MIMO OFDM with ST Coding and Beamforming Adapted to Partial CSI

Pengfei Xia, Shengli Zhou, and Georgios B. Giannakis[1]
Dept. of Electrical and Computer Engr., University of Minnesota
200 Union St. SE, Minneapolis, MN 55455

Abstract — **We study adaptive multi-input multi-output (MIMO) space-time (ST) OFDM transmissions capable of adapting to partial channel state information (CSI) that is available at the transmitter. Our proposed transmitter includes an outer stage (adaptive modulation), and an inner stage (adaptive beamforming). Specifically, the inner stage switches between a one-dimensional (1D) and a balanced two-dimensional (B2D) coder-beamformer, whereas the outer stage carries out joint power and bit loading across OFDM subcarriers. The adaptive transmitter based on 1D/B2D beamformer switching performs close to its counterpart based on a general 2D beamformer structure, while obviating the need for numerical search when carrying out the power splitting between two basis beams.**

I. INTRODUCTION

Transmitter designs adapted to the intended Multi-Input Multi-Output (MIMO) fading channels are capable of improving both performance and rate of communication links. As data rates increase, the underlying MIMO channels exhibit strong frequency-selectivity. By transforming frequency-selective channels to an equivalent set of flat fading subchannels, orthogonal frequency division multiplexing (OFDM) has emerged as an attractive transmission modality, because it comes with low-complexity (de)modulation, and equalization.

These considerations motivate well adaptive MIMO OFDM, but the challenge is on what type of CSI can be made practically available to the transmitter in a wireless setting, where fading channels are randomly varying. Both single-input single-output (SISO) as well as MIMO Discrete Multi Tone (DMT) systems [4, 11] assume *perfect CSI* at the transmitter. Albeit reasonable for wireline links, perfect CSI based adaptive transmissions developed for wireless systems, can be justified only when the fading is sufficiently slow. On the other hand, the proliferation of space-time coding research testifies to the efforts towards the other extreme: non-adaptive designs requiring *no CSI* at the transmitter.

As no-CSI leads to robust but rather conservative designs, and perfect CSI is probably a utopia for most wireless links, recent efforts geared towards quantification and exploitation of *partial (or statistical) CSI* [7, 10, 13, 17] promise to have great practical value, because they are capable of offering the "jack of both trades." For example, outdated CSI (caused e.g., by feedback delay), uncertain CSI (induced e.g., by channel estimation or prediction errors), and limited CSI (appearing e.g., with quantized feedback), all can be accounted for statistically under partial CSI, but are ignored when perfect CSI is assumed. Using terms such as mean-, or, covariance-feedback to specify the type of partial CSI, existing designs have focused on MIMO transmissions over flat fading channels, where transmitter adaptation is based on a capacity-oriented [10, 13], or, a performance-oriented criterion [7, 17].

In this paper, we develop an adaptive MIMO OFDM scheme for frequency-selective fading channels, based on partial CSI at the transmitter. Our adaptive transmitter includes an outer stage (adaptive modulation), and an inner stage (adaptive beamforming). More specifically, the inner stage switches between a one-dimensional (1D) and a balanced two-dimensional (B2D) coder-beamformer, whereas the outer stage carries out joint power and bit loading across OFDM subcarriers. The adaptive transmitter based on 1D/B2D beamforming performs close to its counterpart based on a general 2D beamformer structure, while obviating the need for numerical search when carrying out the power splitting between two basis beams.

Throughout the paper, we adopt the following notational conventions: bold upper and lower case letters denote matrices and column vectors, respectively; $(\cdot)^T$ and $(\cdot)^{\mathcal{H}}$ denote transpose and Hermitian transpose, respectively; $\|\cdot\|_F$ stands for the Frobenius norm; $\mathrm{E}[\cdot]$ denotes the ensemble average; and $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a complex Gaussian distribution with mean $\boldsymbol{\mu}$, and covariance matrix $\boldsymbol{\Sigma}$.

II. SYSTEM MODEL AND PROBLEM STATEMENT

We deal with an OFDM system equipped with K subcarriers, N_t transmit-, and N_r receive-antennas, signaling over a MIMO frequency selective fading channel. Per OFDM subcarrier, we adopt a two-dimensional (2D) coder-beamformer structure; see [16, 17] for detailed derivation of the adaptive 2D coder-beamformer, which combines Alamouti's space time block coding (STBC) with transmit beamforming. Fig. 1 depicts the equivalent discrete-time baseband model of the system under consideration. To begin with, we pair two consecutive OFDM symbols, to form one ST coded OFDM block, and the transmitter adaptation is carried out on a per block basis. We reserve k to index the subcarriers; i.e., $k \in \{0, 1, \ldots, K-1\}$. If $P[k]$ stands for the power allocated to the kth subcarrier, then depending on $P[k]$, we will select a constellation $\mathcal{A}[k]$, consisting of $M[k]$ constellation points. We will focus on square and rectangular Quadrature Amplitude Modulations (QAM) that can be implemented with two independent Pulse Amplitude Modulations (PAM) on the I-Q branches [16]. For each block time-slot, the input to the 2D coder-beamformer on every subcarrier entails two information symbols, $s_1[k]$ and $s_2[k]$, drawn from $\mathcal{A}[k]$, with each one conveying $b[k] = \log_2(M[k])$ bits of information. These two information symbols will be ST coded, power-loaded, and multiplexed by the 2D coder-beamformer to generate an $N_t \times 2$

[1]This work was supported by the NSF Grant No. 01-0516, and by the ARL/CTA Grant no. DAAD19-01-2-011.

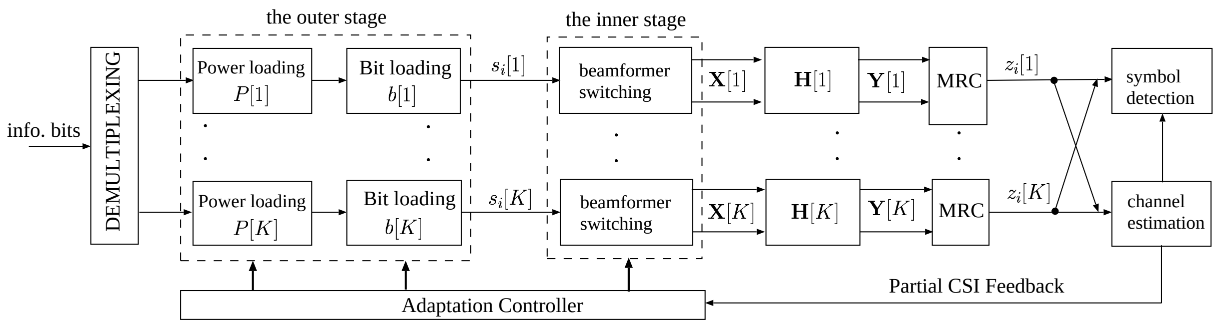

Figure 1: Discrete time baseband equivalent system model

ST matrix as:

$$\mathbf{X}[k] = \underbrace{\begin{bmatrix}\mathbf{u}_1^*[k], \mathbf{u}_2^*[k]\end{bmatrix}}_{:=\mathbf{U}^*[k]} \underbrace{\begin{bmatrix}\sqrt{\delta_1[k]} & 0 \\ 0 & \sqrt{\delta_2[k]}\end{bmatrix}}_{:=\mathbf{D}[k]} \underbrace{\begin{bmatrix}s_1[k] & -s_2^*[k] \\ s_2[k] & s_1^*[k]\end{bmatrix}}_{:=\mathbf{S}[k]}, \qquad (1)$$

where $\mathbf{S}[k]$ is the well-known Alamouti ST code matrix [1]; $\mathbf{U}[k]$ is the multiplexing matrix formed by two $N_t \times 1$ basis-beam vectors $\mathbf{u}_1[k]$ and $\mathbf{u}_2[k]$; and $\mathbf{D}[k]$ is the corresponding power splitting matrix on these two basis-beams with $0 \le \delta_1[k], \delta_2[k] \le 1$, and $\delta_1[k] + \delta_2[k] = 1$. In the two time slots corresponding to the two OFDM symbols, the two columns of $\mathbf{X}[k]$ are transmitted on the kth subcarrier over N_t transmit-antennas.

We suppose that the MIMO channel is invariant during each ST coded block, but is allowed to vary from block to block. Let $\mathbf{h}_{\mu\nu} := [h_{\mu,\nu}[0], \ldots, h_{\mu,\nu}[L]]^T$ be the baseband equivalent FIR channel between the μth transmit- and the νth receive-antenna for the current block, where $1 \le \mu \le N_t$, $1 \le \nu \le N_r$, and L is the maximum channel order of all $N_t N_r$ channels. With $\mathbf{f}_k := [1, e^{j2\pi k/N}, \ldots, e^{j2\pi kL/N}]^T$, the frequency response of $\mathbf{h}_{\mu\nu}$ on the kth subcarrier is:

$$H_{\mu,\nu}[k] = \sum_{l=0}^{L} h_{\mu\nu}[l] e^{-j2\pi kl/N} = \mathbf{f}_k^{\mathcal{H}} \mathbf{h}_{\mu\nu}.$$

Let $\mathbf{H}[k]$ be the $N_t \times N_r$ matrix having $H_{\mu\nu}[k]$ as its (μ, ν)th entry. To isolate the transmitter design from channel estimation issues at the receiver, we suppose that:

AS0) *the receiver has perfect knowledge of* $\mathbf{H}[k], \forall k$.

With $\mathbf{Y}[k]$ denoting the received block on the kth subcarrier, we have the input-output relationship per subcarrier as

$$\mathbf{Y}[k] = \mathbf{H}^T[k]\mathbf{U}^*[k]\mathbf{D}[k]\mathbf{S}[k] + \mathbf{W}[k], \qquad (2)$$

where $\mathbf{W}[k]$ stands for the additive white Gaussian noise (AWGN) at the receiver with each entry having variance $N_0/2$ per real and imaginary dimension. Based on (2), one can view our coded-beamformed MIMO OFDM transmissions per subcarrier as an Alamouti transmission with ST matrix $\mathbf{S}[k]$ passing through an equivalent channel matrix $\mathbf{B}^T[k] := \mathbf{H}^T[k]\mathbf{U}^*[k]\mathbf{D}[k]$. With knowledge of this equivalent channel and maximum ratio combining at the receiver, it is not difficult to verify that each information symbol is thus passing through an equivalent scalar channel with I/O relationship [1]

$$z_i[k] = h_{\text{eqv}}[k] s_i[k] + w_i[k], \quad i = 1, 2, \qquad (3)$$

where in our case the equivalent channel is:

$$h_{\text{eqv}}[k] = \left[\delta_1[k] \|\mathbf{H}^{\mathcal{H}}[k]\mathbf{u}_1[k]\|_F^2 + \delta_2[k] \|\mathbf{H}^{\mathcal{H}}[k]\mathbf{u}_2[k]\|_F^2\right]^{\frac{1}{2}}. \qquad (4)$$

Having introduced the system model, we next specify the partial CSI used at the transmitter. Specifically, we adopt the following model:

AS1) *On each subcarrier k, the transmitter obtains an unbiased channel estimate $\overline{\mathbf{H}}[k]$ either through a feedback channel, or, by predicting the channel from past blocks. The transmitter treats this "nominal channel" $\overline{\mathbf{H}}[k]$ as deterministic, and in order to account for CSI uncertainty, it adds a "perturbation" term. The partial CSI of the true $N_t \times N_r$ MIMO channel $\mathbf{H}[k]$ at the transmitter is thus perceived as:*

$$\check{\mathbf{H}}[k] = \overline{\mathbf{H}}[k] + \mathbf{\Xi}[k], \quad k = 0, 1, \ldots, K-1, \qquad (5)$$

where $\mathbf{\Xi}[k]$ is a random matrix Gaussian distributed according to $\mathcal{CN}\left(\mathbf{0}_{N_t \times N_r}, N_r \sigma_\epsilon^2[k] \mathbf{I}_{N_t}\right)$. The variance $\sigma_\epsilon^2[k]$ encapsulates the CSI reliability on the kth subcarrier.

The "nominal-plus-perturbation" or "mean feedback" model of AS1) has been documented for flat-fading channels [7,10,13,17]. For MIMO frequency selective fading channels, the validity of AS1) has been justified in [15] in the delayed feedback scenario.

Notwithstanding, the *partial* CSI described by AS1) has also unifying value. When $K = 1$, it boils down to the partial CSI for flat fading channels [7, 10, 13, 17]. With $\sigma_\epsilon^2 = 0$, it reduces to the *perfect* CSI of the MIMO setup considered in [11]. When $N_t = N_r = 1$, it simplifies to the partial CSI feedback used for SISO FIR channels [12, 14]. Furthermore, with $N_t = N_r = 1$ and $\sigma_\epsilon^2 = 0$, it is analogous to perfect CSI feedback for wireline DMT channels [2, 4].

A. Formulation of a Constrained Optimization Problem

Our objective is to optimize the MIMO-OFDM transmissions, based on partial CSI available at the transmitter. Specifically, we want to maximize the transmission rate subject to a power constraint, while maintaining a target BER performance on each subcarrier. Let $\overline{\mathrm{BER}}[k]$ and $\overline{\mathrm{BER}}_0[k]$ denote the average BER and the prescribed target BER on the kth subcarrier respectively. The target BERs can be identical, or, different across subcarriers, depending on system specifications. Recall that each space-time coded block conveys two symbols, $s_1[k], s_2[k]$, and thus $2b[k]$ bits of information on the kth subcarrier. Our goal can thus be formulated as the following constrained optimization problem:

$$\begin{aligned} \max \quad & 2\textstyle\sum_{k=0}^{K-1} b[k] \\ \text{s.t.} \quad & C1.\ \ \overline{\mathrm{BER}}[k] = \overline{\mathrm{BER}}_0[k], \quad \forall k \\ & C2.\ \ \textstyle\sum_{k=0}^{K-1} P[k] = P_{\text{total}} \ \text{ and } \ P[k] \geq 0, \ \ \forall k \\ & C3.\ \ b[k] = \{0, 2, 4, 6, \ldots\}, \ \ \forall k \end{aligned} \tag{6}$$

where P_{total} is the total power available per block. Compared with the constant-power transmissions over flat-fading MIMO channels [16], the problem here is more challenging, due to the needed power loading across OFDM subcarriers, which in turn depends on the 2D beamformer optimization per subcarrier.

III. ADAPTIVE MIMO-OFDM

Our transmitter includes an inner stage (adaptive beamforming), and an outer stage (adaptive modulation). Instrumental to both stages is a threshold metric, $d_0^2[k]$, which determines *allowable combinations* of $(P[k], b[k])$, so that the prescribed $\overline{\mathrm{BER}}_0[k]$ is guaranteed.

A. Adaptive Beamforming

Let T_s be the OFDM symbol duration with the cyclic prefix removed, and without loss of generality, we set $T_s = 1$. With this normalization, the constellation chosen for the kth subcarrier has average energy $\mathcal{E}_s[k] = P[k]T_s = P[k]$, and contains $M[k] = 2^{b[k]}$ signaling points. We find it convenient to work with the scaled distance metric $d^2[k] := d_{\min}^2[k]/4$, where $d_{\min}^2[k]$ denotes the minimum square Euclidean distance for this constellation. It is verified in [16] that for QAM constellations:

$$d_{\min}^2[k] = 4d^2[k] = 4\, g\big(b[k]\big)\, \mathcal{E}_s[k] = 4\, g\big(b[k]\big)\, P[k], \tag{7}$$

where the constellation-specific constant $g(b)$ is defined as:

$$g(b) := \begin{cases} \frac{6}{5\cdot 2^b - 4}, & b = 1, 3, 5, \ldots \\ \frac{6}{4\cdot 2^b - 4}, & b = 2, 4, 6, \ldots . \end{cases} \tag{8}$$

Notice that $d^2[k]$ summarizes the power and bit loading information, $\{P[k], b[k]\}$, that the adaptive modulator passes on to the adaptive coder-beamformer. The latter relies on $d^2[k]$ and the partial CSI to adapt its design so as to meet the performance constraint C1. To proceed with the adaptive beamformer design, we therefore need to analyze the BER performance per subcarrier, with input $s_i[k]$ and output $z_i[k]$, as described by (3). For each (deterministic) realization of $h_{\text{eqv}}[k]$, the BER when detecting $s_i[k]$ in the presence of AWGN in (3), can be approximated as

$$\mathrm{BER}[k] \approx 0.2 \exp(-h_{\text{eqv}}^2[k] d^2[k]/N_0), \tag{9}$$

where the validity of the approximation is confirmed in [16]. Based on the partial CSI model in AS1), the *average* BER performance on the kth subcarrier is evaluated as:

$$\overline{\mathrm{BER}}[k] \approx 0.2\, \mathrm{E}\big[\exp(-h_{\text{eqv}}^2[k] d^2[k]/N_0)\big]. \tag{10}$$

To minimize $\overline{\mathrm{BER}}[k]$ for a given $d^2[k]$, the basis beams are to be adapted based on partial CSI. To this end, we consider the eigen decomposition of the "nominal channel":

$$\begin{aligned} \overline{\mathbf{H}}[k]\overline{\mathbf{H}}^{\mathcal{H}}[k] &= \overline{\mathbf{U}}_{\mathrm{H}}[k]\boldsymbol{\Lambda}_{\mathrm{H}}[k]\overline{\mathbf{U}}_{\mathrm{H}}^{\mathcal{H}}[k], \ \text{ with} \\ \overline{\mathbf{U}}_{\mathrm{H}}[k] &:= \Big[\overline{\mathbf{u}}_{\mathrm{H},1}[k], \ldots, \overline{\mathbf{u}}_{\mathrm{H},N_t}[k]\Big], \\ \boldsymbol{\Lambda}_{\mathrm{H}}[k] &:= \mathrm{diag}\Big(\lambda_1[k], \ldots, \lambda_{N_t}[k]\Big), \end{aligned} \tag{11}$$

where $\overline{\mathbf{U}}_{\mathrm{H}}[k]$ is unitary, and $\boldsymbol{\Lambda}_{\mathrm{H}}[k]$ contains on its diagonal the eigenvalues in a non-increasing order: $\lambda_1[k] \geq \ldots \geq \lambda_{N_t}[k] \geq 0$. As proved in [16, 17], the optimal $\mathbf{u}_1[k]$ and $\mathbf{u}_2[k]$ minimizing the $\overline{\mathrm{BER}}[k]$ in (10) are:

$$\mathbf{u}_1[k] = \overline{\mathbf{u}}_{\mathrm{H},1}[k], \quad \mathbf{u}_2[k] = \overline{\mathbf{u}}_{\mathrm{H},2}[k]. \tag{12}$$

Notice that the columns of $\overline{\mathbf{U}}_{\mathrm{H}}[k]$ are also the eigenvectors of the channel correlation matrix $\mathrm{E}\{\check{\mathbf{H}}[k]\check{\mathbf{H}}^{\mathcal{H}}[k]\} = \overline{\mathbf{H}}[k]\overline{\mathbf{H}}^{\mathcal{H}}[k] + N_r\sigma_\epsilon^2[k]\mathbf{I}_{N_t}$ [17]. Hence, the basis beams $\mathbf{u}_1[k]$ and $\mathbf{u}_2[k]$ should adapt to the two eigenvectors of the perceived channel correlation matrix, corresponding to the two largest eigenvalues.

With the optimal basis beams $\mathbf{u}_1[k], \mathbf{u}_2[k]$ as in (12), we evaluate (10) to obtain the average BER as:

$$\begin{aligned} \overline{\mathrm{BER}}[k] \approx 0.2 \prod_{\mu=1}^{2} \Bigg[& \frac{1}{(1 + \delta_\mu[k] d^2[k] \sigma_\epsilon^2[k]/N_0)^{N_r}} \\ & \times \exp\left(\frac{-\lambda_\mu[k]\delta_\mu[k]d^2[k]/N_0}{1 + \delta_\mu[k]d^2[k]\sigma_\epsilon^2[k]/N_0}\right)\Bigg]. \end{aligned} \tag{13}$$

Next, we need to decide appropriate power splitting between the two basis beams on each subcarrier. We show in [15] that, if optimal power splitting $\{\delta_1[k], \delta_2[k]\}$ is employed on each subcarrier, the average performance $\overline{\mathrm{BER}}[k]$ depends on power and bit loading only through $d^2[k]$, as defined in (7). Furthermore, we show that $\overline{\mathrm{BER}}[k]$ is a monotonically decreasing function of $d^2[k]$. Hence, there exists a threshold $d_0^2[k]$ for which $\overline{\mathrm{BER}}[k] \leq \overline{\mathrm{BER}}_0[k]$ if and only if $d^2[k] \geq d_0^2[k]$, [15].

Generally, the threshold distance metric $d_0^2[k]$ is found by inverting the BER expression in (13) with respect to $d^2[k]$ when $\overline{\mathrm{BER}}[k] = \overline{\mathrm{BER}}_0[k]$. However, closed-form solutions are not available, and thus a one-dimensional numerical search is needed.

Notice that numerical search is required for all subcarriers and for each adaptation instant. The high computational complexity associated with the numerical search constitutes the major drawback, if the optimal 2D beamformer is to be used per subcarrier. We are therefore motivated to look for simpler beamforming designs that do not call for numerical search. In this paper, we propose to switch the beamformer depending on the following two power-splitting options:

Option 1—One-dimensional (1D) beamforming: In this case, we set $\delta_1[k]=1, \delta_2[k]=0$; the 2D beamformer hence reduces to a 1D beamformer with only $\mathbf{u}_1[k]$ activated, and (13) simplifies to

$$\overline{\mathrm{BER}}[k] \approx \frac{1}{5}\left[\frac{1}{(1+\frac{d^2[k]\sigma_\epsilon^2[k]}{N_0})}\exp\left(-\frac{\frac{\mathcal{K}_1[k]d^2[k]\sigma_\epsilon^2[k]}{N_0}}{1+\frac{d^2[k]\sigma_\epsilon^2[k]}{N_0}}\right)\right]^{N_r}, \tag{14}$$

which is nothing but the average BER performance of an N_r-branch diversity combining system with each branch undergoing Rician fading with Rician factor $\mathcal{K}_1[k]=\lambda_1[k]/N_r\sigma_\epsilon^2[k]$. Utilizing the well-known approximation of a Rician distribution by a Nakagami-m distribution, we can approximate the average BER performance in (14) as [17]:

$$\overline{\mathrm{BER}}[k] \approx \frac{1}{5}\left[1+d^2[k]\sigma_\epsilon^2[k]\frac{1+\mathcal{K}_1[k]}{m_1[k]N_0}\right]^{-m_1[k]N_r}, \tag{15}$$

where $m_1[k]=(1+\mathcal{K}_1[k])^2/(1+2\mathcal{K}_1[k])$. Therefore, when 1D beamforming is employed on the kth subcarrier, $d_0^2[k]$ can be found in closed-form as:

$$d_{0,\mathrm{1D}}^2[k]=\frac{\left(5\overline{\mathrm{BER}}_0[k]\right)^{-1/(m_1[k]N_r)}-1}{\sigma_\epsilon^2[k](1+\mathcal{K}_1[k])/(m_1[k]N_0)}. \tag{16}$$

Option 2— Balanced two-dimensional (B2D) beamforming: In this case, we set $\delta_1[k]=\delta_2[k]=1/2$; the 2D beamformer allocates equal (balanced) power to the two basis beams $\mathbf{u}_1[k]$ and $\mathbf{u}_2[k]$, and (13) simplifies to

$$\overline{\mathrm{BER}}[k] \approx \frac{1}{5}\left[\frac{1}{(1+\frac{d^2[k]\sigma_\epsilon^2[k]}{2N_0})}\exp\left(-\frac{\frac{\mathcal{K}_1[k]d^2[k]\sigma_\epsilon^2[k]}{2N_0}}{1+\frac{d^2[k]\sigma_\epsilon^2[k]}{2N_0}}\right)\right]^{2N_r}. \tag{17}$$

Similarly, we approximate the average BER in (17) as:

$$\overline{\mathrm{BER}}[k] \approx \frac{1}{5}\left[1+d^2[k]\sigma_\epsilon^2[k]\frac{1+\mathcal{K}_2[k]}{2m_2[k]N_0}\right]^{-2m_2[k]N_r}, \tag{18}$$

where $m_2[k]:=(1+\mathcal{K}_2[k])^2/(1+2\mathcal{K}_2[k])$, and $\mathcal{K}_2[k]=(\lambda_1[k]+\lambda_2[k])/(2N_r\sigma_\epsilon^2[k])$. Therefore, if B2D beamforming is employed on the kth subcarrier, $d_2^2[k]$ can be obtained in closed form as

$$d_{0,\mathrm{B2D}}^2[k]=\frac{\left(5\overline{\mathrm{BER}}_0[k]\right)^{-1/(2m_2[k]N_r)}-1}{\sigma_\epsilon^2[k](1+\mathcal{K}_2[k])/(2m_2[k]N_0)}. \tag{19}$$

Having specified the 1D and B2D options, we now need to decide on each subcarrier which option is to be chosen. Notice that the smaller the $d_0^2[k]$, the larger the signal constellation that can be adopted with the same power. Therefore, we just need to compare the two $d_0^2[k]$'s obtained in these two different options, and switch to the beamformer configuration with the smaller $d_0^2[k]$. Specifically, the final threshold distance is obtained as:

$$d_0^2[k]=\min(d_{0,\mathrm{1D}}^2[k], d_{0,\mathrm{B2D}}^2[k]). \tag{20}$$

Notice that the choice of beamforming schemes can be different across subcarriers. The 1D beamforming is chosen on the kth subcarrier if $d_{0,\mathrm{1D}}^2[k] \le d_{0,\mathrm{B2D}}^2[k]$, while the B2D beamforming is preferred otherwise.

The 1D and B2D beamforming are two extremes cases of the general 2D beamforming. When perfect CSI is available (extremely reliable CSI quality), the general 2D beamforming reduces to the 1D beamforming [17]. On the other hand, the general 2D beamforming boils down to the B2D beamforming when no CSI is available (extremely poor CSI quality). Hence, we expect that 1D beamforming will be used as the CSI quality improves, while the B2D beamforming will be activated when CSI worsens.

Apparently, the achieved transmission rates based on 1D/B2D switching is inferior to that based on the general 2D beamforming with numerical search, since the power splitting across the two eigen beams is carried out suboptimally. However, as we will demonstrate later, the rate difference is not significant.

B. Adaptive Modulation based on Partial CSI

With $d_0^2[k]$ encapsulating the allowable $(P[k], b[k])$ pairs per subcarrier, we are ready to pursue joint power and bit loading across OFDM subcarriers. It turns out that after suitable interpretations, many existing power and bit loading algorithms developed for DMT systems [3, 6, 9], can be applied to our adaptive MIMO-OFDM system based on partial CSI. We first show how the classical Hughes-Hartogs algorithm (HHA) [6] can be utilized to obtain the optimal power and bit loadings.

As the loaded bits in (6) assume finite (non-negative integer) values, a globally optimal power and bit allocation exists. The idea behind the HHA is that, at each step, it tries to find which subcarrier supports one additional bit with the least required additional power. Notice that the HHA belongs to the class of greedy algorithms [5, Chapter 16] that have found many applications such as the minimum spanning tree, and Huffman encoding.

Recalling our results in Subsection III-A, the minimum required power to maintain i bits on the kth subcarrier with threshold metric $d_0^2[k]$, is $d_0^2[k]/g(i)$. Therefore, the power cost incurred when loading the ith bit to the kth subcarrier is

$$c(k,i)=\frac{d_0^2[k]}{g(i)}-\frac{d_0^2[k]}{g(i-1)}, \quad i\ge 1, \forall k. \tag{21}$$

For $i=1$, we set $c(k,1)=d_0^2[k]/g(1)$. In the following, we use P_{rem} to record the remaining power after each bit loading step, $b_c[k]$ to store the number of bits already loaded on the kth subcarrier, and $P_c[k]$ to denote the amount of power currently loaded on the kth subcarrier.

The greedy algorithm for joint power and bit loading of our adaptive MIMO-OFDM follows these steps:

The Greedy Algorithm:

1. Initialization: Set $P_{\mathrm{rem}}=P_{\mathrm{total}}$. For each subcarrier, set $b_c[k]=P_c[k]=0$, and compute $d_0^2[k]$.

2. Choose the subcarrier that requires the least power to

load one additional bit; i.e., select

$$k_0 = \arg\min_k \; c(k, b_c[k]+1). \tag{22}$$

3. If the remaining power cannot accommodate it, i.e., if $P_{\rm rem} < c(k_0, b_c[k_0]+1)$, then exit with $P[k] = P_c[k]$, and $b[k] = b_c[k]$. Otherwise, load one bit to subcarrier k_0, and update state variables as

$$P_{\rm rem} = P_{\rm rem} - c(k_0, b_c[k_0]+1), \tag{23}$$

$$P_c[k_0] = P_c[k_0] + c(k_0, b_c[k_0]+1), \tag{24}$$

$$b_c[k_0] = b_c[k_0] + 1. \tag{25}$$

4. Loop back to step 2.

The greedy algorithm yields a "1-bit optimal" solution, since it offers the optimal strategy at each step when only a single bit is considered. In general, the 1-bit optimal solution obtained by a greedy algorithm may not be overall optimal. However, for our problem at hand, we establish in [15] that:

Proposition 1: *Given $\{d_0^2[k]\}$, the power and bit loading solution $\{P[k], b[k]\}_{k=0}^{K-1}$ obtained from the greedy algorithm is overall optimal.*

Furthermore, we prove in [15] that only marginal rate loss is incurred if only square QAMs are selected during our adaptive modulation stage.

Proposition 2: *Relative to allowing for both rectangular and square QAMs, the adaptive MIMO-OFDM with only square QAMs incurs up to one bit loss (on the average) per transmitted space-time coded block, that contains two OFDM symbols.*

Compared to the total number of bits conveyed by two OFDM symbols, the one bit loss is negligible when using only square QAM constellations. However, reducing the number of possible constellations by 50% simplifies the practical adaptive transmitter design. These considerations advocate only square QAM constellations for adaptive MIMO-OFDM modulation (this excludes also the popular BPSK choice).

The reason behind Proposition 2 is that square QAMs are more power efficient than rectangular QAM. With K subcarriers at our disposal, it is always possible to avoid usage of less efficient rectangular QAMs, and save the remaining power for other subcarriers to use power-efficient square QAMs. Interestingly, this is different from the adaptive modulation over flat fading channels considered in [16], where the transmit power is constant and considerable loss (one bit every two symbols on average) is involved, if only square QAM constellations are adopted.

The complexity of the optimal greedy algorithm is linearly dependent on the total number of bits loaded, and is considerably large in practice. Alternative low-complexity power and bit loading algorithms have been developed for DMT applications [3, 9]. Notice that [3] studies a dual problem: optimal allocation of the power and bits to minimize the total transmission power with a target number of bits. The truncated water-filling solution of [3] can be modified and used in our transmitter design. In spite of low-complexity, the algorithm in [3] is suboptimal, and may result in a considerable rate loss due to the truncation operation. The most interesting algorithm is the fast Lagrange bi-sectional search proposed in [9] that provides an optimal power and bit loading solution, while having complexity comparable to [3]. Hence, for our adaptive transmissions, we recommend [9] in practice.

Before we conclude this section, we briefly summarize the overall adaptation procedure for our adaptive MIMO-OFDM design based on partial CSI.

1. Basis beams per subcarrier $\{\mathbf{u}_1[k], \mathbf{u}_2[k]\}_{k=0}^{K-1}$ are adapted first using (12).
2. Switching between 1D and B2D beamforming, takes place depending on $d_0^2[k]$ that is determined from (20).
3. Power and bit loading $\{b[k], P[k]\}_{k=0}^{K-1}$ is jointly performed across all subcarriers, using either the optimal greedy algorithm, or, low-complexity one in [9].

IV. Numerical Results

We present numerical results in this section, based on the delayed-feedback paradigm in [15]. We set $K = 64$, $L = 5$, and assume that the channel taps in $\mathbf{h}_{\mu\nu}$ are i.i.d. with covariance matrix $\mathbf{\Sigma}_{\mu\nu} = (1/(L+1))\mathbf{I}_{L+1}$. Suppose that the feedback delay is τ, and the channels are slowly-varying according to the Jakes' model with Doppler frequency f_d. Then $\rho := J_0(2\pi f_d \tau)$ denotes the correlation coefficient between the channel feedback and the true channel, where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. The feedback quality can then be measured by ρ, and the variance $\sigma_\epsilon^2[k] = (1 - |\rho|^2)$ [15]. Throughout this section, the average transmit-SNR (signal to noise ratio) across subcarriers is defined as: $\text{SNR} = P_{\rm total} T_s/(K N_0)$. The transmission rate (the loaded number of bits) is counted every two OFDM symbols as: $\sum_{k=0}^{K-1} 2b[k]$.

Test case 1 — Comparison between exact and approximate solutions for $d_0^2[k]$: We simulate a typical MIMO multipath channel with $N_t = 4$, $N_r = 2$, and $N_0 = 1$. Assuming balanced 2D beamforming on each subcarrier, we plot in Fig. 2 the thresholds $d_0^2[k]$ by solving (17) numerically and by the closed-form solution (19), with $\rho = 0.7, 0.9$, and a target BER $= 10^{-4}$. The non-negative eigenvalues $\lambda_1[k]$ and $\lambda_2[k]$ of the nominal channels are also plotted in dash-dotted lines for illustration purposes. We observe that the solutions of $d_0^2[k]$ obtained via these two different approaches are generally very close to each other. And the discrepancy decreases as the feedback quality ρ increases. Notice that the suboptimal closed-form solution tends to underestimate $d_0^2[k]$. To deploy the suboptimal solution in practice, some SNR margins may be needed to ensure that the target BER performance is met. Nevertheless, we will use the suboptimal closed-form solution for $d_0^2[k]$ in our ensuing numerical results.

Fig. 2 also reveals that on subchannels with large eigenvalues (indicating "good quality"), the resulting $d_0^2[k]$ is small; hence, large size constellations can be afforded on those subchannels.

Test case 2 — Adaptive MIMO OFDM based on partial CSI: With $\overline{\text{BER}}_0 = 10^{-4}$, we compare non-adaptive transmission schemes (that use fixed constellations per OFDM

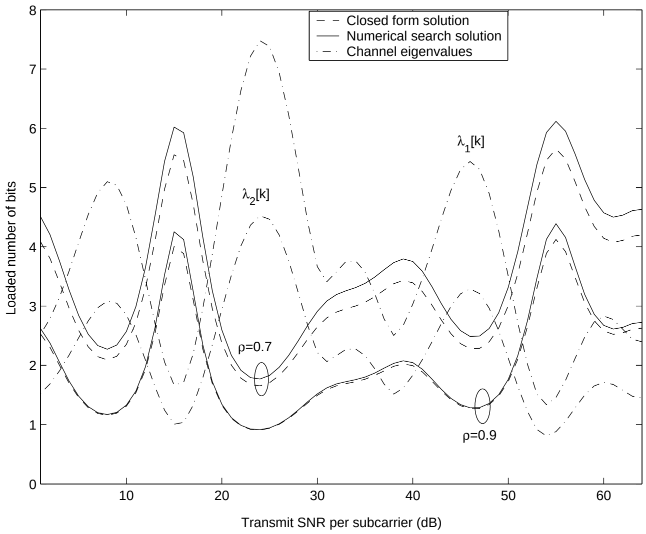

Figure 2: Threshold distances $d_0^2[k]$ for a certain channel.

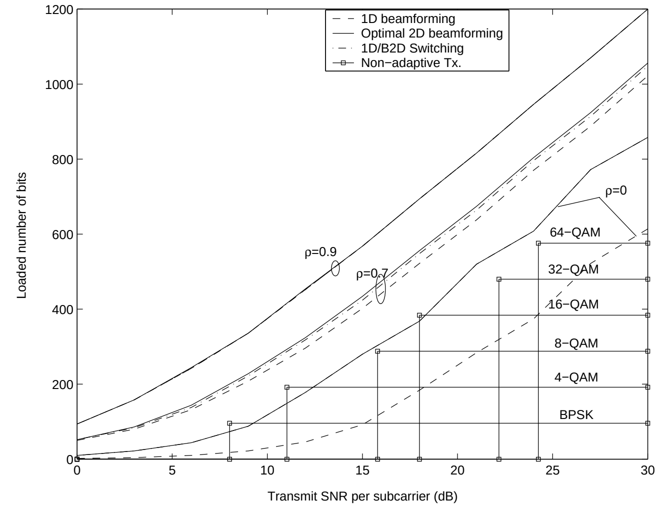

Figure 3: Rate comparisons with $N_t = 4, N_r = 2$

subcarrier), and adaptive MIMO-OFDM schemes based on 1D/B2D switching, 1D, and optimal 2D beamforming in Fig. 3 with $N_t = 4, N_r = 2$. The transmission rates for adaptive MIMO-OFDM are averaged over 200 feedback realizations.

As confirmed by Fig. 3, the non-adaptive transmitter could outperform the adaptive transmitters at the low SNR range, with extremely low feedback quality ($\rho = 0$). However, as the SNR increases, or, the feedback quality improves, the adaptive transmitters outperform their non-adaptive counterparts considerably.

Furthermore, we observe that the adaptive transmitter based on 1D/B2D beamforming achieves almost the same data rate as that based on optimal 2D beamforming, for variable partial CSI quality (they coincide at the two extremes with $\rho = 0$ and $\rho = 1$, as explained in Section III.A). On the other hand, the 1D beamforming is considerably inferior to both the optimal 2D beamforming and the 1D/B2D beamforming, when low quality CSI is present at the transmitter. But as CSI quality improves (e.g., $\rho \geq 0.9$), the rate difference diminishes. In short, the 1D/B2D beamforming is preferred in various scenarios, thanks to its reduced complexity.

REFERENCES

[1] S. M. Alamouti, "A Simple Transmit Diversity Technique for wireless Communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

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

[3] P. Chow, J. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Transactions on Communications*, vol. 43, pp. 773–775, Feb.-March-April 1995.

[4] J. M. Cioffi, "A multicarrier primer," downloadable from *http://www-isl.stanford.edu/ cioffi/pdf/multicarrier.pdf*

[5] T. Cormen, C. Leiserson, R. Rivest, and C. Stein, *Introduction to Algorithms,* (Second Edition), The MIT Press, 2001.

[6] D. Hughes-Hartogs, *Ensemble Modem Structure for Imperfect Transmission Media*, U.S. Patents Nos. 4,679,227 (July 1987), 4,731,816 (March 1988), and 4,833,706, (May 1989).

[7] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 611–627, March 2002.

[8] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framework for time-frequency processing in wireless communications," *Proceedings of the IEEE*, vol. 88, no. 5, pp. 611–640, May 2000.

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

[10] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1423–1436, October 1998.

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

[12] Z. Song, K. Zhang, and Y. Guan, "Statistical adaptive modulation for QAM-OFDM systems," in *Proceedings of IEEE Globecom Conference*, Taipei, China, November 17-21, 2002.

[13] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Transactions on Information Theory*, vol. 47, no. 6, pp. 2632–2639, September 2001.

[14] S. Ye, R. Blum, and L. Cimini, "Adaptive modulation for variable rate OFDM systems with imperfect channel information," in *Proc. of Vehicular Technology Conference*, 2002, vol. 2, pp. 767–771, Birmingham, AL, May 6-9, 2002.

[15] P. Xia and S. Zhou and G. B. Giannakis, "Adaptive MIMO OFDM based on Partial Channel State Information," *IEEE Trans. on Signal Processing*, submitted Dec. 2002

[16] S. Zhou and G. B. Giannakis, "Adaptive modulation for multi-antenna transmissions with channel mean feedback," *Proc. of Intl. Conf. on Communications,* Anchorage, Alaska, May 11-15, 2003.

[17] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2599–2613, October 2002.