

# Performance Analysis and Transceiver Design for Adaptive BIC-MIMO-OFDM Systems

Fu-Hsuan Chiu<sup>†</sup>, Sau-Hsuan Wu<sup>‡</sup> and C.-C. Jay Kuo<sup>†</sup>

<sup>†</sup> University of Southern California, Los Angeles, CA 90089-2564, USA

<sup>‡</sup> National Chiao Tung University, Hsinchu, Taiwan

E-mails: fchiu@usc.edu, sauhuan@cm.nctu.edu.tw and cckuo@sipi.usc.edu

**Abstract**—The performance analysis and transceiver design of an adaptive bit-interleaved-coded multi-input multi-output orthogonal frequency division multiplexing (BIC-MIMO-OFDM) system without channel state information (CSI) in priori are investigated in this research. First, the noncoherent capacity of a MIMO-OFDM system with  $N$  subcarriers,  $n_t$  transmit antennas and  $n_r$  receive antennas is analyzed. It is shown that the lower bound of noncoherent capacity in the high signal-to-noise ratio (SNR) region is asymptotically equal to  $n^*(1 - n^* \frac{L}{N}) \log \rho$ , where  $n^* = \min(n_t, n_r)$ ,  $L$  is the number of channel taps and  $\rho$  is the SNR value. Then, an adaptive BIC-MIMO-OFDM system based on transmitter beamforming and variable-rate-variable-power (VRVP) QAM is proposed. The throughput of adaptive scheme can achieve the noncoherent capacity, however, the gap to the coherent capacity cannot be totally compensated due to the overhead of training. The effect of imperfect channel state information (CSI) feedback is studied. It is demonstrated by numerical simulation that the proposed BIC-MIMO-OFDM scheme has better performance than adaptive trellis coded modulation (TCM) when outdated channel estimation is used.

## I. INTRODUCTION

Space-time coding offers a bandwidth- and power-efficient solution to wireless communications [1]. It exploits spatial diversity from multi-input multi-output (MIMO) channels to combat fading effects [2]. On the other hand, orthogonal frequency division multiplexing (OFDM) has been adopted in several commercial applications due to its robustness against frequency-selective fading. Therefore, MIMO-OFDM is a promising scheme for broadband wireless communications.

Adaptive modulation using variable-rate-variable-power (VRVP) QAM was originally proposed and studied in [3]. A general framework of MIMO adaptive modulation was presented and analyzed in [4] based on transmitter beamforming and VRVP QAM. Alternative solution to this problem is the “mercury/water-filling” strategy based on the universal formula of mutual information and minimum-mean-square-errors (MMSE) [5], where the “mercury” is to compensate the gap between discrete constellations and the Gaussian input. Adaptive bit-interleaved coded modulation was proposed in [6] to provide a reliable adaptive transmission scheme when outdated channel information is used. More recently, bit-interleaved coded multiple beamforming (BICMB) [7] has been proposed and shown to achieve the full spatial multiplexing gain while maintaining full diversity. Pioneering work on applying adaptive modulation to MIMO-OFDM systems

can be found in [8], [9] and the references therein.

In this work, we consider an  $n_t \times n_r$  MIMO-OFDM system with  $N$  subcarriers. Following the techniques developed in [10], we first derive theoretical capacity provided by an MIMO-OFDM system without channel state information (CSI) at the receiver end. We show that, in the high SNR region, the lower bound of noncoherent capacity is asymptotically equal to  $n^*(1 - n^* \frac{L}{N}) \log \rho$ , where  $n^*$  is  $\min(n_t, n_r)$ ,  $L$  is the number of channel taps, and  $\rho$  is the signal-to-noise ratio (SNR). A similar result was also reported in [11], where the behavior of noncoherent capacity of a wideband MIMO-OFDM was analyzed as a function of the total bandwidth. Then, we propose an adaptive bit-interleaved coded modulation scheme to achieve the noncoherent capacity while still maintain consistently reliable performance. However, the gap to the coherent capacity cannot be totally compensated due to the overhead of training.

The rest of this paper is organized as follow. A general MIMO-OFDM system over time-varying frequency-selective fading channels is discussed in Sec. II. Theoretically achievable capacity provided by a MIMO-OFDM system without the knowledge of channel state information (CSI) is analyzed in Sec. III. The transceiver of a BIC-MIMO-OFDM system using transmitter beamforming is presented in Sec. IV. Three adaptive squared QAM schemes are considered in Sec. V. The imperfect CSI effect is discussed in Sec. VI. The performance of the proposed scheme is evaluated and contrasted with adaptive TCM in Sec. VII. Concluding remarks are given in Sec. VIII.

## II. SYSTEM MODEL

We consider a MIMO-OFDM system with  $n_t$  transmit antennas and  $n_r$  receive antennas. Let  $N$  be the total number of subcarriers and  $\mathbf{s}_i(m) = [s_{i,0}(m), \dots, s_{i,N-1}(m)]^T$  the  $m$ th block of frequency-domain symbols sent by the  $i$ th antenna. The corresponding time-domain vector is given by  $\mathbf{x}_i(m) = \mathbf{W}^H \mathbf{s}_i(m)$ , where  $\mathbf{W}$  is the  $N$ -point discrete Fourier transform (DFT) matrix.

Let the discrete-time channel response from the  $i$ th antenna to the  $j$ th antenna be  $\mathbf{h}_{i,j}(m) = [h_{i,j,0}(m), \dots, h_{i,j,L-1}(m)]^T$  where  $L$  is the maximum channel length. The covariance matrix of channel taps is  $\Sigma_{\mathbf{h}_{i,j}} = \mathbf{E}(\mathbf{h}_{i,j} \mathbf{h}_{i,j}^H)$ . The Rayleigh fading is assumed and  $\{\mathbf{h}_{i,j}\}_{i=1, j=1}^{n_t, n_r}$  are i.i.d zero-mean complex Gaussian

random vectors with covariance matrix  $\Sigma_{\mathbf{h}_{ij}} = \Sigma_{\mathbf{h}}$ . Under the wide-sense-stationary-uncorrelated-scattering (WSSUS) assumption,  $\Sigma_{\mathbf{h}}$  is a diagonal matrix with  $[\sigma_0^2, \sigma_1^2, \dots, \sigma_{L-1}^2]^T$  as diagonal entries with  $\sum_{l=0}^{L-1} \sigma_l^2 = 1$ . A cyclic prefix (CP) of length  $L_{CP} > L$  is appended in front of  $\mathbf{x}_i(m)$  to eliminate the interblock interference (IBI). Assuming  $N \gg L$ , the loss of spectral efficiency due to CP inserted is neglected throughout the paper. The resulting vector  $\mathbf{u}_i(m)$  with length  $L_u = N + L_{CP}$  is then transmitted over the channel.

After matched filtering and sampling and CP removal, the frequency domain signal, which is obtained by applying DFT to the received discrete time vector, can be represented as

$$\mathbf{y}_m^{(k)} = \mathbf{H}_m^{(k)} \mathbf{s}_m^{(k)} + \mathbf{n}_m^{(k)}, \quad (1)$$

where  $\mathbf{y}_m^{(k)}$  and  $\mathbf{s}_m^{(k)}$  are the  $n_r \times 1$  output vector and  $n_t \times 1$  input vector, respectively,  $\mathbf{H}_m^{(k)}$  is the  $n_r \times n_t$  channel frequency response matrix corresponding to the  $k$ th subcarrier and the  $m$ th OFDM symbol, and AWGN noise vectors  $\{\mathbf{n}_m^{(k)}\}$  are i.i.d.  $\mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$  distributed.

### III. CAPACITY ANALYSIS

In the following, we use  $I(x, y)$ ,  $\mathbf{Y}$ ,  $\mathbf{S}$  and  $\mathbf{H}$  to represent the mutual information of  $x$  and  $y$ , the received signals, transmitted signals and channel realization, respectively.

#### A. Coherent Capacity

Assume that the channel state information (CSI) is only known at the receiver and the transmitter power is constrained by  $P$ . The ergodic capacity thus be obtained by averaging all channel realizations as

$$C_{\text{coh}} = \mathbb{E}_{\mathcal{H}_w} \left( \log \left( \det \left( \mathbf{I}_{n_r} + \frac{P}{n_t N \sigma^2} \mathcal{H}_w \mathcal{H}_w^H \right) \right) \right), \quad (2)$$

where  $\mathcal{H}_w$  is of dimension  $n_r \times n_t$  and i.i.d.  $\mathcal{CN}(0,1)$  entries. It is well known that this is achievable using i.i.d. Gaussian codewords [12]. The property of the Wishart distribution can be used to derive its closed form representation. In the high SNR region, we have the following approximation

$$\lim_{\rho \rightarrow \infty} \frac{C_{\text{coh}}}{\log \rho} = \min(n_t, n_r) = n^*, \quad (3)$$

where

$$\rho = \frac{P}{n_t N \sigma^2}$$

is the transmit SNR. This spatial multiplexing gain is identical to the capacity of using the water-filling technique when CSI is known at the transmitter in the high SNR region.

#### B. Noncoherent Capacity

If the CSI is unknown at the receiver, the noncoherent capacity can be represented as

$$C_{\text{non}} = \sup_{p(\mathbf{S})} I(\mathbf{Y}, \mathbf{S}). \quad (4)$$

Using the chain rule, we can decompose the mutual information as

$$I(\mathbf{Y}, \mathbf{S}) = I(\mathbf{Y}, \mathbf{S}, \mathbf{H}) - I(\mathbf{Y}, \mathbf{H} | \mathbf{S}). \quad (5)$$

According to the property that mutual information is greater than conditional mutual information, we have  $I(\mathbf{Y}, \mathbf{S}, \mathbf{H}) > I(\mathbf{Y}, \mathbf{S} | \mathbf{H})$ . It is straightforward to show that

$$I(\mathbf{Y}, \mathbf{H} | \mathbf{S}) \leq \frac{n_t n_r}{N} \sum_{l=0}^{L-1} \log(1 + \rho N \sigma_l^2). \quad (6)$$

Substituting (6) into (5), we have the lower bound of the noncoherent capacity as

$$C_{\text{non}} \geq C_{\text{nonLB}} = C_{\text{coh}} - \frac{n_t n_r}{N} \sum_{l=0}^{L-1} \log(1 + \rho N \sigma_l^2). \quad (7)$$

In the high SNR region, the lower bound is asymptotically equal to

$$\lim_{\rho \rightarrow \infty} \frac{C_{\text{nonLB}}}{\log \rho} = n^* \left( 1 - n^* \frac{L}{N} \right), \quad (8)$$

where  $n^* = \min(n_t, n_r)$  and detailed proof is shown in Appendix. Since (8) is a second order polynomial, it can be maximized by choosing  $n^* = \frac{N}{2L}$  with the maximum lower bound equal to  $\frac{N}{4L}$ . If  $\frac{N}{2L}$  is not an integer, we can choose the nearest integer to  $\frac{N}{2L}$  to be  $n^*$  to maximize the lower bound. Fig. 1 illustrates this high SNR lower bound as a function of  $n^*$  and  $L/N$ .

Based on the above analysis, we have the following design rules for MIMO-OFDM systems.

- 1) The number of transmit antennas should not be larger than that of receive antennas.
- 2) The maximum number of transmit antennas should not be larger than  $N/2L$ .
- 3) If the channels are fixed during  $M$  OFDM blocks, (8) still holds with  $N$  replaced by  $NM$ . Thus, the noncoherent capacity is larger when the ratio of the number of unknown channel taps to the number of symbols suffering equal fading gains becomes smaller.

It is well known that Cramer-Rao bound (CRB) decays with SNR, which implies that we can obtain reliable channel estimation in the high SNR region using the MMSE estimator. The identification of all  $n^* 2L$  channel taps requires at least the same number of pilots. Therefore, in high SNR region, the above lower bound can be achieved using  $n^* L$  pilots at each transmit antenna. This overhead will cause a penalty with ratio  $\frac{n^* L}{N}$ .

### IV. TRANSCIEVER DESIGN FOR ADAPTIVE BIC MODULATION

We consider the bit-interleaved coded (BIC) modulation scheme for MIMO-OFDM systems in this section. The MIMO channel model for a particular tone  $k$  can be represented by

$$\mathbf{y}_m^{(k)} = \mathbf{H}_m^{(k)} \mathbf{s}_m^{(k)} + \mathbf{n}_m^{(k)}. \quad (9)$$

Let the singular value decomposition of  $\mathbf{H}_m^{(k)}$  be  $\mathbf{H}_m^{(k)} = \mathbf{U}_m^{(k)} \mathbf{D}_m^{(k)} \mathbf{V}_m^{(k)H}$ , where  $\mathbf{U}_m^{(k)}$  and  $\mathbf{V}_m^{(k)}$  are unitary matrices, and  $\mathbf{D}_m^{(k)}$  is a diagonal matrix with diagonal entries  $\{d_i^{(k)}(m)\}_{i=1}^{n^*}$ . Multiplying (9) by  $\mathbf{U}_m^{(k)H}$ , we get

$$\mathbf{U}_m^{(k)H} \mathbf{y}_m^{(k)} = \mathbf{y}_m^{(k)'} = \mathbf{D}_m^{(k)} \mathbf{s}_m^{(k)'} + \mathbf{n}_m^{(k)'}, \quad (10)$$

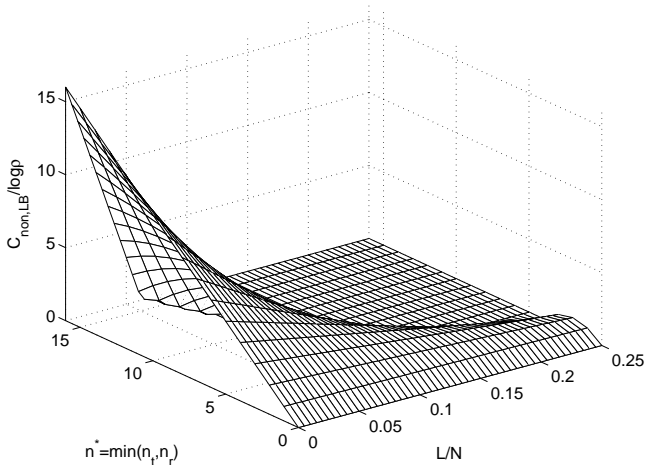


Fig. 1. The lower bound of noncoherent capacity of MIMO-OFDM in high SNR region as a function of  $n^*$  and  $L/N$ .

where  $\mathbf{s}_m^{(k)'} = \mathbf{V}_m^{(k)H} \mathbf{s}_m^{(k)}$  and  $\mathbf{n}_m^{(k)'} = \mathbf{U}_m^{(k)H} \mathbf{n}_m^{(k)}$ . Please note that  $\mathbf{s}_m^{(k)'} and  $\mathbf{n}_m^{(k)'}$  have the same distribution as  $\mathbf{s}_m^{(k)}$  and  $\mathbf{n}_m^{(k)}$ , respectively, due to the unitary property of  $\mathbf{U}_m^{(k)}$  and  $\mathbf{V}_m^{(k)H}$ . Thus, we have decomposed the MIMO channel for a particular tone into  $n^*$  parallel subchannels, and there are totally  $n^*N$  subchannels in a MIMO-OFDM systems. To ensure that the BIC-MIMO-OFDM system provides full diversity, the bit-interleaver should be carefully designed such that consecutive coded bits will be mapped to different antennas and subcarriers.$

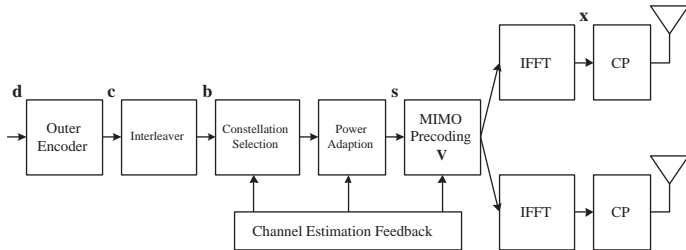


Fig. 2. The transmitter of the adaptive BIC-MIMO-OFDM system.

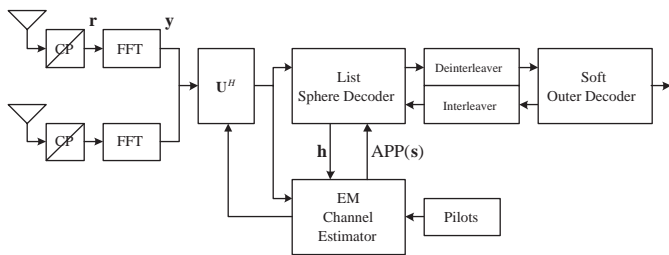


Fig. 3. The receiver of the adaptive BIC-MIMO-OFDM system.

To obtain reliable estimates of initial channel taps even

in the presence of deep fading, we select  $L$  pilot tones in each OFDM block. Since pilot tones are known at the receiver, the sequential MMSE estimator can be derived using the standard Kalman filter. The turbo receiver in the data mode proposed in [13] can be directly adopted to perform joint channel estimation and symbol detection. The predicted channel information is then fed back to the transmitter to determine the adaptive scheme. Thus, the system can allocate power and bits based on imperfect and outdated CSI.

## V. SQUARED QAM SCHEMES WITH VRVP, VR AND VP

We incorporate the adaptive variable-rate variable-power (VRVP) QAM in the BIC modulation in this section.

### A. Adaptive QAM

Since the feedback channel information is outdated with estimation errors, it may not be meaningful to optimize the performance based on such information. Instead, we propose to adjust the modulation based on uncoded QAM, and the coding gain provided by BIC aims to compensate the estimation error. The bit-error-rate (BER) of uncoded square QAM constellation of size  $2^r$ , where  $r$  is an even integer, can be approximated by [3]

$$\text{BER}_{\text{QAM}} \approx 0.2 \exp\left(\frac{-1.5\gamma}{2^r - 1}\right), \quad (11)$$

where  $\gamma$  is the received SNR per symbol. Given a target BER at  $\text{BER}_t$ , the number of bits per constellation symbol can be determined by

$$r = 2j, \quad \text{if } 2j \leq \log_2\left(1 + \frac{1.5\gamma}{\ln\left(\frac{0.2}{\text{BER}_t}\right)}\right) < 2(j+1). \quad (12)$$

To improve spectral efficiency in the low SNR region, we allow BPSK if the 4-QAM cannot meet target BER constraint. That is,

$$r = 1, \quad \text{if } (Q^{-1}(\text{BER}_t))^2 \leq \gamma < \frac{(Q^{-1}(\text{BER}_t))^2}{2}, \quad (13)$$

where  $Q^{-1}(\cdot)$  is the inverse  $Q$ -function and  $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-z^2/2} dz$ .

### B. Variable-Rate Variable-Power (VRVP) QAM

Let the power and the constellation rate allocated at the  $i$ -th subchannel of the  $k$ -th subcarrier be  $P_i^{(k)}$  and  $r_i^{(k)}$ , respectively. The optimization problem can be represented as

$$\begin{aligned} & \max \sum_{i,k} r_i^{(k)} \\ & \text{subject to } \sum_{i,k} P_i^{(k)} \leq P, \text{BER}_i^{(k)} \leq \text{BER}_t, \\ & r_i^{(k)} \geq 0, P_i^{(k)} \geq 0. \end{aligned} \quad (14)$$

Several algorithms have been proposed to solve this problem; for example, methods described in [3], [4] and [9]. However, their complexity are too high to be adopted in a real-time adaptive system. We consider two suboptimum solutions in the next two subsections.

### C. Variable-Rate (VR) QAM

In the variable-rate (VR) QAM scheme, each subchannel is allocated with the same power; namely,  $P_i^{(k)} = \frac{P}{n_t N}$ , but with a variable constellation. The constellation size allocated to each subchannel can be found as

$$r_i^{(k)} = 2j, \text{ if } 2j \leq \log_2 \left( 1 + \frac{1.5d_i^{(k)2} \frac{P}{n_t N \sigma^2}}{\ln\left(\frac{0.2}{\text{BER}_t}\right)} \right) < 2(j+1), \quad (16)$$

which is obtained using (12). The throughput per channel use of discrete rate adaption scheme with  $J$  different constellation size of square QAM can be obtained by averaging this instantaneous rate with respect to the pdf of  $\lambda = d^2$ , which is the eigenvalue of Wishart matrix [14]:

$$R_{\text{dis.VR}} = n^* \sum_{j=1}^J \int_{\lambda_{L,j}(\rho)}^{\lambda_{U,j}(\rho)} 2j p_\lambda(\lambda) d\lambda + n^* \int_{\lambda_{U,J}(\rho)}^{\infty} 2J p_\lambda(\lambda) d\lambda,$$

where  $\lambda_{L,j}(\rho) = \frac{2^{2j}-1}{\rho/K}$  and  $\lambda_{U,j}(\rho) = \frac{2^{2(j+1)}-1}{\rho/K}$ . The selection of VR constellations based on  $d_i^{(k)}$ ,  $\text{BER}_t$  and transmitter SNR ( $P/n_t N \sigma^2$ ) is summarized in Table I. An example of VR-QAM selection is given in Table II, which has the following system parameters:  $\text{BER}_t = 10^{-3}$ ,  $K = \frac{\ln(0.2/\text{BER}_t)}{1.5} = 4.893$ ,  $n_t = 4$ ,  $N = 64$ , and  $P/\sigma^2 = 44\text{dB}$

TABLE I  
ADAPTIVE VR QAM SELECTION.

Size	$d^2$ Region, $K = \frac{\ln(0.2/\text{BER}_t)}{1.5}$
0	$d^2 < \frac{(Q^{-1}(\text{BER}_t))^2}{P/n_t N \sigma^2}$
BPSK	$\frac{(Q^{-1}(\text{BER}_t))^2}{P/n_t N \sigma^2} \leq d^2 < \frac{(Q^{-1}(\text{BER}_t))^2}{P/2n_t N \sigma^2}$
4-QAM	$\frac{(Q^{-1}(\text{BER}_t))^2}{P/2n_t N \sigma^2} \leq d^2 < \frac{15K}{P/n_t N \sigma^2}$
16-QAM	$\frac{15K}{P/n_t N \sigma^2} \leq d^2 < \frac{63K}{P/n_t N \sigma^2}$
64-QAM	$\frac{63K}{P/n_t N \sigma^2} \leq d^2 < \frac{255K}{P/n_t N \sigma^2}$
256-QAM	$\frac{255K}{P/n_t N \sigma^2} \leq d^2$

TABLE II

AN EXAMPLE OF ADAPTIVE VR QAM SELECTION WITH  $\text{BER}_t = 10^{-3}$ ,  $K = 4.893$ ,  $n_t = 4$ ,  $N = 64$ , AND  $P/\sigma^2 = 44\text{dB}$ .

Size	Example
0	$d < 0.31$
BPSK	$0.31 \leq d < 0.44$
4-QAM	$0.44 \leq d < 0.86$
16-QAM	$0.86 \leq d < 1.76$
64-QAM	$1.76 \leq d < 3.53$
256-QAM	$3.53 < d$

### D. Variable-Power (VP) QAM

We employ a fixed constellation for all subcarriers in the variable-power (VP) QAM scheme and consider the power allocation problem. The optimization problem can be formulated

TABLE III

THE PSEUDO CODE FOR THE VP ADAPTIVE SCHEME.

1. Order  $\{d_i^{(k)}\}$  from the largest to the smallest  
The ordered set is written as  $\{\tilde{d}_j\}_{j=1}^{n_t N}$ ,  
where  $(i, k) \leftrightarrow j$  is a 1 to 1 mapping.
2. For  $r=1,2,4,6,8$   
 $j = 0; \tilde{P} = 0; \{\tilde{P}_{j,r}\}_{j=1}^{n_t N} = 0;$   
if  $\tilde{P} < P$   
     $\{j = j + 1;$   
    calculating  $\tilde{P}_{j,r}$  from (19);  
     $\tilde{P} = \tilde{P} + \tilde{P}_{j,r}\};$   
     $\tilde{R}_r = (j-1)r; \tilde{P}_{j,r} = 0;$
3. Return  $r^* = \arg \max_r \tilde{R}_r$  and set  $\{P_i^{(k)}\} \leftrightarrow \{\tilde{P}_{j,r^*}\}_{j=1}^{n_t N}$ .

as

$$\begin{aligned} & \max_{r_i^{(k)}} \sum_{i,k} r_i^{(k)} & (17) \\ & \text{subject to } \sum_{i,k} P_i^{(k)} \leq P, \text{BER}_i^{(k)} \leq \text{BER}_t, \\ & r_i^{(k)} = r \text{ or } 0, P_i^{(k)} \geq 0. & (18) \end{aligned}$$

It can be solved as follows. First, for each  $r$ , we calculate the required power in each subchannel using (12) and (13) from the largest  $d_i^{(k)}$  to the smallest one. Then, we have

$$P_i^{(k)} = \begin{cases} Q^{-1}(\sqrt{\text{BER}_t})\sigma^2/d_i^{(k)2}, & r=1; \\ 2Q^{-1}(\sqrt{\text{BER}_t})\sigma^2/d_i^{(k)2}, & r=2; \\ (2^r - 1)K\sigma^2/d_i^{(k)2}, & r=4,6,8. \end{cases} \quad (19)$$

Next, add  $P_i^{(k)}$  from the smallest one up to their summation just below the power constraint,  $\sum_{i,k} P_i^{(k)} < P$ . Set other  $P_i^{(k)} = 0$  and their corresponding  $r_i^{(k)} = 0$ . Finally, choose the constellation with  $2^r$  points that maximizes  $\sum_{i,k} r_i^{(k)}$ . Table III presents the pseudo code of the VP adaptive scheme.

It is easy to see that VRVP has the highest computational complexity, next VP and finally, VR. Note that, in the high SNR region, the capacity limit with CSI at the transmitter (waterfilling) is equal to that without CSI at the transmitter. Thus, VR has the same asymptotic performance as VRVP yet with a much lower complexity.

### VI. ANALYSIS OF IMPERFECT CSI EFFECT

It is assumed that the true channel and the estimated channel with imperfect CSI is related via

$$\mathbf{H}_m^{(k)} = \hat{\mathbf{H}}_m^{(k)} + \mathbf{R}_m^{(k)}, \quad (20)$$

where  $\mathbf{H}_m^{(k)}$ ,  $\hat{\mathbf{H}}_m^{(k)}$  and  $\mathbf{R}_m^{(k)}$  are the true channel, the imperfect estimated channel and the estimation error, respectively. We adopt the partial CSI model derived in [9]. That is,  $\mathbf{R}_m^{(k)}$  is i.i.d. and  $\mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I})$  distributed, where  $\sigma_e^2 = (1 - \beta^2)n_r$  and  $\beta$  is the correlation coefficient specified by the Jakes' model. Both the transmitter and the receiver use the SVD of estimated CSI,  $\hat{\mathbf{H}}_m^{(k)} = \hat{\mathbf{U}}_m^{(k)} \hat{\mathbf{D}}_m^{(k)} \hat{\mathbf{V}}_m^{(k)}$ , to perform beamforming and

recover the transmitted symbol. Then, (10) becomes

$$\begin{aligned} \mathbf{y}_m^{(k)''} &= \hat{\mathbf{U}}_m^{(k)H} (\mathbf{H}_m^{(k)} \hat{\mathbf{V}}_m^{(k)} \mathbf{x}_m^{(k)'} + \mathbf{n}_m^{(k)}) \\ &= \hat{\mathbf{D}}_m^{(k)} \mathbf{x}_m^{(k)'} + \hat{\mathbf{U}}_m^{(k)H} \mathbf{R}_m^{(k)} \hat{\mathbf{V}}_m^{(k)} \mathbf{x}_m^{(k)'} + \hat{\mathbf{n}}_m^{(k)''} \end{aligned} \quad (21)$$

where  $\mathbf{n}_m^{(k)''} = \hat{\mathbf{U}}_m^{(k)H} \mathbf{n}_m^{(k)}$  has the same distribution as  $\mathbf{n}_m^{(k)}$ . Consequently, imperfect CSI not only affects the subchannel gain but also introduces subchannel interference.

The effective SNR of the  $i$ -th subchannel can be found from (21) as

$$\begin{aligned} \gamma_{i,\text{eff}}^{(k)} &= \frac{\mathbb{E}\{|y_{i,m}^{(k)}|^2\}_{|x_{j,m}^{(k)}=0, j \neq i, \mathbf{n}_m^{(k)''}=0}}{\mathbb{E}\{|y_{i,m}^{(k)}|^2\}_{|x_{i,m}^{(k)}=0}} \\ &= \frac{P_i^{(k)} \hat{d}_i^{(k)2}}{\sigma^2} \frac{1 + \sigma_e^2 / \hat{d}_i^{(k)2}}{1 + \sum_{j=1, j \neq i}^{n_t} P_j^{(k)} \sigma_e^2 / \sigma^2}. \end{aligned} \quad (22)$$

However, since the rate adaption scheme is based on the estimated SNR,  $\gamma_{i,\text{est}}^{(k)} = \frac{P_i^{(k)} \hat{d}_i^{(k)2}}{\sigma^2}$ , we have

$$\frac{\gamma_{i,\text{est}}^{(k)}}{\gamma_{i,\text{eff}}^{(k)}} = \frac{1 + \sum_{j=1, j \neq i}^{n_t} P_j^{(k)} \sigma_e^2 / \sigma^2}{1 + \sigma_e^2 / \hat{d}_i^{(k)2}}. \quad (23)$$

When the ratio is larger than 1, the SNR is over-estimated. On the other hand, when the ratio is smaller than 1, the SNR is underestimated.

Consider the scenario that the system is operated under the VR scheme. Using (22), the effective SNR of the  $i$ -th subchannel can be found as

$$\gamma_{i,\text{eff,VR}}^{(k)} = \frac{P \hat{d}_i^{(k)2}}{n_t N \sigma^2} \left( \frac{1 + \sigma_e^2 / \hat{d}_i^{(k)2}}{1 + \frac{P(n_t-1)\sigma_e^2}{N n_t \sigma^2}} \right). \quad (24)$$

and the estimated SNR,  $\gamma_{i,\text{est,VR}}^{(k)} = \frac{P \hat{d}_i^{(k)2}}{n_t N \sigma^2}$ . We observe that

$$\frac{\gamma_{i,\text{est,VR}}^{(k)}}{\gamma_{i,\text{eff,VR}}^{(k)}} = \frac{1 + \frac{P(n_t-1)\sigma_e^2}{N n_t \sigma^2}}{1 + \sigma_e^2 / \hat{d}_i^{(k)2}}. \quad (25)$$

The ratio becomes larger when the transmitter SNR,  $\frac{P}{n_t N \sigma^2}$ , increases. In the high SNR region, the ratio will be larger than 1 for all subchannels. Consequently, the system attempts to allocate constellations of higher spectral efficiency to strong subchannels, which tend to fail to meet the target BER. Thus, increasing the transmit SNR will not solve the problem of subchannel interference totally although it provides more reliable channel estimation (with smaller  $\sigma_e^2$ ). For fixed transmit SNR, the ratio becomes larger in subchannels with larger  $\hat{d}_i^{(k)}$  while becomes smaller in subchannels with smaller  $\hat{d}_i^{(k)}$ . It means that it is more likely to overestimate the subchannel SNR value in strong channels and underestimate the subchannel SNR value in weak channels due to the imperfect CSI effect. For example, consider the case with  $\sigma_e^2 = 0.01$  and other parameters given in Table II. When  $\hat{d}_i^{(k)} = 1$ , the VR system allocates 16-QAM based on the estimated SNR,  $\hat{\gamma}_{i,\text{est}}^{(k)} = 20$ , in this subchannel. On the other hand, according to (23), the

effective SNR,  $\hat{\gamma}_{i,\text{eff}}^{(k)}$ , using noisy SVD is 14dB which can support 4-QAM only at given BER<sub>t</sub>.

Note that (25) is slightly different from the result derived using a similar approach in [4] since we extract all components corresponding to the desirable symbol while all components in the second term in (21) are treated as interference in [4].

## VII. SIMULATION RESULTS

In the simulation, we set  $N = 64$ ,  $L = 4$  and  $n_t = n_r = 4$  with a target BER at  $10^{-3}$ . The receiver does not have the knowledge of CSI so that we reserve  $n_r \times L = 16$  pilots from a total of  $N = 64$  subcarriers for the training and tracking purpose. Fig. 4 shows that the VRVP scheme performs the best among three schemes since it has two degrees of freedom in this optimization problem. The VR scheme performs the worst. However, the performance gaps narrows as the SNR value becomes higher. The VP scheme performs almost the same as the VRVP scheme in the low SNR region but the performance gap increases with SNR. It is due to the rate loss from the constant constellation constraint. By taking the implementation complexity into account, it is better to choose the VR scheme in the low SNR region and the VP scheme in the high SNR region.

It is observed that the gap between the lower bound of noncoherent capacity and coherent capacity increases as SNR becomes larger. This is due to a penalty of  $L/N$  paid as shown in (8), which reduces the slope from  $n_t = 4$  to  $n_t(1 - n_t \frac{L}{N}) = 3$ . We see that both VRVP and VP schemes can achieve noncoherent capacity at any SNR while the VR scheme can only achieve the capacity in the high SNR region. It implies that power allocation is more useful to enhance throughput than rate adaption when the power is limited. The throughput of adaptive scheme can achieve the noncoherent capacity, however, the gap to the coherent capacity cannot be totally compensated due to the overhead of training.

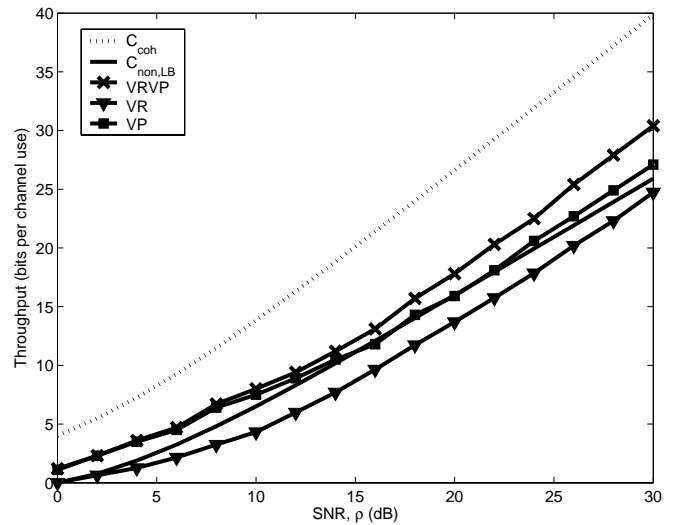


Fig. 4. Average rate for MIMO-OFDM systems of VRVP, VR and VP.

Since TCM and BICM have a similar trellis structure at the same coding rate, their decoding complexities are almost the same. Thus, we choose adaptive TCM as a benchmark for performance comparison. Fig. 5 compares the BER performance of MIMO-OFDM with adaptive TCM and BICM. The adaptive scheme is designed with respect to an uncoded  $\text{BER}_t = 10^{-3}$ . It is obvious that adaptive BIC-MIMO-OFDM performs better in the presence of channel estimation (or prediction) errors. This is because that each bit is protected in BIC-MIMO-OFDM and has the full diversity through a carefully designed bit-interleaver. In contrast, the adaptive TCM scheme demands the knowledge of uncoded bits to choose the constellation size. Also, the interleaver in TCM is operating at the symbol level and does not achieve the same time diversity as the bit interleaver.

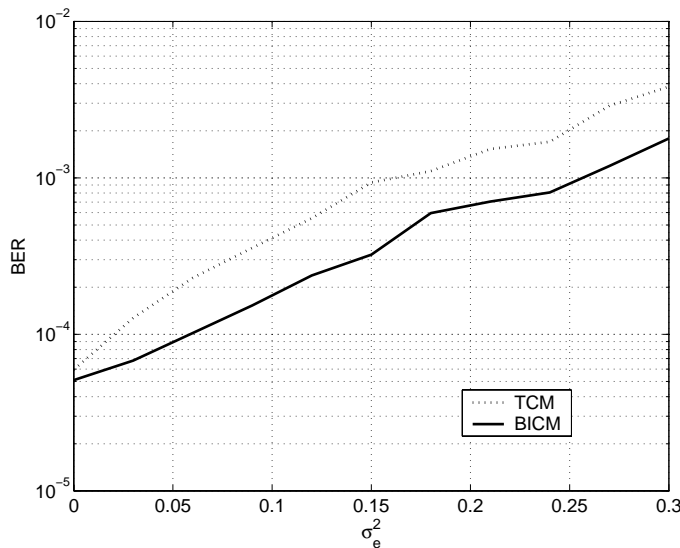


Fig. 5. Comparison of BER performance for two MIMO-OFDM systems using adaptive TCM and adaptive BICM, respectively, in the presence of imperfect CSI.

## VIII. CONCLUSION

The noncoherent capacity of MIMO-OFDM systems was first derived in this work. Then, an adaptive BIC-MIMO-OFDM scheme that exploit transmitter beamforming and variable-rate-variable-power QAM was proposed to enhance the throughput in a time-varying environment. The performance of adaptive VRVP, VR and VP schemes was analyzed and evaluated via computer simulation. The imperfect CSI effect on the effective channel SNR was explained. It was shown by simulation results that the adaptive BICM outperformed adaptive TCM when outdated channel estimation was used. Simulation results also suggested that power allocation is a better choice to enhance throughput than rate adaption when the total power is limited.

## APPENDIX

### A. Proof of (8)

In the high SNR region, an improved lower bound can be obtained by turning off the “dispensable antennas”, *i.e.*, only  $n^* \times n^*$  channels being used. Then, we have

$$\lim_{\rho \rightarrow \infty} \sum_{l=0}^{L-1} \log(1 + \rho N \sigma_l^2) = L \log \rho. \quad (26)$$

By substituting (3) and (26) into (7) and replacing  $n_t n_r$  with  $n^{*2}$ , we obtain the result in (8).

### ACKNOWLEDGEMENTS

This research has been funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152, and in part by the National Science Foundation under Grant NSF/ANI-0087761.

### REFERENCES

- [1] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment using multi-element antennas,” *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41–59.
- [2] V. Tarokh, N. Seshadri, and A. R. Calderbank, “Space-time codes for high data rate wireless communications: Performance criterion and code construction,” *IEEE Trans. on Information Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [3] A. J. Goldsmith and S.-G. Chua, “Variable-rate variable-power MQAM for fading channels,” *IEEE Trans. on Communications*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.
- [4] Z. Zhou, B. Vucetic, M. Dohler, and Y. Li, “MIMO systems with adaptive modulation,” *IEEE Trans. on Vehicular Technology*, vol. 54, no. 5, pp. 1828–1842, Sept. 2005.
- [5] A. Lozano, A. Tukino, and S. Verdu, “Mercury/Waterfilling: optimum power allocation with arbitrary input constellations,” in *Proc. IEEE ISIT*, Adelaide, Australia, Sep. 2005.
- [6] P. Ormeci, X. Liu, D. L. Goeckel, and R. D. Wesel, “Adaptive bit-interleaved coded modulation,” *IEEE Trans. on Communications*, vol. 49, no. 9, pp. 1572–1581, Nov. 2001.
- [7] E. Akay, E. Senqul, and E. Ayanoglu, “Performance of MIMO techniques to achieve full diversity and maximum spatial multiplexing,” in *ITA Inaugural Workshop*. UCSD, La Jolla, CA, Feb. 2006.
- [8] Y. Yao and G. B. Giannakis, “Rate-maximizing power allocation in OFDM based on partial channel knowledge,” *IEEE Trans. on Wireless Communications*, vol. 4, no. 3, pp. 1073–1083, May 2005.
- [9] P. Xia, S. Zhou, and G. B. Giannakis, “Adaptive MIMO-OFDM based on partial channel state information,” *IEEE Trans. on Signal Processing*, vol. 52, no. 1, pp. 202–212, Jan. 2004.
- [10] L. Zheng and D. Tse, “Communication on the Grassmann Manifold: a geometric approach to the noncoherent multiple-antenna channel,” *IEEE Trans. on Information Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
- [11] M. Borgmann and H. Bolcskei, “On the capacity of noncoherent wideband MIMO-OFDM systems,” in *Proc. IEEE ISIT*, Adelaide, Australia, Sep. 2005, pp. 651–655.
- [12] H. Bolcskei, D. Gesbert, and A. J. Paulraj, “On the capacity of OFDM-based spatial multiplexing systems,” *IEEE Trans. on Communications*, vol. 50, no. 2, pp. 225–234, Feb. 2002.
- [13] F.-H. Chiu, S.-H. Wu, and C.-C. J. Kuo, “Receiver design for bit-interleaved MIMO-OFDM systems over time-varying channels,” in *Proc. IEEE WCNC*, Las Vegas, Apr. 2006.
- [14] E. Telatar, “Capacity of multi-antenna Gaussian channels,” *Europe Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.