Adaptive Modulation for Multiantenna Transmissions With Channel Mean Feedback

Shengli Zhou, *Member, IEEE* and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Adaptive modulation has the potential to increase the system throughput significantly by matching transmitter parameters to time-varying channel conditions. However, adaptive modulation schemes that rely on perfect channel state information (CSI) are sensitive to CSI imperfections induced by estimation errors and feedback delays. In this paper, we design adaptive modulation schemes for multiantenna transmissions based on partial CSI, that models the spatial fading channels as Gaussian random variables with nonzero mean and white covariance, conditioned on feedback information. Based on a two-dimensional beamformer, our proposed transmitter optimally adapts the basis beams, the power allocation between two beams, and the signal constellation, to maximize the transmission rate, while maintaining a target bit-error rate. Adaptive trellis-coded multiantenna modulation is also investigated. Numerical results demonstrate the rate improvement, and illustrate an interesting tradeoff that emerges between feedback quality and hardware complexity.**

***Index Terms*—Adaptive modulation, channel feedback, eigen-beamforming, multiantenna transmissions, space–time block coding, trellis-coded modulation.**

I. Introduction

BY MATCHING transmitter parameters to time varying channel conditions, adaptive modulation can increase the transmission rate considerably, which justifies its popularity for future high-rate wireless applications; see e.g., [4], [6], [8]–[12], [14], [17], [22], [25], and references therein. Crucial to adaptive modulation is the requirement of channel state information (CSI) at the transmitter, which may be obtained through a feedback channel. Adaptive designs assuming perfect CSI work well only when CSI imperfections induced by channel estimation errors and/or feedback delays are limited [4], [9]. For example, an adaptive system with delayed error-free feedback should maintain a feedback delay $\tau \leq 0.01/f_d$, where f_d denotes the Doppler frequency [4]. Such a stringent constraint is hard to ensure in practice, unless channel fading is sufficiently slow. Long range channel predictors relax this delay constraint considerably [7], [12]. An alternative approach is to account for CSI imperfections explicitly, when designing the adaptive modulator [8].

On the other hand, antenna diversity has been well established as an effective fading counter measure for wireless applications. Due to size and cost limitations, mobile units can only afford one or two antennas, which motivates multiple transmit-antennas at the base station. With either perfect or partial CSI at the transmitter, the capacity and performance of multiantenna transmissions can be further improved [16], [23]. Mean feedback is a special form of partial CSI, that is suitable for slowly time-varying channels. Through a feedback channel, the transmitter is assumed able to track the channel variations. However, to account for the uncertainty due to channel estimation errors and/or channel variations during the feedback delay, the transmitter models the spatial channels as Gaussian random variables with nonzero mean and white covariance, conditioned on the instantaneous feedback [23]. Optimal transmitters based on mean feedback have been studied using either capacity [16], [23], or, performance-oriented criteria [13], [28].

In this paper, we design *adaptive* modulation schemes for multiantenna transmissions with channel mean feedback. We base our transmitter on a two-dimensional (2-D) beamformer we derived recently in [28], where Alamouti coded [3] data streams are power loaded and transmitted along two orthogonal basis beams. Different from [28] where performance is optimized for a fixed constellation, our transmitter here optimally adjusts the basis beams, the power allocation between two beams, and the signal constellation, to maximize the system throughput while maintaining a prescribed bit-error rate (BER). We also investigate adaptive trellis coded modulation, to further increase the transmission rate. Numerical results demonstrate the rate improvement. Interestingly, adaptive multiantenna modulation turns out to be less sensitive to channel imperfections, compared to its single-antenna counterpart. To achieve the same transmission rate, an interesting tradeoff emerges between feedback quality and hardware complexity. As an example, the rate achieved by one transmit antenna when $f_d\tau \leq 0.01$ can be provided by two transmit antennas, but with a relaxed feedback delay $f_d\tau = 0.1$, representing an order of magnitude improvement.

The rest of this paper is organized as follows. Section II presents a unifying BER approximation, that comes handy for adaptive modulation. Section III specifies the system and channel models. Uncoded adaptive multiantenna modulation is designed in Section IV, and adaptive trellis coded modulation is developed in Section V. Numerical results are collected in Section VI, and conclusions are drawn in Section VII.

Manuscript received July 27, 2002; revised May 18, 2003. The editor coordinating the review of this paper and approving it for publication was J. K. Cavers. This work was supported by the National Science Foundation under Grant 0105612, and by the Army Research Laboratory under the Collaborative Technology Allliance Program, Cooperative Agreement DAAD19-01-2-011. The material in this paper was presented in part at the International Conference on Communications, Anchorage, AK, May 11–14, 2003.

S. Zhou is with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269 USA (e-mail: shengli@engr.uconn.edu).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Digital Object Identifier 10.1109/TWC.2004.833411

Notation: Bold uppercase (lowercase) letters denote matrixes (column vectors); $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote conjugate, transpose, and Hermitian transpose, respectively; $\mathrm{E}\{\cdot\}$ stands for expectation, $\mathbf{I}_K$ denotes the identity matrix of size K; $\mathbf{0}_{K\times P}$ denotes an all-zero matrix with size $K \times P$. The special notation $\mathbf{h} \sim \mathcal{CN}(\bar{\mathbf{h}}, \mathbf{\Sigma}_h)$ indicates that $\mathbf{h}$ is complex Gaussian distributed with mean $\bar{\mathbf{h}}$, and covariance matrix $\mathbf{\Sigma}_h$.

II. Unifying BER Approximation

Our goal in this section is to present a unifying approximation to BER for M-ary quadrature amplitude modulation (M-QAM). Gray mapping from bits to symbols is assumed. Notice that closed-form BER expressions are available in e.g., [27]. However, in order to facilitate adaptive modulation, approximate BERs, that are very simple to compute, are particularly attractive; see also [9]. In addition to square QAMs with $M = 2^{2i}$, we will also consider rectangular QAMs with $M = 2^{2i+1}$. We will focus on those rectangular QAMs that can be implemented with two independent pulse-amplitude-modulations (PAMs): one on the in-phase branch with size $\sqrt{2M}$, and the other on the quadrature-phase branch with size $\sqrt{M/2}$, as those studied in [24], [27].

In this section, we consider a nonfading channel with additive white Gaussian noise (AWGN), having variance $N_0/2$ per real and imaginary dimension. For a constellation with average energy E_s, let $d_0 := \min(|s - s'|)$ be its minimum Euclidean distance. For each constellation, we define a constant g as

$$g = \frac{3}{2(M-1)}, \quad \text{for square } M\text{−QAM} \tag{1}$$

$$g = \frac{6}{5M-4}, \quad \text{for rectangular } M\text{−QAM.} \tag{2}$$

The symbol energy E_s is then related to d_0^2 through the identity

$$d_0^2 = 4gE_s. \tag{3}$$

We adopt the following unifying BER approximation for all QAM constellations:

$$\mathrm{P_b} \approx 0.2\exp\left(-\frac{d_0^2}{4N_0}\right) \tag{4}$$

which can be re-expressed as

$$\mathrm{P_b} \approx 0.2\exp\left(-\frac{gE_s}{N_0}\right). \tag{5}$$

The BER approximation in (5) with g in (1) for square M-QAM was first proposed in [9]. We here extend the result in [9] to rectangular QAMs. BPSK is a special case of rectangular QAM with $M = 2$, corresponding to $g = 1$ in (5). Hence, no special treatment is needed for BPSK, as opposed to [4], [9]. We next verify the approximate BER in (5).

Example 1 (BER Approximation): In Fig. 1, we compare the exact BERs evaluated using [24] against the approximate BERs of (5) for QAM constellations with $M = 2^i, i \in [1, 8]$. The approximation is within 2 dB, for all constellations at $\mathrm{P_b} \leq 10^{-2}$, as confirmed by Fig. 1. ■

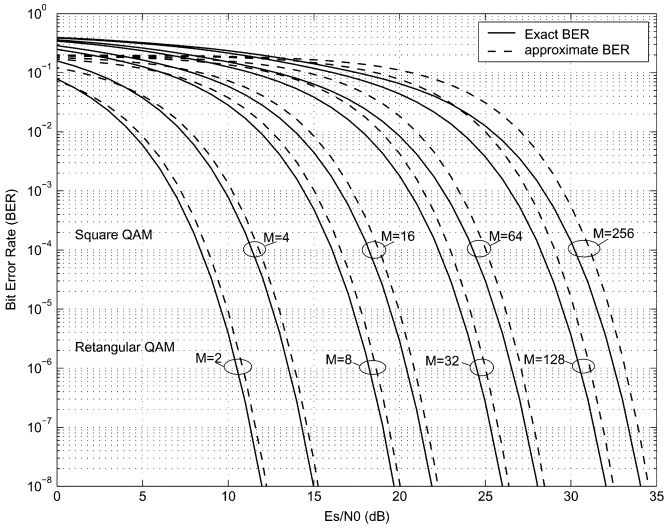

Fig. 1. BER approximation for QAM constellations.

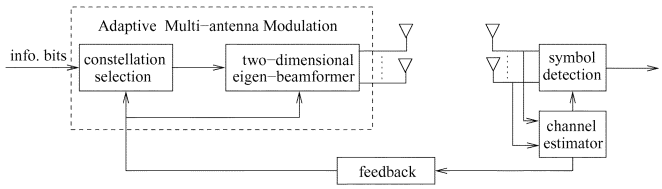

Fig. 2. System diagram.

III. System Description

With reference to Fig. 2, we study a wireless communication system with N_t transmit- and N_r receive-antennas. We focus on flat-fading channels, and let $h_{\mu\nu}$ denote the channel coefficient between the μth transmit- and the νth receive-antenna, where $\mu \in [1, N_t]$ and $\nu \in [1, N_r]$. For the extension to frequency selective multipath channels, we refer the readers to, e.g., [26]. We collect channel coefficients in an $N_t \times N_r$ channel matrix $\mathbf{H}$ having (μ, ν)th entry $h_{\mu\nu}$. For each receive antenna ν, we also define the channel vector $\mathbf{h}_\nu := [h_{1\nu}, \ldots, h_{N_t\nu}]^T$.

The wireless channels are slowly time-varying. The receiver obtains instantaneous channel estimates, and feeds them back to the transmitter regularly. Based on the available channel knowledge, the transmitter optimizes its transmission to improve the performance, and increase the overall system throughput. We next specify our channel feedback setup, and develop our adaptive multiantenna transmission structure.

A. Channel Mean Feedback

Similar to [23], we focus on channel mean feedback, where spatial fading channels are modeled as Gaussian random variables with nonzero mean and white covariance conditioned on the feedback. Specifically, we adopt the following assumption throughout.

AS0) The transmitter models the channel as

$$\mathbf{H} = \bar{\mathsf{H}} + \mathbf{\Xi} \tag{6}$$

where $\bar{\mathsf{H}}$ is the conditional mean of $\mathbf{H}$ given feedback information, and $\mathbf{\Xi} \sim \mathcal{CN}(\mathbf{0}_{N_t \times N_r}, N_r \sigma_\epsilon^2 \mathbf{I}_{N_t})$ is the associated zero-mean error matrix. The deterministic pair $(\bar{\mathsf{H}}, \sigma_\epsilon^2)$ parameterizes the partial CSI, which is updated regularly given feedback information from the receiver. ■

The partial CSI parameters $(\bar{\mathsf{H}}, \sigma_\epsilon^2)$ can be provided in many different ways; see [28] for a brief summary. For illustration purposes, we elaborate next on a specific application scenario with delayed channel feedback [8], [16], [23], that we will use in our simulations.

Example 2 (Delayed Channel Feedback): Suppose that: 1) the channel coefficients $\{h_{\mu\nu}\}_{\mu=1,\nu=1}^{N_t,N_r}$ are independent and identically distributed with Gaussian distribution $\mathcal{CN}(0, \sigma_h^2)$; 2) the channels are slowly time varying according to Jakes' model with Doppler frequency f_d; and 3) the channels are acquired perfectly at the receiver and are fed back to the transmitter with delay τ, but without errors. Perfect channel estimation at the receiver (with infinite quantization resolution), and error-free feedback, which can be approximated by using error-control coding and automatic repeat request (ARQ) protocol in the feedback channel, are commonly assumed [8], [9]. Notice that the channel feedback $\mathbf{H}_\mathrm{f}$ is drawn from the same Gaussian process as $\mathbf{H}$, but in τ seconds ahead of $\mathbf{H}$. The corresponding entries of $\mathbf{H}_\mathrm{f}$ and $\mathbf{H}$ are then jointly zero-mean Gaussian, with correlation coefficient $\rho := J_0(2\pi f_d \tau)$ specified from the Jakes' model, where $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. For each realization of $\mathbf{H}_\mathrm{f}$, the parameters needed in the mean feedback model of (6) are obtained as [8], [16], [23]

$$\bar{\mathsf{H}} = \mathrm{E}\{\mathbf{H} \,|\, \mathbf{H}_\mathrm{f}\} = \rho \mathbf{H}_\mathrm{f}, \quad \sigma_\epsilon^2 = \sigma_h^2(1 - |\rho|^2). \tag{7}$$

B. Adaptive 2-D Transmit-Beamforming

The adaptive multiantenna transmitter in this paper is based on the 2-D beamformer that we developed recently in [28] for fixed constellations. Depending on channel feedback, the information bits will be mapped to symbols drawn from a suitable constellation. The symbol stream $s(n)$ will then be fed to the 2-D beamformer, and transmitted through N_t antennas. The structure of the 2-D beamformer is depicted in Fig. 3. It uses the Alamouti code [3], to generate two data streams $\bar{s}_1(n)$ and $\bar{s}_2(n)$ from the original symbol stream $s(n)$ as follows:

$$\begin{bmatrix} \bar{s}_1(2n) & \bar{s}_1(2n+1) \\ \bar{s}_2(2n) & \bar{s}_2(2n+1) \end{bmatrix} = \begin{bmatrix} s(2n) & -s^*(2n+1) \\ s(2n+1) & s^*(2n) \end{bmatrix}. \tag{8}$$

The total transmission power E_s is allocated to these streams: $\delta_1 E_s$ to $\bar{s}_1(n)$, and $\delta_2 E_s = (1 - \delta_1) E_s$ to $\bar{s}_2(n)$, where $\delta_1 \in [0, 1]$. Each power-loaded symbol stream is weighted by an $N_t \times 1$ beam-steering vector $\mathbf{u}_j, j = 1, 2$, and transmitted simultaneously. Collecting symbols across N_t antennas, the transmitted vector $\mathbf{x}(n) := [x_1(n), \ldots, x_{N_t}(n)]^T$ at the nth time slot is

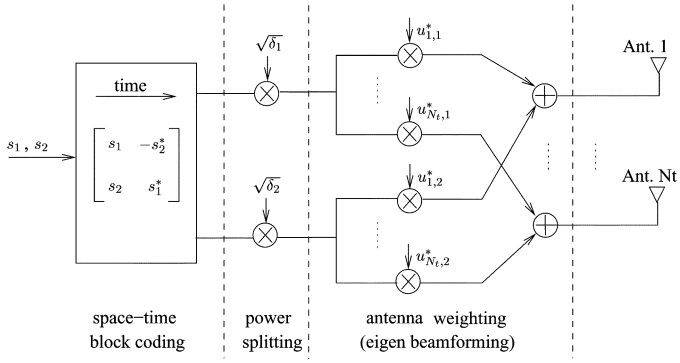

Fig. 3. The 2-D eigen-beamformer, $u_{p,q} := [\mathbf{U}_\mathrm{H}]_{p,q}$.

$$\mathbf{x}(n) = \bar{s}_1(n)\sqrt{\delta_1}\mathbf{u}_1^* + \bar{s}_2(n)\sqrt{\delta_2}\mathbf{u}_2^*. \tag{9}$$

Moving from single to multiple transmit-antennas, a number of spatial multiplexing and space time coding options are possible, at least when no CSI is available at the transmitter. We are motivated to pursue an adaptive transmitter based on our 2-D beamforming approach for the following reasons.

1) Based on channel mean feedback, the optimal transmission strategy (in the uncoded case) is to combine beamforming (with $N_t \geq 2$ beams) with orthogonal space–time block coding (STBC) [3], [20], where the optimality pertains to an upper-bound on the pairwise error probability [13], or, an upper-bound on the symbol error rate [28]. However, orthogonal STBC loses rate when $N_t > 2$, which is not appealing for adaptive modulation whose ultimate goal is to increase the data rate given a target BER performance. On the other hand, the 2-D beamformer can achieve the best possible performance when the channel feedback quality improves [28]. Furthermore, the 2-D beamformer is suboptimal only at very high SNR [28]. In such cases, the achieved BER is already below the target, rendering further effort on BER improvement by sacrificing the rate unnecessary. In a nutshell, the 2-D beamformer is preferred because of its full-rate property, and its robust performance across the practical SNR range.
2) Our 2-D beamformer structure is general enough to include existing adaptive multiantenna approaches; e.g., the special case of $(N_t, N_r) = (2, 1)$ with perfect CSI considered in [12]. To verify this, let us denote the channels as h_1 and h_2. Setting $(\delta_1, \delta_2) = (1, 0), \mathbf{u}_1 = [1, 0]^T$ when $|h_1| > |h_2|$ and $\mathbf{u}_1 = [0, 1]^T$ otherwise, our 2-D beamformer reduces to the selective transmitter diversity (STD) scheme of [12]. Setting $(\delta_1, \delta_2) = (1, 0)$ and $\mathbf{u}_1 = [h_1, h_2]^T / \sqrt{|h_1|^2 + |h_2|^2}$, our 2-D beamformer reduces to the transmit adaptive array (TxAA) scheme of [12]. Finally, setting $(\delta_1, \delta_2) = (1/2, 1/2), \mathbf{u}_1 = [1, 0]^T$ and $\mathbf{u}_2 = [0, 1]^T$ leads to the space–time transmit diversity (STTD) scheme of [12].

3) Thanks to the Alamouti structure, our optimal receiver processing is simple. The received symbol $y_\nu(n)$ on the νth antenna is

$$\begin{aligned} y_\nu(n) &= \mathbf{x}^T(n)\mathbf{h}_\nu + w_\nu(n) \\ &= \bar{s}_1(n)\sqrt{\delta_1}\mathbf{u}_1^H\mathbf{h}_\nu \\ &\quad + \bar{s}_2(n)\sqrt{\delta_2}\mathbf{u}_2^H\mathbf{h}_\nu + w_\nu(n) \end{aligned} \tag{10}$$

where $w_\nu(n)$ is the additive white noise with variance $N_0/2$ per real and imaginary dimension. Equation (10) suggests that the receiver only observes two virtual transmit antennas, transmitting $\bar{s}_1(n)$ and $\bar{s}_2(n)$, respectively. The equivalent channel coefficient from the jth virtual transmit antenna to the νth receive-antenna is $\sqrt{\delta_j}\mathbf{u}_j^H\mathbf{h}_\nu$. Supposing that the channels remain constant at least over two symbols, the linear maximum ratio combiner (MRC) in [3] is directly applicable to our receiver, ensuring maximum likelihood optimality. Symbol detection is performed separately for each symbol; and each symbol is equivalently passing through a scalar channel with

$$\begin{aligned} y(n) &= h_{\rm eqv}s(n) + w(n) \\ h_{\rm eqv} &:= \left[\delta_1\sum_{\nu=1}^{N_r}\left|\mathbf{u}_1^H\mathbf{h}_\nu\right|^2 + \delta_2\sum_{\nu=1}^{N_r}\left|\mathbf{u}_2^H\mathbf{h}_\nu\right|^2\right]^{\frac{1}{2}} \end{aligned} \tag{11}$$

where $w(n)$ has variance $N_0/2$ per dimension. The transmitter influences the quality of the equivalent scalar channel $h_{\rm eqv}$ through the 2-D beamformer adaptation of $(\delta_1, \delta_2, \mathbf{u}_1, \mathbf{u}_2)$.

4) Individually, Alamouti's coding and transmit-beamforming have been proposed into standards [1], [2]. The 2-D beamformer offers a neat combination of these two existing components.

We next specify our adaptive transmitter based on the 2-D beamformer structure.

IV. Adaptive Modulation Based on 2-D Beamforming

Based on mean feedback, the transmitter will adjust the basis beams ($\mathbf{u}_1$ and $\mathbf{u}_2$), the power allocation (δ_1 and δ_2), and the signal constellation of size M and energy E_s, to maximize the transmission rate while maintaining the target BER: $\mathrm{P_{b,\,target}}$. As in [8]–[10], we will adopt QAM constellations. Suppose we have N different QAM constellations with $M_i = 2^i$, where $i = 1, 2, \ldots, N$, as those specified in Example 1. Correspondingly, we denote the constellation-specific constant g as g_i. The value of g_i is evaluated from (1), or (2), depending on the constellation M_i. When the channel experiences deep fades, we will allow our adaptive design to suspend data transmission (this will correspond to $M_0 = 0$).

Under AS0), the transmitter perceives a random channel matrix $\mathbf{H}$ as in (6). The BER for each realization of $\mathbf{H}$ is obtained from (11) and (5) as

$$\mathrm{P_b}(\mathbf{H}, M_i) \approx 0.2\exp\left(-h_{\rm eqv}^2\frac{g_iE_s}{N_0}\right). \tag{12}$$

Since the realization of $\mathbf{H}$ is not available, the transmitter relies on the average BER

$$\bar{\mathrm{P}}_{\rm b}(M_i) = \mathrm{E}\{\mathrm{P_b}(\mathbf{H}, M_i)\} \approx 0.2\,\mathrm{E}\left\{\exp\left(-h_{\rm eqv}^2\frac{g_iE_s}{N_0}\right)\right\} \tag{13}$$

and uses $\bar{\mathrm{P}}_{\rm b}(M_i)$ as a performance metric to select a constellation of size M_i.

A. Adaptive Beamforming

Let the eigen decomposition of $\bar{\mathbf{H}}\bar{\mathbf{H}}^H$ be

$$\begin{aligned} \bar{\mathbf{H}}\bar{\mathbf{H}}^H &= \mathbf{U}_{\rm H}\mathbf{D}_{\rm H}\mathbf{U}_{\rm H}^H \\ \mathbf{D}_{\rm H} &:= \mathrm{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_{N_t}) \end{aligned} \tag{14}$$

where $\mathbf{U}_{\rm H} := [\mathbf{u}_{\rm H,1}, \ldots, \mathbf{u}_{\rm H,N_t}]$ contains N_t eigenvectors, and $\mathbf{D}_{\rm H}$ has the corresponding N_t eigenvalues on its diagonal in a nonincreasing order: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_t}$. Since $\{\mathbf{u}_{{\rm H},\mu}\}_{\mu=1}^{N_t}$ are also eigenvectors of $\bar{\mathbf{H}}\bar{\mathbf{H}}^H + N_r\sigma_\epsilon^2\mathbf{I}_{N_t}$, the correlation matrix of the perceived channel $\mathbf{H}$ in (6), we term them as eigendirections or eigenbeams [28].

For any power allocation with $\delta_1 \geq \delta_2 \geq 0$, we have established in [28, Proposition 2] and [28, eq. (63)] that the optimal $\mathbf{u}_1$ and $\mathbf{u}_2$ minimizing $\bar{\mathrm{P}}_{\rm b}(M_i)$ in (13) are

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{u}_{\rm H,1} \\ \mathbf{u}_2 &= \mathbf{u}_{\rm H,2}. \end{aligned} \tag{15}$$

In other words, the optimal basis beams for our 2-D beamformer are eigenbeams corresponding to the two largest eigenvalues λ_1 and λ_2. Hereafter, we term our adaptive 2-D beamformer as 2-D eigenbeamformer.

B. Adaptive Power Allocation Between Two Beams

With the optimal eigenbeams in (15), the average BER can be obtained similar to [28, eq. (51)], but with only two virtual antennas. Formally, the expected BER is

$$\bar{\mathrm{P}}_{\rm b}(M_i) \approx 0.2\prod_{\mu=1}^{2}\left[\frac{1}{1+\delta_\mu\beta_i}\exp\left(-\frac{\lambda_\mu\delta_\mu\beta_i}{N_r\sigma_\epsilon^2(1+\delta_\mu\beta_i)}\right)\right]^{N_r} \tag{16}$$

where for notational brevity, we define

$$\beta_i := g_i\sigma_\epsilon^2E_s/N_0. \tag{17}$$

For a given β_i, the optimal power allocation that minimizes (16) can be found in closed-form, following derivations in [28]. Specifically, with two virtual antennas, we simplify [28, eq. (53)] to

$$\delta_2 = \max\left(\delta_2^0, 0\right)$$

and

$$\delta_1 = 1 - \delta_2 \tag{18}$$

where δ_2^0 is obtained from [28, eq. (54)] as

$$\delta_2^0 := \frac{1+\frac{N_r\sigma_\epsilon^2+\lambda_1}{(N_r\sigma_\epsilon^2+2\lambda_1)\beta_i}+\frac{N_r\sigma_\epsilon^2+\lambda_2}{(N_r\sigma_\epsilon^2+2\lambda_2)\beta_i}}{1+\frac{(N_r\sigma_\epsilon^2+2\lambda_2)(N_r\sigma_\epsilon^2+\lambda_1)^2}{(N_r\sigma_\epsilon^2+2\lambda_1)(N_r\sigma_\epsilon^2+\lambda_2)^2}} - \frac{N_r\sigma_\epsilon^2+\lambda_2}{(N_r\sigma_\epsilon^2+2\lambda_2)\beta_i}.$$

The optimal solution in (18) guarantees that $\delta_1 \geq \delta_2 \geq 0$; thus, more power is allocated to the stronger eigenbeam. If two eigenbeams are equally important ($\lambda_1 = \lambda_2$), the optimal solution is $\delta_1 = \delta_2 = 1/2$. On the other hand, if the channel feedback quality improves as $\sigma_\epsilon^2 \to 0$, we have $\delta_1 = 1$, and $\delta_2 = 0$. We underscore that the optimal δ_1 and δ_2 are constellation dependent.

C. Adaptive Rate Selection With Constant Power

With perfect CSI, using the probability density function (pdf) of the channel fading amplitude, the optimal rate and power allocation for single antenna transmissions has been provided in [9]. Optimal rate and power allocation for our multiantenna transmission with imperfect CSI turns out to be complicated. We will thus focus on constant power transmission, and only adjust the modulation level, as in [4], [12], [17], [22]. Constant power transmission simplifies the transmitter design, and obviates the need for knowing the channel pdf.

With fixed transmission power and a given constellation, the transmitter computes the expected BER with optimal power splitting on two eigenbeams, per channel feedback. It then chooses the rate-maximizing constellation, while maintaining the target BER. Since the BER performance decreases monotonically with the constellation size, the transmitter finds the optimal constellation to be

$$M = \arg \max_{M\in\{M_i\}_{i=0}^N} \bar{\mathrm{P}}_\mathrm{b}(M) \leq \mathrm{P}_{\mathrm{b,target}}. \tag{20}$$

Equation (20) can be simply solved by trial and error: we start with the largest constellation $M_i = M_N$, and then decrease i until we find the optimal M_i.

Interestingly, although we have N_tN_r entries in $\bar{\mathsf{H}}$, our constellation selection depends only on the first two eigen-values λ_1 and λ_2. We can split the 2-D space of (λ_1, λ_2) into $N+1$ disjoint regions $\{D_i\}_{i=0}^N$, each associated with one constellation. Specifically, we choose

$$M = M_i, \quad \text{when } (\lambda_1, \lambda_2) \in D_i, \quad \forall i = 0, 1, \ldots, N. \tag{21}$$

The rate achieved by our system is, therefore,

$$R = \sum_{i=1}^N \log_2(M_i) \iint_{D_i} p(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2 \tag{22}$$

where $p(\lambda_1, \lambda_2)$ is the joint pdf of λ_1 and λ_2. The outage probability is thus

$$\mathrm{P}_\mathrm{out} = \iint_{D_0} p(\lambda_1, \lambda_2) d\lambda_1 d\lambda_2. \tag{23}$$

We next specify the fading regions. Since $\lambda_2 \leq \lambda_1$, we have $a := \lambda_2/\lambda_1 \in [0, 1]$. To specify the region D_i in the (λ_1, λ_2) space, we will specify the intersection of D_i with each straight line $\lambda_2 = a\lambda_1$, where $a \in [0, 1]$. Specifically, the fading region D_i on each line will reduce to an interval. We denote this interval on the line $\lambda_2 = a\lambda_1$ as $[\alpha_i(a), \alpha_{i+1}(a))$, during which the constellation M_i is chosen. Obviously, $\alpha_0(a) = 0$, and $\alpha_{N+1}(a) = \infty$. What is left to specify are the boundary points $\{\alpha_i(a)\}_{i=1}^N$.

For a given constellation M_i and power allocation factors $(\delta_1, \delta_2 = 1-\delta_1)$, we determine from (16) the minimum value of λ_1 on the line of $\lambda_2 = a\lambda_1$, so that $\bar{\mathrm{P}}_\mathrm{b}(M_i) \leq \mathrm{P}_{\mathrm{b,target}}$ as

$$\lambda_1(a, \delta_1, M_i) = \sigma_\epsilon^2 \left(\frac{\delta_1\beta_i}{1+\delta_1\beta_i} + \frac{a\delta_2\beta_i}{1+\delta_2\beta_i}\right)^{-1} \times \ln\left(\frac{0.2}{\mathrm{P}_{\mathrm{b,target}}[(1+\delta_1\beta_i)(1+\delta_2\beta_i)]^{N_r}}\right). \tag{24}$$

Since the optimal $\delta_1 \in [1/2, 1]$ will lead to the minimal λ_1 that satisfies the BER requirement, we find the boundary point $\alpha_i(a)$ as

$$\alpha_i(a) = \min_{\delta_1\in[1/2,1]} \lambda_1(a, \delta_1, M_i). \tag{25}$$

The minimization in (25) is a simple one-dimensional (1-D) search, and we carry it out numerically. Having specified the boundaries on each line, we are now able to plot the fading regions associated with each constellation in the 2-D space, as we will illustrate later.

D. Special Cases

In the general multiple input multiple output (MIMO) case, each constellation M_i is associated with a fading region D_i on the 2-D plane (λ_1, λ_2). We will discuss several special cases, where the fading region is effectively determined by fading intervals on the first eigenvalue λ_1. In such cases, we denote the boundary points as $\{\bar{\alpha}_i\}_{i=0}^{N+1}$. The constellation M_i is chosen when $\lambda_1 \in [\bar{\alpha}_i, \bar{\alpha}_{i+1})$. We then obtain

$$\begin{aligned} R &= \sum_{i=1}^N \log_2(M_i) \int_{\bar{\alpha}_i}^{\bar{\alpha}_{i+1}} p(\lambda_1) d\lambda_1 \\ &= \sum_{i=1}^N \log_2(M_i)[F(\bar{\alpha}_{i+1}) - F(\bar{\alpha}_i)] \end{aligned} \tag{26}$$

where $F(x) := \int_0^x p(\lambda_1) d\lambda_1$ is the cumulative distribution function (cdf) of λ_1. The outage in (23) becomes

$$\mathrm{P}_\mathrm{out} = F(\bar{\alpha}_1). \tag{27}$$

To calculate the rate and outage, it suffices to determine the pdf of λ_1, and the boundaries $\{\bar{\alpha}_i\}_{i=1}^N$.

1) Multiple Input and Single Output (MISO): We consider here multiple transmit- and a single receive-antennas. With $N_r = 1$, we have only one nonzero eigenvalue λ_1, and thus $a = \lambda_2/\lambda_1 = 0$. The boundary points are

$$\bar{\alpha}_i = \alpha_i(0), \quad \forall i = 0, 1, \ldots, N \tag{28}$$

where $\alpha_i(a)$ is specified in (25).

Example 3 (i.i.d. Channels): When $N_r = 1$, the channel $\mathbf{h}_1$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_t})$. With delayed feedback considered in Example 2, we have $\lambda_1 = (|\rho|^2)\|\mathbf{h}_1\|^2 = |\rho|^2 \sum_{\mu=1}^{N_t} |h_{\mu 1}|^2$,

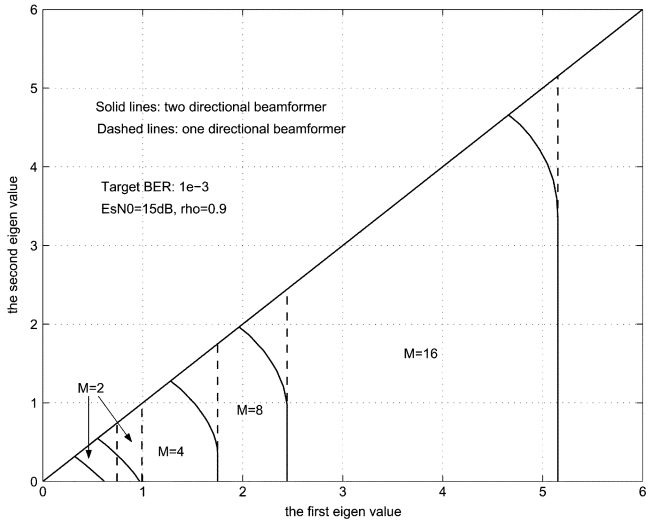

Fig. 4. Fading regions for different constellations.

which is Gamma distributed with parameter N_t and mean $\mathrm{E}\{\lambda_1\} = |\rho|^2 N_t$. The pdf and cdf of λ_1 are (see also [4])

$$p(\lambda_1) = \left(\frac{1}{|\rho|^2}\right)^{N_t} \frac{\lambda_1^{N_t-1}}{(N_t-1)!} \exp\left(-\frac{\lambda_1}{|\rho|^2}\right), \quad \lambda_1 \geq 0$$

$$F(x) = \int_0^x p(\lambda_1) d\lambda_1 \tag{29}$$

$$= 1 - e^{-x/|\rho|^2} \sum_{j=0}^{N_t-1} \frac{1}{j!} \left(\frac{x}{|\rho|^2}\right)^j, \quad x \geq 0. \tag{30}$$

Plugging (30) and (28) into (26), the rate becomes readily available. ∎

2) One-Dimensional Eigen-Beamforming: We now return to the MIMO case. Instead of using two basis beams, the conventional beamforming uses only one eigen-beam; an adaptive system based on 1-D beamforming is studied in [29]. Our 2-D beamformer subsumes the 1-D beamformer by setting $\delta_1 = 1$ and $\delta_2 = 0$. Numerical search in (25) is now unnecessary, and $\alpha_i(a)$ does not depend on a anymore. We simplify (24) to

$$\bar{\alpha}_i = \lambda_1(a, 1, M_i) = \frac{\sigma_\epsilon^2}{\beta_i}(1+\beta_i) \ln\left(\frac{0.2}{\mathrm{P}_{\mathrm{b,target}}(1+\beta_i)^{N_r}}\right). \tag{31}$$

The fading region, thus, depends only on λ_1.

Example 4 (Fading Region): The optimal regions for different signal constellations are plotted in Fig. 4 with $\mathrm{P_b} = 10^{-3}, E_s/N_0 = 15$ dB, and $\rho = 0.9$. As the constellation size increases, the difference between 1- and 2-D beamforming decreases.

3) Perfect CSI: With perfect CSI ($\sigma_\epsilon^2 = 0, \bar{\mathsf{H}} = \mathbf{H}$), the optimal loading in (18) ends up being $\delta_1 = 1, \delta_2 = 0$. Therefore, the optimal transmission strategy in this case is 1-D eigen-beamforming. Our results apply to 1-D beamforming, but with $\sigma_\epsilon^2 = 0$. Specifically, we simplify (16) to

$$\mathrm{P_b}(M_i) \approx 0.2 \exp\left(-\lambda_1 \frac{g_i E_s}{N_0}\right) \tag{32}$$

and (31) to

$$\bar{\alpha}_i = \lambda_1(a, 1, M_i) = \frac{1}{g_i E_s/N_0} \ln\left(\frac{0.2}{\mathrm{P}_{\mathrm{b,target}}}\right). \tag{33}$$

Equation (32) reveals that the MIMO antenna gain is introduced solely through λ_1, the maximum eigenvalue of $\bar{\mathsf{H}}\bar{\mathsf{H}}^H$ (or, $\mathbf{H}\mathbf{H}^H$).

Remark 1: Notice that with *perfect* CSI, one can enhance spectral efficiency by adaptively transmitting parallel data streams over as many as N_t eigenchannels of $\mathbf{H}\mathbf{H}^H$. These data streams can be decoded separately at the receiver. However, this scheme can not be applied when the available CSI is *imperfect*, since the eigendirections of $\bar{\mathsf{H}}\bar{\mathsf{H}}^H$ are no longer the eigendirections of the true channel $\mathbf{H}\mathbf{H}^H$. As a result, these parallel streams will be coupled at the receiver side, and will interfere with each other. This coupling calls for higher receiver complexity to perform joint detection, and also complicates the transmitter design, since no approximate BER expressions are readily available. Adaptive transmitter design with parallel transmissions based on imperfect CSI is an interesting topic, but beyond the scope of this paper.

V. Adaptive Trellis-Coded Modulation (TCM)

In this section, we consider coded modulation. We recall from (11) that each information symbol $s(n)$ is equivalently passing through a scalar channel in the proposed transmitter. Thus, conventional channel coding can be applied. As in [8], [10], [11], [14], we focus on TCM, where a fixed trellis code is superimposed on uncoded adaptive modulation for fading channels. Our goal in this section is to extend the single antenna design with perfect CSI [10], to our MIMO system with partial (i.e., imperfect) CSI.

The adaptive TCM diagram is plotted in [10, Fig. 2]. Out of n information bits, k bits are passing through a trellis encoder to generate $k+r$ coded bits. A constellation of size 2^{n+r} is partitioned into 2^{k+r} subsets with size 2^{n-k} each. The $k+r$ coded bits will decide which subset to be used, and the remaining $n-k$ uncoded bits will specify one signal point from the subset to be transmitted. Similar to [10], we fix the trellis code, and adapt the signal constellation according to channel conditions. Different from the uncoded case, the minimum constellation size now is 2^{k+r} with each subset containing only one point. With a constellation of size M_i, only $\log_2(M_i) - r$ bits are transmitted.

A. BER Approximation for AWGN Channels

Let d_{free} denote the minimum Euclidean distance between any pair of valid codewords. At high SNR, the error probability resulting from nearest neighbor codewords dominates. The dominant error events have probability [18, Fig.8.3-1], [10]

$$\mathrm{P_E} \approx N(d_{\mathrm{free}}) Q\left(\sqrt{\frac{d_{\mathrm{free}}^2}{2N_0}}\right) \approx 0.5 N(d_{\mathrm{free}}) \exp\left(-\frac{d_{\mathrm{free}}^2}{4N_0}\right) \tag{34}$$

where $N(d_{\text{free}})$ is the number of nearest neighbor codewords with Euclidean distance d_{free}. Along with (4) for the uncoded case, we propose to approximate the BER by

$$\mathrm{P}_{\mathrm{b,TCM}} \approx c_2 \mathrm{P}_{\mathrm{E}} \approx c_3 \exp\left(-\frac{d_{\text{free}}^2}{4N_0}\right) \tag{35}$$

where the constants c_2 and c_3 need to be determined. For each chosen trellis code, we will use one constant c_3 for all possible constellations, to facilitate the adaptive modulation process. For each chosen trellis code and signal constellation M_i, the ratio of d_{free}^2/d_0^2 is fixed. For each prescribed trellis code, we define

$$g_i' = \frac{d_{\text{free}}^2}{d_0^2} g_i, \quad \text{for the constellation } M_i. \tag{36}$$

Substituting (36) and (3) into (35), we obtain the approximate BER for constellation M_i as

$$\mathrm{P}_{\mathrm{b,TCM}}(M_i) \approx c_3 \exp\left(-\frac{g_i' E_s}{N_0}\right). \tag{37}$$

Example 5 (BER Approximation for Four-State TCM): We check the four-state trellis code with $k = r = 1$ [18, Fig. 8.3-6]. The constellations of size $M_i = 2^i, \forall i \in [2, 8]$ are divided into four subsets, following the set partitioning procedure in [18, Fig. 8.3-2]. Let d_j denote the minimum distance after the jth set partitioning. For QAM constellations, we have $d_{j+1}/d_j = \sqrt{2}$ [18, p. 524]. When $M > 4$, parallel transitions dominate with $d_{\text{free}}^2 = d_2^2 = 4d_0^2$. With $M = 4$, no parallel transition exists, and we have $d_{\text{free}}^2 = d_0^2 + 2d_1^2 = 5d_0^2$. We find the parameter $c_3 = 1.5 = 0.375N(d_{\text{free}})$ for the four-state trellis, where $N(d_{\text{free}}) = 4$ [18, Table 8.3.3]. The simulated BER and the approximate BER in (37) are plotted in Fig. 5. The approximation is within 2 dB for BER less than 10^{-1}.

Example 6 (BER Approximation for Eight-State TCM): We also check the eight-state trellis code with $k = 2$ and $r = 1$ [18, Fig. 8.3-10]; the trellis is also plotted in Fig. 6. The constellations of size $M = 2^i, \forall i \in [3, 8]$ are divided into eight subsets. The subset sequences dominate the error performance with $d_{\text{free}}^2 = d_0^2 + 2d_1^2 = 5d_0^2$ for all constellations. We choose $c_3 = 6 = 0.375N(d_{\text{free}})$ for the eight-state trellis code, where $N(d_{\text{free}}) = 16$ [18, Table 8.3.3]. The approximation is within 2 dB for BER less than 10^{-1}. We skip the plot for brevity. ■

B. Adaptive TCM for Fading Channels

We are now ready to specify the adaptive coded modulation with mean feedback. Since the transmitted symbols are correlated in time, we explicitly associate a time index t for each variable, e.g., we use $\mathbf{H}(t)$ to denote the channel perceived at time t. We calculate the following average error probability at time t based on (11) and (37):

$$\begin{aligned}\bar{\mathrm{P}}_{\mathrm{b,TCM}}(M_i, t) &= \mathrm{E}\left\{\mathrm{P}_{\mathrm{b,TCM}}(\mathbf{H}(t), M_i)\right\} \\ &\approx c_3 \mathrm{E}\left\{\exp\left(-h_{\mathrm{eqv}}^2(t)\frac{g_i' E_s}{N_0}\right)\right\}. \end{aligned} \tag{38}$$

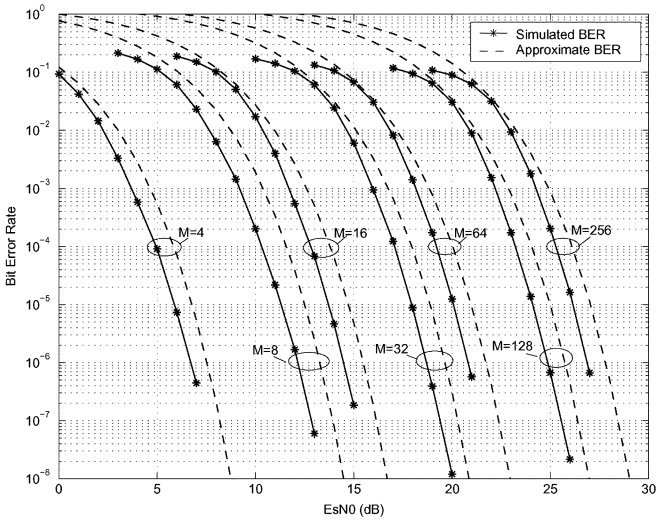

Fig. 5. BER approximation for trellis coded modulation (four-state trellis).

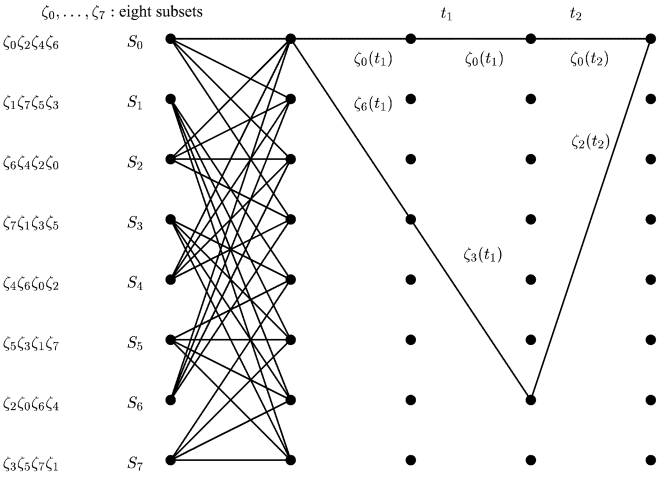

Fig. 6. One possible error path in adaptive TCM (eight-state trellis).

At each time t when updated feedback arrives, our transmitter chooses the constellation

$$M(t) = \arg\max_{M \in \{M_i\}_{i=k+r}^{N}} \bar{\mathrm{P}}_{\mathrm{b,TCM}}(\mathrm{M}, \mathrm{t}) \le \mathrm{P}_{\mathrm{b,target}}. \tag{39}$$

By the similarity of (37) and (5), we end up with an uncoded problem, with constellation M_i having a modified constant g_i' and conveying $\log_2(M_i) - r$ bits.

However, distinct from uncoded modulation, the coded transmitted symbols are correlated in time. Suppose that the channel feedback is frequent. The subset sequences may span multiple feedback updates, and thus different portions of one subset sequence may use subsets partitioned from different constellations. We show an example in Fig. 6. Our transmitter design in (39) implicitly assumes that all dominating error events are confined within one feedback interval. Nevertheless, we show next that this design guarantees the target BER for all possible scenarios. Since the dominating error events may occur between

parallel transitions, or, between subset sequences, we explore all the possibilities in the following.

1) Parallel Transitions Dominate: The parallel transitions occur in one symbol interval, and thus depend only on one constellation selection. The transmitter adaptation in (39) is in effect.

2) Subset Sequences Dominate: The dominating error events may be limited to one feedback interval, or, may span multiple feedback intervals. If the dominating error events are within one feedback interval, the transmitter adaptation in (39) is certainly effective. On the other hand, the error path may span multiple feedback intervals, with different portions of the error path using subsets partitioned from different constellations.

We focus on any pair of subset sequences $\mathbf{c}_1$ and $\mathbf{c}_2$. For brevity, we assume that the error path spans two feedback intervals (or updates), at time t_1 and t_2. Different constellations are chosen at time t_1 and t_2, resulting in different $d_0^2(t_1)$ and $d_0^2(t_2)$. We resort to Fig. 6 to describe a simple example. The distance between $\mathbf{c}_1$ and $\mathbf{c}_2$ can be partitioned as: $d^2(\mathbf{c}_1,\mathbf{c}_2\,|\,t_1,t_2)=\tilde{d}^2(t_1)+\tilde{d}^2(t_2)$. The contribution of $\tilde{d}^2(t_1)$ at time t_1 is the minimum distance between subsets $\zeta_0(t_1)$ and $\zeta_2(t_1)$ plus the minimum distance of subsets $\zeta_0(t_1)$ and $\zeta_3(t_1)$, i.e., $\tilde{d}^2(t_1)=d_1^2(t_1)+d_0^2(t_1)=3d_0^2(t_1)$. Similarly, we have $\tilde{d}^2(t_2)=d_1^2(t_2)=2d_0^2(t_2)$.

Now, we construct two virtual events that the error path between $\mathbf{c}_1$ and $\mathbf{c}_2$ experiences only one feedback: one at t_1 and the other at t_2. For $j=1,2$, the average pairwise error probability is defined as

$$\bar{\mathrm{P}}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_j)=0.5\mathrm{E}\left\{\exp\left(-\frac{h_{\mathrm{eqv}}^2(t_j)d^2(\mathbf{c}_1,\mathbf{c}_2\,|\,t_j)}{4N_0}\right)\right\}. \tag{40}$$

We define the constants as

$$b_1:=\frac{\tilde{d}(t_1)}{d^2(\mathbf{c}_1,\mathbf{c}_2\,|\,t_1)}$$
$$b_2:=\frac{\tilde{d}(t_2)}{d^2(\mathbf{c}_1,\mathbf{c}_2\,|\,t_2)}. \tag{41}$$

It is clear that $b_1+b_2=1$, and $0<b_1,b_2\le 1$.

We next show that when the error path between $\mathbf{c}_1$ and $\mathbf{c}_2$ spans multiple feedback intervals, the average pairwise error probability (PEP) decreases, relative to the case of one feedback interval. Since the conditional channels at different times are independent, we have

$$\begin{aligned}&\mathrm{E}\{\mathrm{P}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_1,t_2)\}\\&\quad=0.5\mathrm{E}\left\{\exp\left(-\frac{h_{\mathrm{eqv}}^2(t_1)\tilde{d}^2(t_1)}{4N_0}\right)\right\}\\&\qquad\times\mathrm{E}\left\{\exp\left(-\frac{h_{\mathrm{eqv}}^2(t_2)\tilde{d}^2(t_2)}{4N_0}\right)\right\}\\&\quad\le 0.5\left[\frac{\bar{\mathrm{P}}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_1)}{0.5}\right]^{b_1}\left[\frac{\bar{\mathrm{P}}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_2)}{0.5}\right]^{b_2}\\&\quad\le\max\left(\bar{\mathrm{P}}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_1),\bar{\mathrm{P}}(\mathbf{c}_1\to\mathbf{c}_2\,|\,t_2)\right)\end{aligned} \tag{42}$$

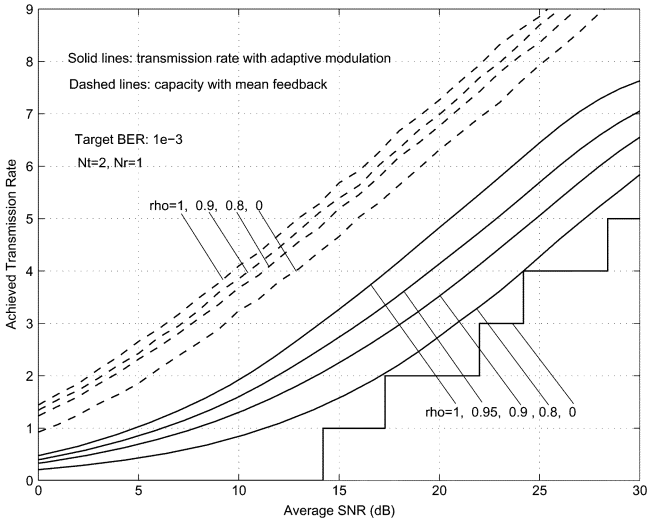

Fig. 7. Transmission rates for variable feedback quality.

where in deriving (42), we have used the inequality in (47) proved in the Appendix. Equation (42) reveals that the worst case happens when the error path between subset sequences spans only one feedback. In such cases, however, we have guaranteed the average BER in (39), for each feedback. When the error path in dominating error events spans multiple feedback intervals, the average pairwise error probability decreases, and thus the average BER [proportional to the dominating pairwise error probability as approximated in (35)] is guaranteed to stay below the target.

In summary, the transmitter adaptation in (39) guarantees the prescribed BER. With perfect CSI, this adaptation reduces to that in [10], where d_0 is maintained for each constellation choice. Compared with [8], our approach here is simpler in the sense that we do not need to check all distances between each pair of subsets. If indeed multiple distances are to be checked, the optimal power allocation between two beams will become involved. Notice that an interleaver is introduced in [8] to improve performance by distributing the error path between subset sequences to multiple channel feedback intervals. This intuition is theoretically verified by (42), which is not available in [8]. However, interleaving may not be feasible due to the inherent lack of large time diversity within a reasonable delay.

VI. Numerical Results

In our simulations, we adopt the channel setup of Example 2, with $\sigma_h^2=1$. Recall that the feedback quality σ_ϵ^2 is related to the correlation coefficient $\rho=J_0(2\pi f_d\tau)$ via: $\sigma_\epsilon^2=1-|\rho|^2$. With $\rho=0.95,0.9,0.8$, we have $\sigma_\epsilon^2=-10.1,-7.2,-4.4$ dB. For fair comparison among different setups, we use in all plots the average received SNR defined as

$$\text{average SNR}:=(1-\mathrm{P}_{\mathrm{out}})E_s/N_0. \tag{43}$$

Case 1 (Distance From Capacity): We set $\mathrm{P}_{\mathrm{b,target}}=10^{-3}$. Fig. 7 plots the rate achieved by the proposed adaptive transmitter with $N_t=2,N_r=1$, and $\rho=1,0.95,0.9,0.8,0$. It is clear that the rate decreases relatively fast as the feedback quality drops.

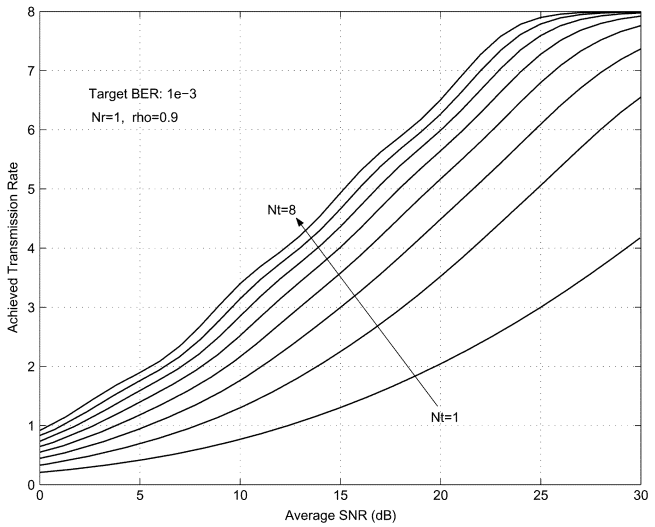

Fig. 8. Rate improvement with multiple transmit antennas.

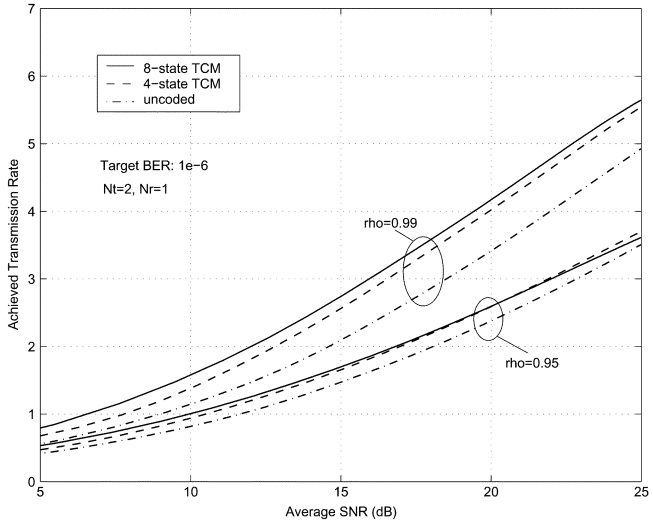

Fig. 10. Rate improvement with trellis coded modulation.

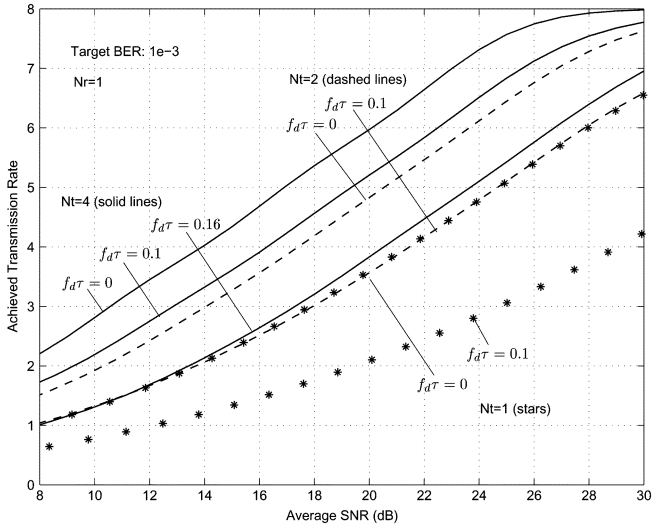

Fig. 9. Tradeoff between feedback delay and hardware complexity.

For comparison, we also plot the channel capacity with mean feedback [23], using the semianalytical result in [15]. As shown in Fig. 7, the capacity is less sensitive to channel imperfections. The capacity with perfect CSI is larger than the capacity with no CSI by about $\log_2(N_t) = 1$ bit at high SNR, as predicted in [21]. With $\rho = 0.9$, the adaptive uncoded modulation is about 11 dB away from capacity.

Case 2 (Rate Improvement With Antennas): We set $N_r = 1, \mathrm{P}_{\mathrm{b,target}} = 10^{-3}$, and $\rho = 0.9$. As shown in Fig. 8, the achieved transmission rate increases as the number of transmit antennas increases. The largest rate improvement occurs when N_t increases from one to two. Notice that the achieved rate with $N_t = 1$ is sightly less than that in [8, Fig. 4], due to the lack of energy adaptation.

Case 3 (Tradeoffs Between Feedback Quality and Hardware Complexity): It is shown in [4] that the critical value is $f_d\tau = 0.01$ for single antenna transmissions. In Fig. 9, we verify that with two transmit antennas, the achieved rate with $f_d\tau = 0.1$ ($\rho = 0.904$) coincides with that corresponding to one transmit antenna with perfect CSI ($f_d\tau \leq 0.01$); hence, more than ten times of feedback delay can be tolerated. The rate with $N_t = 4$ and $f_d\tau = 0.16$ ($\rho = 0.76$) is even better than that of $N_t = 1$ with perfect CSI. To achieve the same rate, the delay constraint with single antenna can be relaxed considerably by using more transmit antennas, an interesting tradeoff between feedback quality and hardware complexity.

Fig. 9 also reveals that the adaptive design becomes less sensitive to CSI imperfections, when the number of transmit antenna increases.

Case 4 (TCM): We test the four-state and eight-state trellis codes listed in Examples 5 and 6. We first set $\mathrm{P}_{\mathrm{b,target}} = 10^{-6}, N_t = 2, N_r = 1$. When the feedback quality is near perfect ($\rho = 0.99$), the rate is considerably increased by using trellis coded modulation instead of uncoded modulation, in agreement with the perfect CSI case [10]. However, the achieved SNR gain decreases quickly as the feedback quality drops, as shown in Fig. 10. This can be predicted from (47), since increasing the Euclidean distance by TCM with set partitioning is less effective for fading channels ($\rho < 1$) than for AWGN channels ($\rho = 1$). If affordable, coded bits can be interleaved to benefit from time diversity, as suggested in [8]. This is suitable for the eight-state TCM, where the subset sequences dominate the error performance.

On the other hand, the Euclidean distance becomes the appropriate performance measure, when the number of receive antennas increases, as established in [5]. The SNR gain introduced by TCM is thus restored, as shown in Fig. 11 with $N_r = 2, 4$.

Comparing Fig. 10 with Fig. 7, one can observe that the adaptive system is more sensitive to noisy feedback when the prescribed bit error rate is small (10^{-6}) as opposed to large (10^{-3}).

VII. Conclusion

In this paper, we introduced adaptive modulation for multiantenna transmissions with channel mean feedback. Based on a 2-D beamformer, the proposed transmitter optimally adapts the basis beams, the power allocation between two beams, and the signal constellation, to maximize the transmission rate while

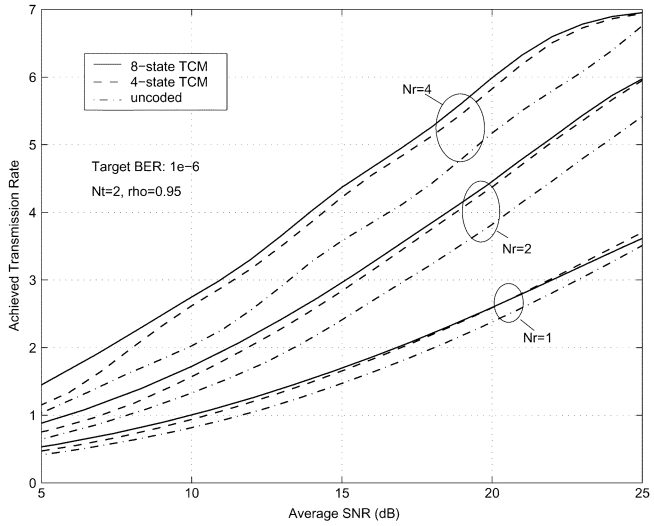

Fig. 11. Impact of receive diversity on adaptive TCM.

guaranteeing a target BER. Both uncoded and TCM were addressed. Numerical results demonstrated the rate improvement enabled by adaptive multiantenna modulation, and pointed out an interesting tradeoff between feedback quality and hardware complexity.

The proposed adaptive modulation maintains low receiver complexity thanks to the Alamouti structure. Relaxing the receiver complexity constraint, adaptive modulation based on spatial multiplexing schemes is an interesting future research topic.

APPENDIX

Let γ denote a nonnegative random variable with pdf $p(\gamma)$, and mean $\bar{\gamma}$. We wish to prove that the following function:

$$\phi(b) = [\mathrm{E}\{e^{-\gamma b}\}]^{\frac{1}{b}} = \left[\int e^{-\gamma b} p(\gamma) d\gamma\right]^{\frac{1}{b}} \tag{44}$$

is a nondecreasing function of $b, \forall b > 0$. Toward this goal, we need to verify

$$\frac{d(\ln(\phi(b))}{db} = \frac{1}{b^2} \underbrace{\left[\ln \frac{1}{\int e^{-\gamma b} p(\gamma) d\gamma} - \frac{b \int \gamma e^{-\gamma b} p(\gamma) d\gamma}{\int e^{-\gamma b} p(\gamma) d\gamma}\right]}_{:=\varphi(b)} \geq 0. \tag{45}$$

Now it remains to show that $\varphi(b) \geq 0$. Using the Cauchy–Schwartz inequality [18, p. 161], we verify that

$$\frac{d(\varphi(b))}{db} = b \frac{\int \gamma^2 e^{-\gamma b} p(\gamma) d\gamma \int e^{-\gamma b} p(\gamma) d\gamma - \left[\int \gamma e^{-\gamma b} p(\gamma) d\gamma\right]^2}{\left[\int e^{-\gamma b} p(\gamma) d\gamma\right]^2} \geq 0. \tag{46}$$

Therefore, $\varphi(b)$ is nondecreasing with b, and $\varphi(b) \geq \varphi(0) = 0$. Hence, we proved that $\phi(b)$ is nondecreasing with b.

Since $\phi(1) = \mathrm{E}\{e^{-\gamma}\}$, we establish the following two inequalities:

$$\mathrm{E}\{e^{-\gamma b}\} \leq [\mathrm{E}\{e^{-\gamma}\}]^b, \quad \forall 0 < b \leq 1 \tag{47}$$

$$\mathrm{E}\{e^{-\gamma b}\} \geq [\mathrm{E}\{e^{-\gamma}\}]^b, \quad \forall b \geq 1. \tag{48}$$

We can also provide an intuitive explanation of these results. If we view γ as the received SNR, then $e^{-\gamma}$ is proportional to the approximate BER of a BPSK transmission, as evidenced from (5). When $b < 1$, the received SNR is decreasing, while the received SNR is increasing when $b > 1$. The equalities in (47) and (48) hold only when $p(\gamma)$ is a delta function $p(\gamma) = \delta(\gamma - \bar{\gamma})$; in other words, when γ is the SNR of a nonfading channel. Equations (47) and (48) simply point out the fact that the average BER $\mathrm{E}\{e^{-\gamma b}\}$ for an arbitrary fading channel deteriorates (when $b < 1$), or, improves (when $b > 1$), with a lower speed compared to a nonfading channel, whose BER is

$$\mathrm{E}\{e^{-\gamma b}\} = e^{-\bar{\gamma} b} = [\mathrm{E}\{e^{-\gamma}\}]^b.$$

In other words, the slopes of BER curves of arbitrary fading channels are no larger than that of a nonfading channel at any fixed BER level.

REFERENCES

[1] *Physical Layer Procedures*, 3GPP RAN 25.214 V1.3.0, 1999.
[2] *Physical Channels and Mapping of Transport Channels Onto Physical Channels*, 3GPP TS 25.211 V2.4.0, 1999.
[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
[4] M. S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *Kluwer J. Wireless Commun.*, vol. 13, no. 1–2, pp. 119–143, May 2000.
[5] E. Biglieri, G. Taricco, and A. Tulino, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1794–1803, July 2002.
[6] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view," *IEEE Trans. Commun.*, vol. 49, pp. 1561–1571, Sept. 2001.
[7] A. Duel-Hallen, S. Hu, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Processing Mag.*, vol. 17, pp. 62–75, May 2000.
[8] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, pp. 844–855, June 1999.
[9] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.
[10] ——, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 595–602, May 1998.
[11] K. J. Hole, H. Holm, and G. E. Oien, "Adaptive multidimensional coded modulation over flat fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1153–1158, July 2000.
[12] S. Hu and A. Duel-Hallen, "Combined adaptive modulation and transmitter diversity using long range prediction for flat fading mobile radio channels," in *Proc. Global Telecommunications Conf.*, vol. 2, San Antonio, TX, Nov. 25–29, 2001, pp. 1256–1261.
[13] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory*, vol. 48, pp. 611–627, Mar. 2002.
[14] V. K. N. Lau and M. D. Macleod, "Variable-rate adaptive trellis coded QAM for flat-fading channels," *IEEE Trans. Commun.*, vol. 49, pp. 1550–1560, Sept. 2001.
[15] A. Moustakas and S. Simon, "Optimizing multitransmitter single-receiver (MISO) communication systems with general Gaussian channels: Nontrivial covariance and nonzero mean," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2770–2880, Oct. 2003.

[16] 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 J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.
[17] S. Otsuki, S. Sampei, and N. Morinaga, "Square-QAM adaptive modulation/TDMA/TDD systems using modulation level estimation with Walsh function," *Electron. Lett.*, vol. 31, pp. 169–171, Feb. 1995.
[18] J. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2000.
[19] M. K. Simon and M.-S. Alouini, *Digital Communication Over Generalized Fading Channels: A Unified Approach to the Performance Analysis*. New York: Wiley, 2000.
[20] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
[21] I. E. Telatar, "Capacity of multiantenna Gaussian channels," *Bell Lab. Tech. Memo.*, 1995.
[22] T. Ue, S. Sampei, N. Morinaga, and K. Hamaguchi, "Symbol rate and modulation level-controlled adaptive modulation/TDMA/TDD system for high-bit-rate wireless data transmission," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 1134–1147, Nov. 1998.
[23] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.
[24] P. K. Vitthaladevuni and M. S. Alouini, "BER computation of generalized QAM constellations," in *Proc. Global Telecommunications Conf.*, vol. 1, San Antonio, TX, Nov. 2001, pp. 632–636.
[25] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.
[26] P. Xia, S. Zhou, and G. B. Giannakis, "Adaptive MIMO OFDM based on partial channel state information," *IEEE Trans. Signal Processing*, vol. 52, pp. 202–213, Jan. 2004.
[27] D. Yoon and K. Cho, "General bit error probability of rectangular quadrature amplitude modulation," *Electron. Lett.*, vol. 38, no. 3, pp. 131–133, Jan. 2002.
[28] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Processing*, vol. 50, pp. 2599–2613, Oct. 2002.
[29] ——, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels?," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1285–1294, July 2004.

Shengli Zhou (M'03) received the B.S. and M.Sc. degrees in electrical engineering and information science from the University of Science and Technology of China, Hefei, in 1995 and 1998, respectively, and the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 2002.

He has been an Assistant Professor with the Department of Electrical and Computer Engineering, the University of Connecticut, Storrs, in 2003. His research interests lie the areas of communications and signal processing, including channel estimation and equalization, multiuser and multicarrier communications, space–time coding, adaptive modulation, and cross-layer designs.

Georgios B. Giannakis (S'84–M'86–SM'91–F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, 1981, and the MSc. degree in electrical engineering, the MSc. degree in mathematics, and the Ph.D. degree in electrical engineering, from the University of Southern California (USC), Los Angeles, in 1983, 1986, and 1986, respectively.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of Electrical Engineering in 1997. Since 1999, he has been a Professor with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 200 journal papers, 350 conference papers, and two edited books. His current research focuses on transmitter and receiver diversity techniques for single-user and multiuser fading communication channels, complex-field and space–time coding, multicarrier, ultrawideband wireless communication systems, cross-layer designs, and sensor networks.

Dr. Giannakis is the (co-) recipient of six best paper awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001, 2003, 2004). He also received the SP Society's Technical Achievement Award in 2000. He has served as Editor in Chief for theIEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as Secretary of the SP Conference Board, as Member of the SP Publications Board, as Member and Vice-Chair of the Statistical Signal and Array Processing Technical Committee, as Chair of the SP for Communications Technical Committee, and as a member of the IEEE Fellows Electrion Committee. He has also served as a member of the IEEE-SP Society's Board of Governors, the Editorial Board of the PROCEEDINGS OF THE IEEE and the Steering Committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.