{K-1} & \cdots & W_K^{(K-1)(L-1)} \end{bmatrix}_{K \times L}, \tag{5}$$

then

$$\mathbf{g}_{K \times 1}^{(p)} \equiv \begin{bmatrix} H_0^{(p)} \\ \vdots \\ H_{K-1}^{(p)} \end{bmatrix} = \mathbf{F} \begin{bmatrix} h_0^{(p)} \\ \vdots \\ h_{L-1}^{(p)} \end{bmatrix} \equiv \mathbf{F} \mathbf{h}_{L \times 1}^{(p)}. \tag{6}$$

The channel can be modeled as an FIR filter with L taps and sampling interval T_s. Combining (4) and (6), we obtain

$$\mathbf{y} = \sum_{p=1}^{N_t} \mathbf{X}^{(p)} \mathbf{F} \mathbf{h}^{(p)} + \zeta = \mathbf{A} \underline{\mathbf{h}} + \zeta, \tag{7}$$

where

$$\mathbf{A}_{K \times LN_t} = \left[\mathbf{X}^{(1)} \mathbf{F}, \ldots, \mathbf{X}^{(N_t)} \mathbf{F} \right],$$
$$\underline{\mathbf{h}}_{LN_t \times 1} = \left[\mathbf{h}^{(1)H}, \ldots, \mathbf{h}^{(N_t)H} \right]^H. \tag{8}$$

By the LS method, the multiple channels can be estimated by

$$\hat{\underline{\mathbf{h}}} = \mathbf{A}^+ \mathbf{y}, \tag{9}$$

where the pseudo-inverse of matrix $\mathbf{A}^+ = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ if $K \geq LN_t$. Because $\text{rank}(\mathbf{A}) = \min(K, KN_t, LN_t) = \min(K, LN_t)$, the necessary and sufficient condition to have unique LS channel-estimate is $K \geq LN_t$. The advantage of this LS approach is its low-complexity and it requires no a priori statistical knowledge about the channel and the noise. In the training mode, since the training sequences have been specified, $(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ can be precomputed and stored such that the operation of the initial channel acquisition can be further speeded up.

3. MSE BOUND OF LS ESTIMATOR

Combining (7) and (9), we obtain

$$\hat{\underline{\mathbf{h}}} = \underline{\mathbf{h}} + \mathbf{A}^+ \zeta, \tag{10}$$

which indicates that the estimate $\hat{\underline{\mathbf{h}}}$ is the combination of the true $\mathbf{h}$ plus a term affected by the estimation error. Since $E\{\hat{\underline{\mathbf{h}}}\} = \underline{\mathbf{h}} + \mathbf{A}^+ E\{\mathbf{1}\} = \mathbf{h}$, $\hat{\underline{\mathbf{h}}}$ is an unbiased estimator. The MSE of estimation is given by

$$\begin{aligned} \text{MSE} &= \frac{1}{LN_t} E\left\{ \left\| \hat{\underline{\mathbf{h}}} - \underline{\mathbf{h}} \right\|^2 \right\} \\ &= \frac{1}{LN_t} E\left\{ \left\| \mathbf{A}^+ \mathbf{1} \right\|^2 \right\} \\ &= \frac{1}{LN_t} \text{Tr} \left\{ E\{ \mathbf{A}^+ \zeta \zeta^H \mathbf{A}^{+H} \} \right\} \\ &= \frac{\sigma_n^2}{LN_t} \text{Tr} \left\{ (\mathbf{A}^H \mathbf{A})^{-1} \right\}, \end{aligned} \tag{11}$$

where the superscript H denotes the Hermitian operation, $\text{Tr}\{\cdot\}$ denotes the trace of a square matrix and $E\{\zeta \zeta^H\} =$

$\sigma_n^2 \mathbf{I}_K$. Define $\mathbf{B} = (\mathbf{A}^H \mathbf{A})^{-1}$, then $\mathbf{B}_{LN_t \times LN_t}$ becomes

$$\begin{bmatrix} \mathbf{F}^H \mathbf{X}^{(1)H} \mathbf{X}^{(1)} \mathbf{F} & \cdots & \mathbf{F}^H \mathbf{X}^{(1)H} \mathbf{X}^{(N_t)} \mathbf{F} \\ \mathbf{F}^H \mathbf{X}^{(2)H} \mathbf{X}^{(1)} \mathbf{F} & \cdots & \mathbf{F}^H \mathbf{X}^{(2)H} \mathbf{X}^{(N_t)} \mathbf{F} \\ \vdots & \ddots & \vdots \\ \mathbf{F}^H \mathbf{X}^{(N_t)H} \mathbf{X}^{(1)} \mathbf{F} & \cdots & \mathbf{F}^H \mathbf{X}^{(N_t)H} \mathbf{X}^{(N_t)} \mathbf{F} \end{bmatrix}^{-1}. \quad (12)$$

Suppose $x_k^{(p)'}s$, for $p = 1, \ldots, N_t$, all n and k, are constant modulus, and $|x_k^{(p)}|^2 = 1$, then $\mathbf{F}^H \mathbf{X}^{(p)H} \mathbf{X}^{(p)} \mathbf{F} = K\mathbf{I}$. Denote the (i, i)th component of $\mathbf{B}$ as b_i, then

$$\text{MSE} = \frac{\sigma_n^2}{LN_t} \sum_{i=1}^{LN_t} b_i \geq \frac{\sigma_n^2}{LN_t} LN_t \sqrt[LN_t]{\prod_{i=1}^{LN_t} b_i}, \quad (13)$$

with equality when $b_1 = b_2 = \cdots = b_{LN_t}$. By Hardamard's inequality,

$$\prod_{i=1}^{LN_t} b_i \geq \det(\mathbf{B}); \qquad K^{LN_t} \geq \det(\mathbf{B}^{-1}) = [\det(\mathbf{B})]^{-1}, \quad (14)$$

with equality if and only if $\mathbf{B}$ is diagonal. Finally, combining (13) and (14), we obtain

$$\text{MSE} \geq \frac{\sigma_n^2}{K}, \quad (15)$$

with equality if and only if $\mathbf{B}$ is a diagonal matrix and all the diagonal components b_i, $i = 1, \ldots, LN_t$ are equal. Note that (15) corroborates the results in [5] where the case $N_t = 2$ was considered. The merits of our derivation is that we can obtain the MSE bound of the channel estimator without computing the inverse of the matrix analytically in (11), which is a computationally-intractable work when $N_t > 2$.

4. OPTIMUM TRAINING SIGNAL

According to (15), in order to make the channel estimator achieve its best performance at the channel acquisition mode, the training signals from multiple antennas should satisfy

$$\mathbf{F}^H \mathbf{X}^{(r)H} \mathbf{X}^{(s)} \mathbf{F} = \begin{cases} \mathbf{0}_{L \times L}, & r \neq s, \\ c\mathbf{I}_{L \times L}, & r = s, \end{cases} \quad (16)$$

where c is a constant, $1 \leq r, s \leq N_t$. Therefore, $\mathbf{B}$ is guaranteed to be a diagonal matrix with the same diagonal components. To satisfy the constraints, the optimal training sequence can be designed as

$$x_k^{(p)} = (-1)^{\lfloor k/2^{(p-2)} \rfloor} x_k^{(1)}, \quad p = 1, \ldots, N_t, \quad (17)$$

where $|x_k^{(1)}| = c$ for all k. The verification that the designed training sequences can satisfy the constraints is given in [8]. It is concluded that given N_t transmit antennas, the channel estimator can achieve MSE lower bound, as long as $K \geq 2^{N_t-1}L$. Figure 1 shows that optimum training
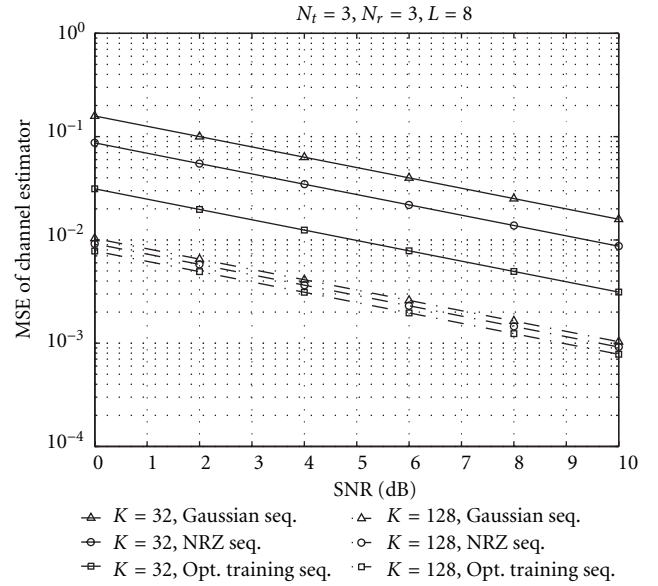

FIGURE 1: Channel acquisition error induced by different training sequences.

sequences provide the minimum mean square channel acquisition error as compared to randomly generated sequences.

5. TRACKING PERFORMANCE IN TIME-VARYING CHANNELS

In the simulations, channels with three-ray and a Doppler frequency of 10 Hz and 100 Hz are used to characterize the time-varying multipath propagation environments. Each resolvable multipath is independent of each other and characterized by Ricean fading where the Rayleigh fading part is generated using Jake's model [8]. Two transmitter antennas ($N_t = 2$) with multiple receiver antennas ($N_r = 2, 3, 4$) are used for diversity. The parameters for the simulated OFDM system are as follows. The entire channel bandwidth $f_s = 240$ kHz is divided into $K = 128$ subchannels. QAM signaling is applied. To make the tones orthogonal to each other, the OFDM frame duration is $K/f_s = 533 \mu s$. An additional 83 μs guard interval is used to provide protection from interblock interference. The channel multipath delay spread is assumed $(L - 1)/f_s = 83 \mu s$ and $L = 3$. This results in a total block length $T_f = 616 \mu s$ and a subcarrier symbol rate 1.623 kBaud. In the data transmission mode, the channel estimates of the previous OFDM block are used to decode the current OFDM symbols. The performance of the system is measured by bit error rate (BER) and the channel estimation mean square error (MSE), each averaged over 2500 OFDM blocks. Figure 2 shows BER performance versus SNR using beamforming plus successive interference cancellation (denoted as QR beamforming [7]) with path gain ordering for $N_t = 2$, $N_r = 2, 3, 4$, and $f_d = 10$ Hz and 100 Hz. Suppose that the average received SNR = 8 dB, then for $f_d = 100$ Hz, the resulting BER is around 10^{-2} for $N_r = 3$ and approaches
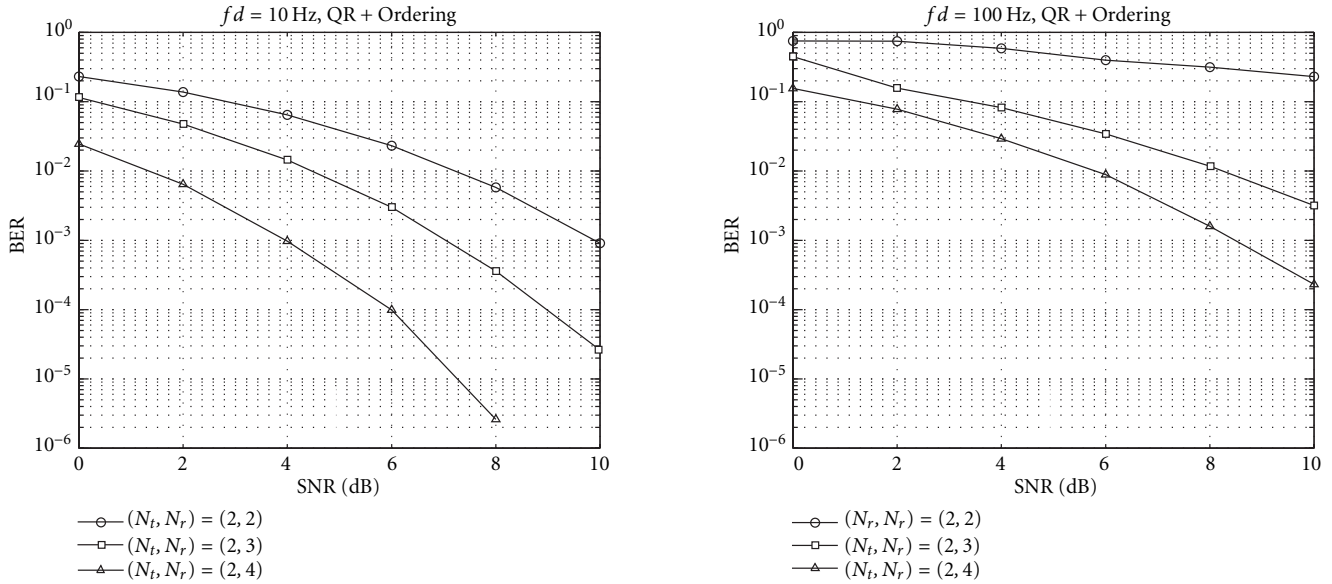
FIGURE 2: BER versus SNR using beamforming with successive interference cancellation (QR beamforming) and path gain ordering.
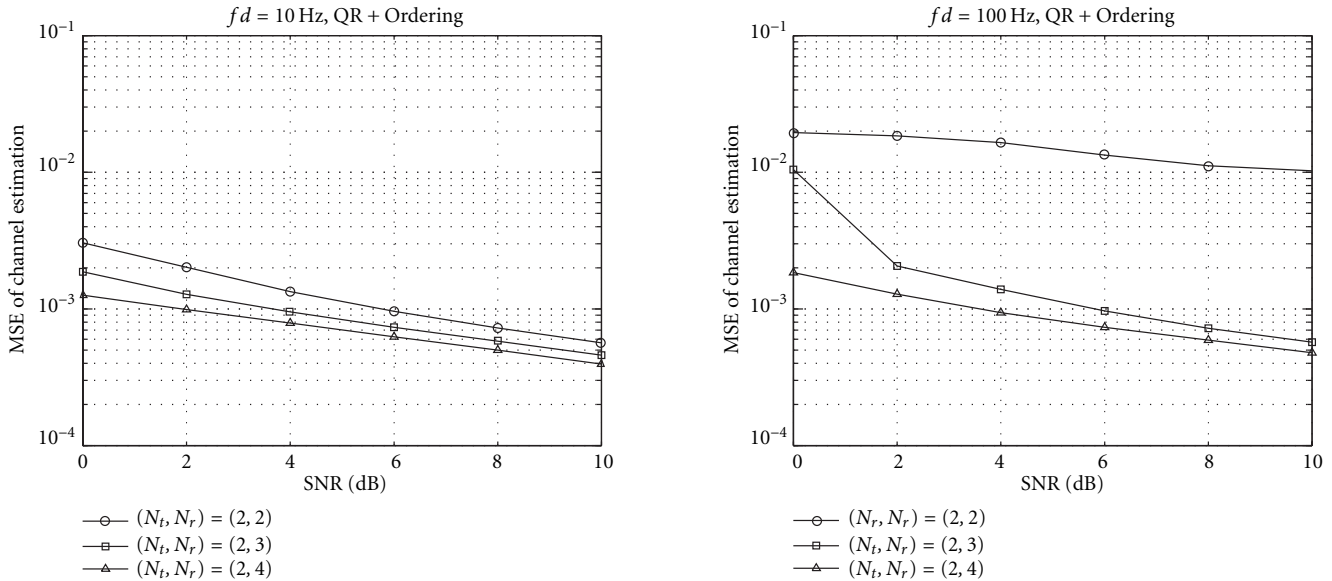

FIGURE 3: Channel estimation error versus SNR using QR beamforming and path gain ordering.

10^{-3} for $N_r = 4$. For $f_d = 10$ Hz, the BER reduces to 0.5×10^{-3} and 10^{-6}, respectively. This shows that the BER reduces when more receive antennas are used, especially under low Doppler frequency conditions. Figure 3 shows the channel estimation error for $N_t = 2$, $N_r = 2, 3, 4$, and $f_d = 10$ Hz and 100 Hz when QR beamforming with path gain ordering is applied. We observe that the channel estimation performance improves when N_r increases, SNR increases and f_d decreases. It is observed that when the Doppler frequency is small, the rate that channel fades is slow; thus, better channel tracking performance can be attained. In addition, when more receive

antennas are used, more spatial diversity can result in better chance to successfully detect the data.

6. CAPACITY OF OFDM MULTIPLE-ANTENNA SYSTEM

In multiple-antenna OFDM system, the demodulated signal in the kth subband at the qth receive antenna is denoted by $y_k^{(q)} = \sum_{p=1}^{N_t} H_k^{(qp)} x_k^{(p)} + \zeta_k^{(q)}$, given N_t transmit antennas and N_r receive antennas, the received vector in the kth subband is given as $\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{\imath}_k$, $k = 0, \ldots, K - 1$,

where $\mathbf{y}_k = [y_k^{(1)}, \ldots, y_k^{(N_r)}]^H$, $\mathbf{x}_k = [x_k^{(1)}, \ldots, x_k^{(N_t)}]^H$, $\mathbf{1}_k = [\zeta_k^{(1)}, \ldots, \zeta_k^{(N_r)}]^H$, and the (q, p)th entry of $\mathbf{H}_k$ is $H_k^{(qp)}$. The information capacity of the kth subchannel is given as

$$
\begin{aligned}
C_k &= \max_{f(\mathbf{x}_k): \sum Ex_k^{(p)} \leq P} I(\mathbf{x}_k; \mathbf{y}_k) \\
&= \max_{f(\mathbf{x}_k): \sum Ex_k^{(p)} \leq P} h(\mathbf{y}_k) - h(\zeta_k).
\end{aligned}
\tag{18}
$$

When the noise ζ_k is Gaussianly distributed, its entropy is given by

$$
h(\zeta_k) = \frac{1}{2} \log_2 (2\pi e)^{N_r} \det(\mathbf{N}_k),
\tag{19}
$$

where $\mathbf{N}_k = E\{\zeta_k \zeta_k^H\}$. If the noise is independent of the input, the capacity is obtained by maximizing the entropy of the received signals which occurs when received signals and input signals have Gaussian distributions. Thus,

$$
\max_{f(\mathbf{x}_k): \sum Ex_k^{(p)} \leq P} h(\mathbf{y}_k) = \frac{1}{2} \log_2 (2\pi e)^{N_r} \det(\mathbf{Y}_k),
\tag{20}
$$

where $\mathbf{Y}_k = E\{\mathbf{y}_k \mathbf{y}_k^H\} = \mathbf{H}_k \mathbf{D}_k \mathbf{H}_k^H + \mathbf{N}_k$, $\mathbf{D}_k = E\{\mathbf{x}_k \mathbf{x}_k^H\}$, $\mathbf{N}_k = E\{\zeta_k \zeta_k^H\}$. Combining (19) and (20), yields

$$
C_k = \log_2 \left(\frac{\det (\mathbf{H}_k \mathbf{D}_k \mathbf{H}_k^H + \mathbf{N}_k)}{\det (\mathbf{N}_k)} \right) \text{(bit/s/Hz).}
\tag{21}
$$

If uniform power allocation is used (i.e., $\mathbf{D}_k = P/(KN_t)$) and the Gaussian noise is spatially white with equal power at different receive antennas (i.e., $\mathbf{N}_k = n_k \mathbf{I}$), (21) reduces to

$$
C_k = \log_2 \left(\det \left(\mathbf{I} + \frac{P}{KN_t n_k} \mathbf{H}_k \mathbf{H}_k^H \right) \right),
\tag{22}
$$

which is consistent to equation (4) in [4]. Assuming that the OFDM frames and carriers are synchronized, and cyclic prefix (CP) is applied, the K subbands in each transmitter can be considered as K parallel channels and the overall channel capacity of the OFDM multiple-antenna system is given by

$$
\begin{aligned}
C &= \sum_{k=0}^{K-1} C_k \\
&= \sum_{k=0}^{K-1} \log_2 \left(\det \left(\mathbf{I} + \frac{1}{n_k} \mathbf{H}_k \mathbf{D}_k \mathbf{H}_k^H \right) \right).
\end{aligned}
\tag{23}
$$

7. OPTIMAL POWER ALLOCATION

It is observed that, given the channel transfer functions $\mathbf{H}_k$, the channel capacity is a function of $\mathbf{D}_k$, whose diagonal elements correspond to the allocated power. In order to maximize the channel capacity, an optimization problem is formulated as

$$
C = \max_{\mathbf{D}_0, \ldots, \mathbf{D}_{K-1}} \sum_{k_0}^{K-1} C_k, \quad \text{s.t.} \sum_{k=0}^{K-1} \text{tr} (\mathbf{D}_k) = P,
\tag{24}
$$

where P is total power allocated to $\{\mathbf{x}_k\}_{k=0,K-1}$. Since CP is applied, the total transmission power is indeed $((K + \nu/K)P)$ where ν is the length of the CP. To solve (24), we first decompose $\mathbf{H}_k$ by singular value decomposition (SVD), $\mathbf{H}_k = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^H$, where $\mathbf{U}_k \mathbf{U}_k^H = \mathbf{I}_{N_r \times N_r}$, $\mathbf{V}_k \mathbf{V}_k^H = \mathbf{I}_{N_t \times N_t}$, the (p, p)th entry of $\mathbf{S}_k$ is $s_k^{(p)}$, $k = 0, \ldots, K - 1$, $p = 1, \ldots, N_t$. Thus, C_k in (21) can be expressed as

$$
C_k = \log_2 \left(\det \left(\mathbf{I} + \frac{\mathbf{S}_k \mathbf{V}_k^H \mathbf{D}_k \mathbf{V}_k \mathbf{S}_k^H}{n_k} \right) \right).
\tag{25}
$$

Define $\mathbf{W}_k = \mathbf{V}_k^H \mathbf{D}_k \mathbf{V}_k$, C_k in (25) is maximized when $\mathbf{W}_k$ is a diagonal matrix (see the appendix). Since $\text{tr}(\mathbf{W}_k) = \text{tr}(\mathbf{V}_k^H \mathbf{D}_k \mathbf{V}_k) = \text{tr}(\mathbf{V}_k \mathbf{V}_k^H \mathbf{D}_k) = \text{tr}(\mathbf{D}_k)$, the optimization problem in (24) is reduced to

$$
\begin{aligned}
C = \max_{\mathbf{W}_0, \ldots, \mathbf{W}_{K-1}} \sum_{k=0}^{K-1} \sum_{p=1}^{N_t} \log_2 \left(1 + \frac{s_k^{(p)^2} w_k^{(p)}}{n_k} \right), \\
\text{s.t.} \sum_{k=0}^{K-1} \sum_{p=1}^{N_t} w_k^{(p)} = P,
\end{aligned}
\tag{26}
$$

where the (p, p)th entry of $\mathbf{W}_k$ is denoted by $w_k^{(p)}$. The extended 2D spatial-spectral water filling solution is then obtained as

$$
w_k^{(p)} = \left(\nu - \frac{n_k}{s_k^{(p)^2}} \right)^+,
\tag{27}
$$

where $(\bullet)^+$ is equal to the argument $\bullet$ if $\bullet$ is positive and set to zero if $\bullet$ is negative. ν is chosen so that

$$
\sum_{k=0}^{K-1} \sum_{p=1}^{N_t} \left(\nu - \frac{n_k}{s_k^{(p)^2}} \right)^+ = P.
\tag{28}
$$

Note that making $\mathbf{W}_k$ diagonal implies a precoloring processing at the transmitter side and a beamforming processing at the receiver side. That is, the transmit signal vector is precolored by setting $\mathbf{x}_k = \mathbf{V}_k \mathbf{w}_k$. Since the input data stream are spatially and temporally white, $\mathbf{W}_k = E\{\mathbf{w}_k \mathbf{w}_k^H\} = \mathbf{I}$ attains a diagonal form. On the other hand, the beamforming weight vector which steers the beam to the pth transmit antenna and places nulls to the others is given by the pth column of $\mathbf{U}_k$. Referring now to [6], a spatio-temporal vector-coding communication structure is suggested as a means for achieving capacity. Here, we tackle the capacity-maximization problem in opposite direction. Upon maximizing the channel capacity using Hadamard's inequality, we show that a spatio-temporal vector coding processing (a precoloring processing) is required to achieve the capacity.

8. OPTIMAL BIT ALLOCATION SUBJECT TO POWER AND LOS (SER) CONSTRAINTS

In this section, we apply the above optimal power allocation method to a practical adaptive QAM modulation scheme for

each subband. The symbol error rate (SER) of a $2^b - QAM$ signaling scheme is given in [2] by

$$P_e = MQ\left(\sqrt{\frac{3P_o|H|^2}{N(2^b-1)}}\right), \tag{29}$$

where M is average number of neighbors for symbols in the QAM constellation. P_o, H, N, and 2^b are the transmit power, path gain, noise power, and QAM size, respectively. For $M = 4$, denote $\gamma = Q^{-1}(P_e/4)$, then the resultant bits per symbol can be expressed as

$$b = \log_2\left(1 + \frac{3}{\gamma^2}\frac{P_o|H|^2}{N}\right). \tag{30}$$

Accordingly, for the multiple antennas OFDM system, the optimal allocated bit in the kth subcarrier for the pth spatial channel is given as

$$b_{pk} = \log_2\left(1 + \frac{3}{\gamma^2}\frac{w_k^{(p)}s_k^{(p)2}}{n_k}\right). \tag{31}$$

Thus, given a total allocated power P, subband bandwidth W, and the required SER, the total transmission rate R is maximized by

$$R = \max_{w_k^{(p)}}\left(\frac{K}{K+\nu}\right)W\sum_{k=0}^{K-1}\sum_{p=1}^{N_t}b_{pk}, \quad \text{s.t.} \sum_{k=0}^{K-1}\sum_{p=1}^{N_t}w_k^{(p)} = P. \tag{32}$$

The optimal allocated power is then solved as

$$w_k^{(p)} = \left(\nu - \frac{\gamma^2 n_k}{3s_k^{(p)2}}\right)^+, \tag{33}$$

and the corresponding optimal allocated bit is

$$b_{pk} = \log_2\left(1 + \left(\frac{3}{\gamma^2}\right)\left(\frac{s_k^{(p)2}}{n_k}\right)\left(\nu - \frac{\gamma^2 n_k}{3s_k^{(p)2}}\right)^+\right), \tag{34}$$

where ν is chosen so that

$$\sum_{k=0}^{K-1}\sum_{p=1}^{N_t}\left(\nu - \frac{\gamma^2 n_k}{3s_k^{(p)2}}\right)^+ = P. \tag{35}$$

Note that the bandwidth efficiency defined by $R/KW = \sum_{k=0}^{K-1}\sum_{p=1}^{N_t}b_{pk}/(K+\nu)$, where KW is the total bandwidth. Also note that the optimal bit provided by (31) need not be an integer. To satisfy the integer constraints, an iterative algorithm proposed in [3] can be used. To reduce the gap r, many other practical coding techniques such as trellis coding [9] and coset coding [10] can also be applied.
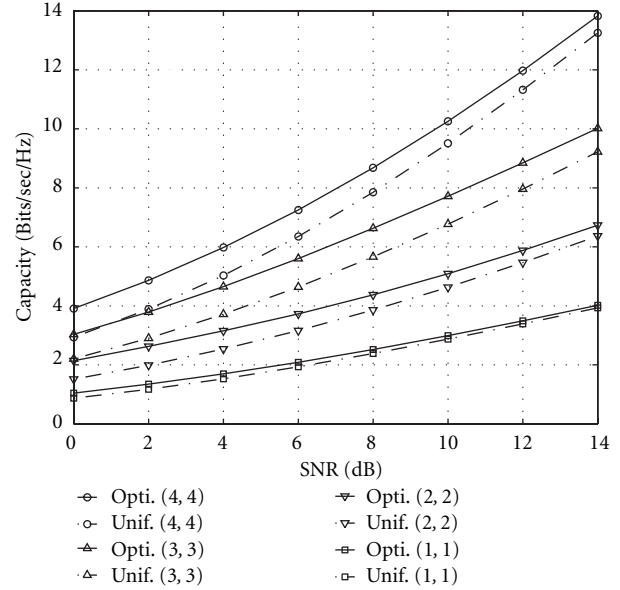

FIGURE 4: Capacities with optimal power allocation (opti.) and uniform power allocation (unif.) versus SNR for an OFDM (N_t, N_r) multiple-antenna system.

9. NUMERICAL RESULTS ON OPTIMAL POWER AND BIT ALLOCATION

In the simulations, sampling frequency f_0 is normalized to 1, the number of subchannels $K = 128$, cyclic prefix length $\nu = 4$, OFDM frame duration is $(K+\nu)/f_0$, the maximum multipath delay spread is $(\nu-1)/f_0$. The channel from the pth transmit antenna to the qth receive antenna is modeled by two rays as $h^{(pq)}(t) = \alpha^{(pq)}(\delta(t) + A^{(pq)}e^{\sqrt{-1}\theta^{(pq)}}\delta(t-\tau^{(pq)}))$, where $\tau^{(pq)} \in \{1/f_0, 2/f_0, 3/f_0\}$, $\alpha^{(pq)}$ is drawn from a uniform distributions $[0,1]$, and under severe frequency selective fading conditions of $A^{(pq)} \rightarrow 1$, and $\theta^{(pq)} \rightarrow \pi$ [11]. Note that the total allocated power is defined by $P = \sum_{k=0}^{K-1}\sum_{p=1}^{N_t}E\{|x_k^{(p)}|^2\}$, and the average SNR ratio at the receiver is defined by

$$\text{SNR} = \sum_{k=0}^{K-1}\sum_{p=1}^{N_t}\frac{E\{|H_k^{(qp)}x_k^{(p)}|^2\}}{\left(N_r\sum_{k=0}^{K-1}n_k\right)}. \tag{36}$$

Figure 4 shows the capacities of an OFDM (N_t, N_r) multiple-antenna system using optimal (solid curves using (27) in (26)) and uniform (dash curves using $w_k^{(p)} = P/(KN_t)$ in (26)) power allocations. It is observed that a significant increase in capacity can be obtained when the optimal power allocation is applied. It also shows that the channel capacity increases with the total allocated power and the number of antennas (N_t, N_r). To illustrate the mechanism of optimal bit allocation using the adaptive QAM modulation, Figure 5 shows both the channel gain and the allocated bits in each subchannel for a specific $N_t = N_r = 2$ and $K = 128$ example. It is of interest to observe that given fixed transmit power, more power is allocated to higher gain subchannels such that
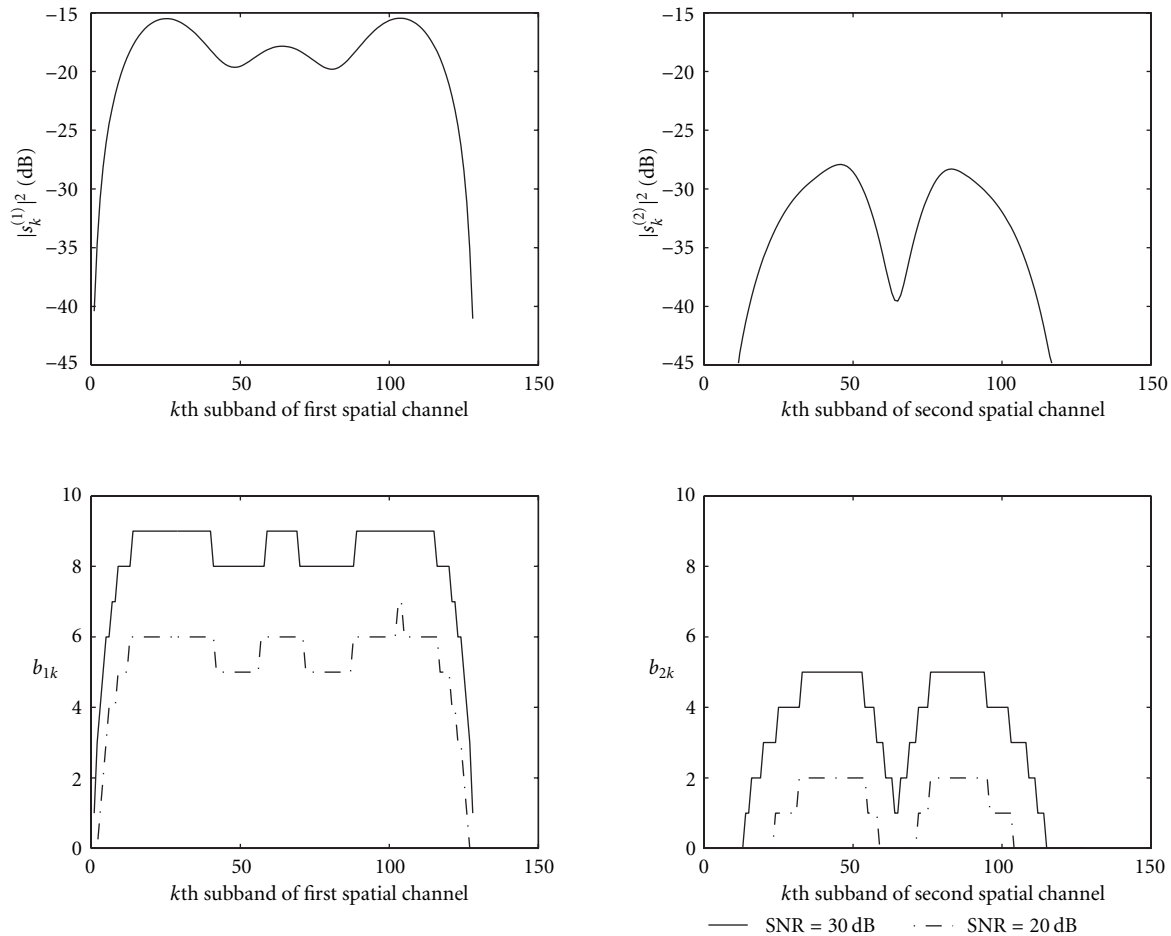
FIGURE 5: Channel state information and optimal bit allocation for OFDM system with 2 transmit and 2 receive antennas with $K = 128$ and SER $= 10^{-5}$.

more bits can be efficiently transmitted. Figure 6 shows the bandwidth efficiency R/KW obtained by different power allocation approaches under different QoS (SER) and received power (SNR) conditions. These plots show that under low SNR and high QoS conditions, optimal power allocation has significant advantage over that of uniform power allocation.

10. CAPACITY IMPROVEMENT

When the system capacity is observed as a random variable in a time-varying environment, its complementary cumulative distribution functions (ccdfs) can be used to show what fraction of time or probability, a specified capacity can be attained [4]. For example, the capacity that can be attained with a 99% probability results in a 1% probability of outage. In order to reduce the outage probability and increase the channel capacity in a time-varying environment, we adaptively performing optimal power allocation among multiple-antenna and multiple-subband based on updated channel state information (CSI). In this application, it is important to keep the updating rate high enough so that the time-varying channel can be tracked.

To evaluate the effect of CSI updating on outage capacity, we assume the estimated CSI at the receivers are sent back to the transmitters every T_d seconds, where the loop delay is included which accounts for the two-way signal propagation delay and the time delay involved in generating and executing the updating command. Based on this mechanism, the mean square error of the channel tracking error is given by $E\{|h(t) - h(t - T_d)|^2\}$, where $h(t)$ is the path gain at time t. Since each of the resolvable multipath components is actually a combination of many unresolvable paths with varying amplitudes, phases, and direction of arrivals. The path gain can be modeled as a zero mean complex-valued colored Gaussian random process with an auto correlation function $J_0(2\pi f_d \tau)$, where f_d is the maximum Doppler shift, and τ is the time difference. Thus, $E\{h(t)h^*(t - \tau)\} = J_0(2\pi f_d \tau)$. Accordingly, $E\{|h(t) - h(t - T_d)|^2\} = 2(1 - J_0(2\pi f_d T_d))$. Since $h(t)$ is zero mean complex Gaussian, $h(t) - h(t - T_d)$ is also zero mean complex Gaussian. Consequently, the variance of $h(t) - h(t - T_d)$, denoted as σ_h^2, is computed as $2(1 - J_0(2\pi f_d T_d))$. Therefore, we model the tracked channel as $\hat{h}(t) = h(t) + \sigma_h(T_d)n(t)$, where $h(t)$ is the true channel and $\sigma_h(T_d)n(t)$ is the channel tracking error. Both $h(t)$
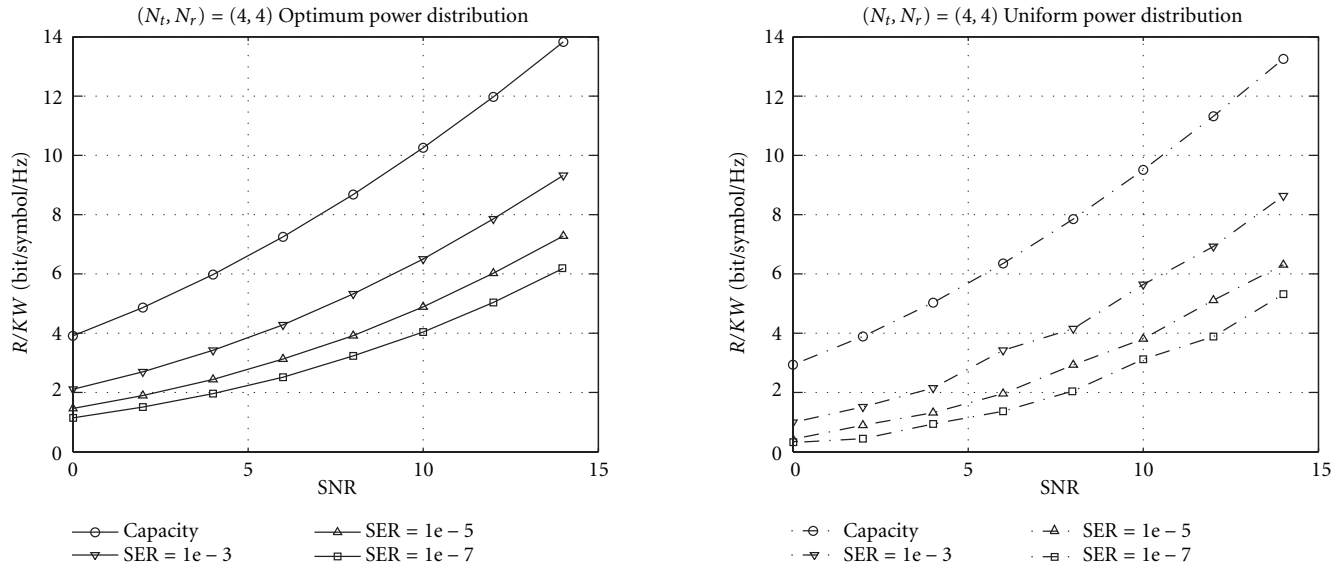
FIGURE 6: Bandwidth efficiency by optimal and uniform power allocations s.t. symbol error rate and power constraints.
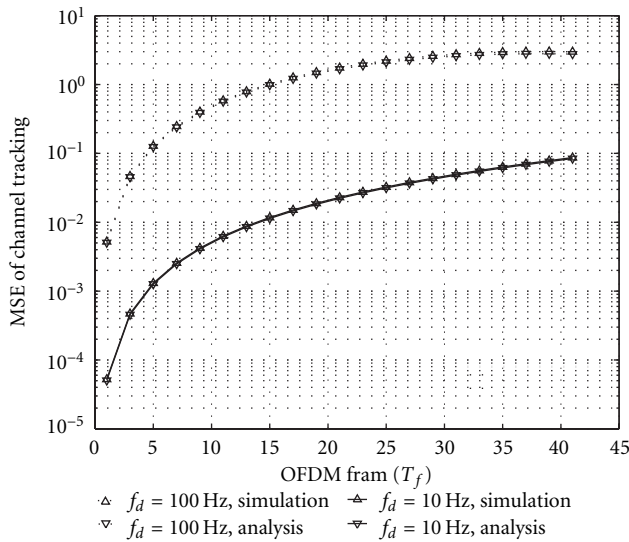

FIGURE 7: Analytical and simulated MSE of channel tracking versus T_d.

FIGURE 8: Capacity improvement using optimal power allocation with various T_d under $f_d = 10$ Hz, SNR = 5 dB, $K = 64$, $(N_r, N_t) = (3, 3)$ or (4.4).

and $n(t)$ are zero mean complex-valued Gaussian random variable's with unit variance. Figure 7 shows both analytical and simulated MSE of channel tracking versus T_d (in units of T_f) for $f_d = 10$ Hz and $f_d = 100$ Hz, respectively. Figures 8 and 9 evaluate the effect of CSI updating on outage capacity. Figure 8 shows the effect of T_d, number of antennas, and power allocation schemes on the outage capacity for $T_d = 20T_f$, $60T_f$, $100T_f$ under $f_d = 10$ Hz, SNR = 5 dB, $K = 64$, $N_r = 3$, and $N_t = 3$ condition. We observe that optimal power allocation results in higher channel capacity compared to uniform power allocation ($w_k^{(p)} = P/(KN_t)$). We can also observe that the outage capacity increases when
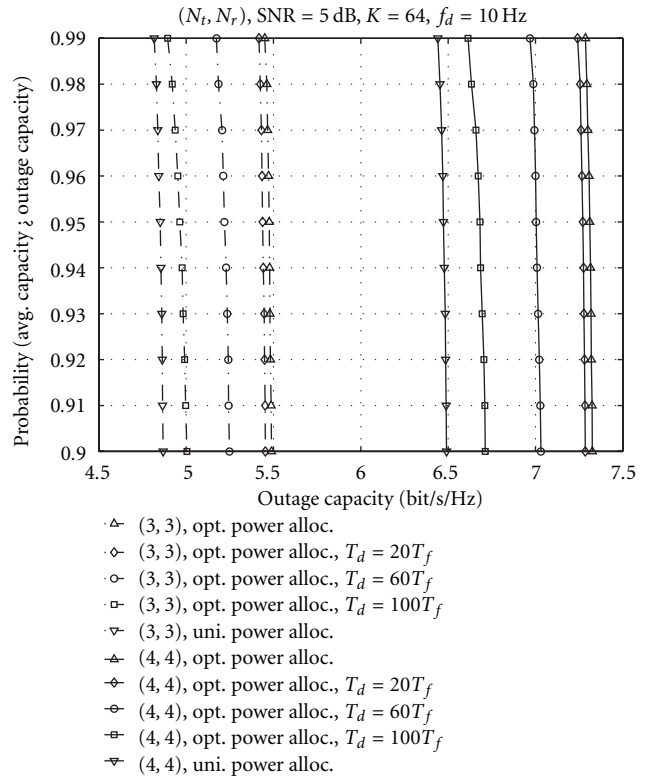
more antennas are exploited and shorter updating period T_d are used. Similarly, Figure 9 shows that the outage capacity uses $T_d = 3T_f$, $6T_f$, and $9T_f$ under $f_d = 100$ Hz.
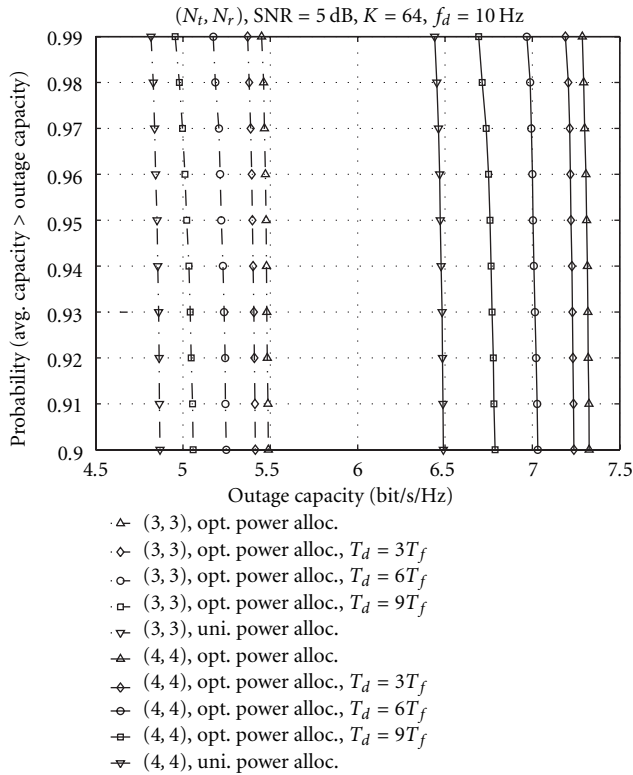
FIGURE 9: Capacity improvement using optimal power allocation with various T_d under $f_d = 100$ Hz, SNR = 5 dB, $K = 64$, $(N_r, N_t) = (3, 3)$ or $(4, 4)$.

11. CONCLUSIONS

In the first part of the paper, we propose channel estimation, acquisition and tracking methods for OFDM multiple-antenna system. Then, we derive the lower bound of channel estimator and design the optimal training sequences for achieving minimum acquisition error. The tracking and BER performance in time-varying channels is then evaluated.

In the second part of the paper, we introduce the channel capacity of an OFDM multiple-antenna system, and the optimal power allocation approach. Then, we apply optimal bit allocation method with adaptive QAM modulation to maximize the bandwidth efficiency subject to fixed transmit power and QoS constraints on the system. Simulation results show that with a multiple-antenna architecture, OFDM modulation, optimal power and bit allocation can dramatically increase the transmission rate and bandwidth efficiency when severe spatial and frequency selective fading occur. Finally, by combining optimal power allocation with real-time CSI updating, the channel outage capacity are shown to be dramatically improved by adaptive power allocation and multiple-antenna architecture in the time-varying environment.

APPENDIX

Hadamard's inequality states that the determinant of any positive definite matrix $\mathbf{K}$ is less than the product of its diagonal elements, that is, $\det(\mathbf{K}) \leq \prod_i \mathbf{K}_{ii}$ with equality if and only if the matrix is diagonal [12, page 233]. Since $\mathbf{D}_k = E\{\mathbf{x}_k \mathbf{x}_k^H\}$, $\mathbf{W}_k = \mathbf{V}_k^H \mathbf{D}_k \mathbf{V}_k \mathbf{a}^H (\mathbf{I} + \mathbf{S}_k \mathbf{W}_k \mathbf{S}_k^H) \mathbf{a} = |\mathbf{a}|^2 + E\{|\mathbf{a} \mathbf{S} \mathbf{V}_k^H \mathbf{x}_k|^2\} > 0$, for all $\mathbf{a} \neq 0$, $(\mathbf{I} + \mathbf{S}_k \mathbf{W}_k \mathbf{S}_k^H)$ is positive definite. By Hadamard's inequality,

$$\det \left(\mathbf{I} + \mathbf{S}_k \mathbf{W}_k \mathbf{S}_k^H\right) \leq \prod_i \left(1 + \left[\mathbf{S}_k \mathbf{W}_k \mathbf{S}_k^H\right]_{ii}\right) \qquad (A.1)$$

with equality if and only if $(\mathbf{I} + \mathbf{S}_k \mathbf{W}_k \mathbf{S}_k^H)$ is diagonal, which happens only when $\mathbf{W}_k$ is also diagonal since $\mathbf{I}$ and $\mathbf{S}_k$ are diagonal.

REFERENCES

[1] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. 45th IEEE Vehicular Technology Conference*, vol. 2, pp. 815–819, Chicago, Ill, USA, July 1995.

[2] A. R. S. Bahai and B. R. Saltzberg, *Multi-carrier Digital Communications—Theory and Application of OFDM*, Kluwer, 1999.

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

[4] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

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

[6] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Communications*, vol. 46, no. 3, pp. 357–366, 1998.

[7] T.-L. Tung and K. Yao, "Spatio-temporal parallel processing and complexity reduction for OFDM multiple antenna system," in *Proc. IEEE Sips2000*, vol. 2, October 2000.

[8] T.-L. Tung, *Adaptive space-time signal processing for wireless communication and sensor systems*, Ph.D. thesis, University of California, Los Angeles, Calif, USA, 2001.

[9] G. Ungerboeck, "Trellis-coded modulation with redundant signal sets: parts I and II," *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 5–21, 1987.

[10] G. D. Forney Jr., "Coset codes—part I: introduction and geometrical classification," *IEEE Transactions on Information Theory*, vol. 34, no. 5, pp. 1123–1151, 1988.

[11] M. Rice, D. de Gaston, A. Davis, G. German, and C. Bettwieser, "ARTM channel sounding results—an investigation of frequency-selective fading on aeronautical telemetry channels," in *Proc. International Telemetering Conference*, Las Vegas, Nev, USA, October 1999.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley and Sons, New York, USA, 1991.

Tai-Lai Tung was born in Taidung, Taiwan, in 1972. He received his B.S. degree in control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1994 and the M.S. and the Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), in December 1997 and March 2001, respectively. From 1994 to 1996, he was an Ensign Instructor at the Naval Communication and Electronics School, Taiwan. Since 1998, he had been a Research Assistant at UCLA. He is currently a Senior DSP Engineer in GlobespanVirata, Irvine Office. His research interests include array signal processing, adaptive signal processing, digital communications, and system optimization.

Kung Yao received his B.S.E. (Highest Honors), M.A., and Ph.D. degrees in electrical engineering all from Princeton University, Princeton, NJ. During the summers, he has worked at the Princeton-Penn Accelerator, the Brookhaven Nat. Lab. and the Bell Telephone Lab. in Murray Hill, NJ. He was a NAS-NRC Post-Doctoral Research Fellow at the University of California, Berkeley. Presently, he is a Professor in the Electrical Engineering Department at UCLA. In 1969, he was a Visiting Assistant Prof. at the Massachusetts Institute of Technology. In 1973–74, he was a Visiting Associate Prof. at the Eindhoven Technical University. In 1985–1988, he served as an Assistant Dean of the School of Engineering and Applied Science at UCLA. His research and professional interests include sensor array system, digital communication theory and system, smart antenna and wireless radio system, chaos communications and system theory, digital and array signal and array processing, systolic and VLSI algorithms, architectures and systems, radar system, and simulation. He has published over 250 journal and conference papers. Dr. Yao received the IEEE Signal Processing Society's 1993 Senior Award in VLSI Signal Processing. He is the co-editor of a two volume series of an IEEE Reprint Book on "High Performance VLSI Signal Processing," IEEE Press, 1997. He is also the co-author of the book, "Processing and Algorithms in Communication and Radar Systems," under preparation. He has served as Program Chair, Secretary, and Chair of the IEEE Inform. Theory Group in L.A. and served two terms as a member of the Board of Governors of the IEEE Inform. Theory Group. He was the Co-Chair of the 1981 Inter. Symposium on Inform. Theory and the Representative of the IT-BOG of the 1987 IEEE Inform. Theory Workshop. Dr. Yao is a member of the VLSI Technical Committee of the IEEE Signal Processing Society. He was also the Chair of the Technical Program and the General Chair of the 1990 and 1992 IEEE Workshop on VLSI Signal Processing. He has served as an Associate Editor for Book Reviews of the IEEE Trans. on Information Theory and was an Associate Editor of the journal, "Probability in the Engineering and Informational Sciences." In 1991–1993, he was the Associate Editor of VLSI Signal Processing of the IEEE Trans. on Circuits and Systems. Since 1999, he is an Associate Editor of the IEEE Communications Letters. He is a member of the Editorial Board of Journal of VLSI Signal Processing and Integration: the VLSI Journal. He is also a Guest Editor of a Special Issue on "Applications of Chaos in Modern Communication Systems" of the IEEE Trans. on Circuits and Systems-Part I to appear in December 2001. In recent years, he has organized various sessions in array processing, wireless radio, and chaotic communication for various conferences. He is a Fellow of IEEE.

Dr. Yao has extensive practical system experiences in digital and satellite communication engineering, DSP, radar system design, link budget analysis, sensor system, and systolic and microphone array designs. He has worked or consulted for AT&T Bell Laboratory, NCR, SRI, TRW, Linknet, Hughes Aircraft Company, and Raytheon Systems, and so on.