Multiantenna Adaptive Modulation With Beamforming Based on Bandwidth-Constrained Feedback

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

***Abstract*—Adaptive modulation has the potential to increase system throughput considerably by adapting transmission parameters to the time-varying channel characteristics. Crucial to adaptive systems is the requirement of a feedback channel, that is often capable of carrying only a limited number of bits. Under such a bandwidth-constrained feedback link, we aim to optimize a multiantenna system based on transmit beamforming and adaptive modulation, where the transmit power, the signal constellation, the beamforming direction, and the feedback strategy, are designed jointly. Our proposed nested iterative approach leads to an approximate, yet practical, solution. Simulation results demonstrate considerable improvement in transmission rate, as the number of feedback bits increases.**

***Index Terms*—Adaptive modulation, beamforming, feedback, vector quantization.**

I. Introduction

By matching transmitter parameters to the time-varying channel conditions, adaptive modulation can increase the transmission rate considerably, which justifies its popularity in high-rate wireless services; see, e.g., [4], [8], and [11]–[13], and references therein. On the other hand, antenna diversity has been well established as an effective fading counter measure for wireless applications [23]. Size and cost limitations of mobile units suggest multiple transmit antennas at the base station, which motivates adaptive designs with multiple transmit antennas.

Crucial to adaptive systems is the requirement of a feedback channel, through which the receiver feeds the estimated channel state information (CSI) to the transmitter. Although perfect CSI is often assumed in the adaptive modulation literature (see, e.g., [8] and [11]), in practice, the CSI at the transmitter suffers from imperfections originating from various sources, which include estimation/quantization errors, feedback delay, and feedback errors. Recent works quantifying and exploiting imperfect (or partial) CSI in multiantenna systems include [5], [6], [15], [17], [19]–[22], [24], [26], [28], and [29].

Partial CSI can be modeled in various ways. One approach is to describe it statistically based on feedback information. For example, CSI at the transmitter can be modeled as Gaussian distributed with local mean and local covariance [22], [24]. Under this model, two special forms of partial feedback have been extensively studied. One is the so-called mean feedback [15], [17], [24], [28], that assumes knowledge of the channel mean and models the channel covariance as white. The other is termed covariance feedback, where the channel mean is set to zero, and the relative geometry of the propagation paths manifests itself in a generally nonwhite covariance matrix [15], [24].

Another class of CSI models imposes a bandwidth constraint on the feedback channel which is only able to communicate a finite number of feedback bits per block [5], [6], [16], [19]–[22]. In particular, finite-rate power control is investigated in [5] to reduce the outage probability that the mutual information falls below a certain rate. Transmit beamforming without power control has been studied using average signal-to-noise ratio (SNR) [20], [22], or outage probability [21] criteria. Moreover, a jointly optimal transmission and feedback strategy has been pursued in [19] (see also [6] for the same formulation) to maximize the forward channel capacity subject to bandwidth constraints on the feedback channel. More recently, an analytical codebook construction method has been presented in [31]

We have studied adaptive multiantenna modulation based on channel mean feedback in [26] and [29]. In this paper, we investigate an adaptive modulation system based on transmit beamforming with *rate-limited* feedback. Beamforming is simple to implement [15], [21], and has been adopted in current standards [1], [2]. Subject to average bit-error rate (BER) and average transmit-power constraints, our objective is to jointly design the feedback strategy and the transmission parameters for a rate-limited feedback. We formulate the problem and link it with existing works, assuming perfect or no CSI in Section II. Due to the inherent difficulty in solving the original problem analytically, we target instead a suboptimal formulation in Section III. We then develop a nested iterative approach that yields a practical (albeit suboptimal) design of the adaptive transmission and the feedback strategy. Numerical results are presented

Paper approved by M. Skoglund, the Editor for Source/Channel Coding of the IEEE Communications Society. Manuscript received September 29, 2003; revised November 4, 2004. This work was supported in part through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011, and in part by the National Science Foundation under Grant 01-0516. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon. This paper was presented in part at the 37th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 9–12, 2003.

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

S. Zhou was with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA. He is now with the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269 USA (e-mail: shengli@engr.uconn.edu).

Digital Object Identifier 10.1109/TCOMM.2005.843431

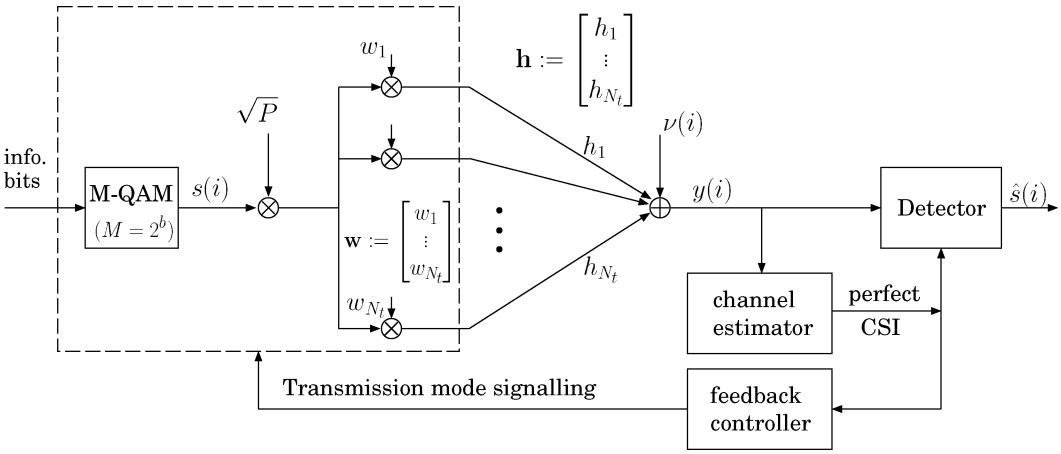


Fig. 1. Overall system model.

in Section IV, which demonstrate the throughput enhancement as the number of feedback bits increases.

Notation: Bold upper- and lower-case letters denote matrices and column vectors, respectively; $\|\cdot\|$ denotes vector norm; $(\cdot)^T$ and $(\cdot)^{\mathcal{H}}$ stand for transpose and Hermitian transpose, respectively; $\mathcal{CN}(0,\sigma_h^2)$ denotes complex normal distribution with zero mean and variance σ_h^2; $\mathrm{E}_A[q]$ denotes the ensemble average of q conditioned on the event A; $\measuredangle$ denotes angle; $\ln(\cdot)$ is the natural logarithm function, and $\delta(\cdot)$ is the Dirac delta function.

II. System Description and Problem Formulation

A. Channel Model

For clarity, we only consider a wireless communication system with multiple (N_t) transmit antennas and a single receive antenna, as depicted in Fig. 1. However, the extension to multiple receive antennas with beamforming is straightforward. Our focus will be on narrowband block transmissions where the wireless channel is frequency nonselective. Let h_l denote the channel coefficient between the lth transmit and the receive antenna, where $1 \le l \le N_t$, and define the channel vector $\mathbf{h} := [h_1, h_2, \ldots, h_{N_t}]^T$. We assume the following.

A1: Channel realizations do not vary within a block, but can change from block to block.

Assumption A1 corresponds to the so-called *block-fading* channel model that is suitable for many practical wireless systems [23]. Furthermore, we assume the following.

A2: The channel coefficients are independently and identically distributed (i.i.d.), with each coefficient being modeled as a circularly complex Gaussian random variable with zero mean and variance σ_h^2; i.e., $h_l \sim \mathcal{CN}(0,\sigma_h^2), \forall l \in [1, N_t]$.

This generally requires a rich scattering environment and transmit antennas placed sufficiently far apart from each other. We assume perfect knowledge of σ_h^2 for our transceiver design. Estimation and quantization error effects of σ_h^2 are interesting to investigate, but go beyond the scope of this paper. We will see that A2 actually is not needed in our transceiver design, where only a statistical description of the channel vector is required to generate the sampling channel values [c.f. (21)]. To assist the receiver with channel estimation and symbol detection, preamble training sequences are placed at the beginning of each block. Similar to [11], [14], and [29], we assume the following.

A3: Perfect CSI is available at the receiver, which allows one to isolate the effects of rate-limited feedback on the transceiver design.

To carry out adaptive modulation and transmit beamforming, CSI needs to be fed back to the transmitter. In accordance with A1, feedback is updated once every block. As in [19]–[22], we focus on a bandwidth-constrained feedback link. Formally, we assume the following.

A4: The feedback channel is error-free, delay-free, but limited in bandwidth, such that the transmitter can acquire only a limited number of, say B, feedback bits per fading block.

The error-free assumption can be well approximated through the use of sufficiently powerful error-control codes; whereas the delay-free assumption is accurate when the processing and feedback delays are small, relative to the channel coherence time.

B. System Model and Problem Statement

We adopt adaptive modulation based on transmit beamforming as the basic transmission strategy. Based on CSI feedback per block, the transmitter draws an information symbol $s(i)$ from an appropriate signal constellation of size M with average energy $E_s = 1$, and transmits the vector $\mathbf{w}^{\mathcal{H}}\sqrt{P}s(i)$ through multiple antennas, where i stands for the symbol index within each block, P denotes transmit power, and $\mathbf{w}$ is the beamsteering vector with unit norm $\|\mathbf{w}\| = 1$. The received signal is thus

$$y(i) = \mathbf{w}^{\mathcal{H}}\mathbf{h}\sqrt{P}s(i) + \nu(i), \tag{1}$$

where $\nu(i)$ is additive white Gaussian noise (AWGN) with zero mean and variance $N_0/2$ per real and imaginary dimension. For notational simplicity, the symbol index i will be omitted in the following. Each triplet of the transmission parameters $(M, P, \mathbf{w})$ constitutes what we will henceforth term a transmission mode. Notice that different transmission modes may adopt the same signal constellation.

According to A4, only B feedback bits are available per block. This decides a feedback index variable $n \in \{1,2,\ldots,N\}$, where $N := 2^B$. Corresponding to each index n, the transmitter will select transmission mode $\mathcal{M}_n$ as

$$\mathcal{M}_n := (M_n, P_n, \mathbf{w}_n), \quad n \in \{1,2,\ldots,N\}.$$

For the candidate signal constellations, we will consider two different cases: continuous rate and discrete rate, as in [8] and [11]. For the continuous-rate case, $M_n = 2^{b_n}$, where $b_n \geq 0$ is an arbitrary real number. For the discrete-rate case, $M_n = 2^2, 2^4, 2^6, \ldots$; i.e., the b_n's are restricted to be even integers, corresponding to square M-ary quadrature amplitude modulation (M-QAM) constellations. To account for deep fading effects, we also allow for no data transmission, which amounts to setting $M_n = 2^0$. As we will see later, investigation of the continuous-rate case provides the basis for system designs in the discrete-rate case.

Let $\mathcal{R}$ denote the N_t-dimensional space of $N_t \times 1$ complex channel vectors. The B feedback bits are used to index each of N nonoverlapping regions $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_N$ of the channel space $\mathcal{R}$. Apparently, selection of these fading regions and transmission modes affects the achievable data rate. Our goal is to design $\{\mathcal{R}_n\}_{n=1}^N$ and $\{\mathcal{M}_n\}_{n=1}^N$ *jointly* to maximize the system transmission rate, subject to certain performance and power constraints, as we formulate next. To simplify the adaptive transceiver design, we rely on the following approximate BER expression [8], [11]:

$$\mathrm{BER} \approx 0.2 \exp(-g_n \cdot \gamma) \tag{2}$$

where $g_n := 1.5/(2^{b_n} - 1)$ is a constellation-dependent constant, and γ is the receive SNR. When the selected transmission mode is $\mathcal{M}_n$, we have from (1)

$$\gamma_n = \left|\mathbf{w}_n^{\mathcal{H}} \mathbf{h}\right|^2 \cdot \frac{P_n}{N_0} \tag{3}$$

and the average BER over the region $\mathcal{R}_n$ can be written as

$$\overline{\mathrm{BER}}_n = \mathrm{E}_{\mathcal{R}_n}\left[0.2 \exp\left(-\frac{g_n \left|\mathbf{w}_n^{\mathcal{H}} \mathbf{h}\right|^2 P_n}{N_0}\right)\right]. \tag{4}$$

Let $A_n := \Pr(\mathbf{h} \in \mathcal{R}_n)$ denote the probability that the channel vector $\mathbf{h}$ lies in region $\mathcal{R}_n$, P_0 stand for the average power, and $\overline{\mathrm{BER}}_0$ be the required BER performance. By jointly designing $\{\mathcal{R}_n\}_{n=1}^N$ and $\{\mathcal{M}_n\}_{n=1}^N$, our ultimate objective is to

$$\begin{aligned}
&\text{maximize} \quad \sum_{n=1}^{N} A_n b_n, \\
&\text{subject to (c1)} \quad \sum_{n=1}^{N} A_n P_n \leq P_0 \\
&\qquad (\mathsf{c2}) \quad \|\mathbf{w}_n\| = 1, \ \forall n \\
&\qquad (\mathsf{c3}) \quad b_n \geq 0 (\text{continuous rate}) \\
&\qquad\qquad\ \ b_n \in \{0,2,4,6,\ldots\} (\text{discrete rate}), \ \forall n \\
&\qquad (\mathsf{c4}) \quad \overline{\mathrm{BER}}_n \leq \overline{\mathrm{BER}}_0, \ \forall n.
\end{aligned} \tag{5}$$

Once the fading regions $\mathcal{R}_1, \ldots, \mathcal{R}_N$ and the transmission modes $\mathcal{M}_1, \ldots, \mathcal{M}_N$ are decided, our adaptive system will operate as follows: per fading block, the receiver first decides which region the current channel $\mathbf{h}$ falls into, say $\mathcal{R}_n$, and feeds back the region index n. The transmitter then uses this index to select a transmission mode, and transmits information symbols with the selected transmission mode $\mathcal{M}_n$.

Different from the joint adaptation strategy studied in this paper, it is also possible to design the adaptive modulation and beamforming strategy separately. For example, one can adopt adaptive modulation along the lines of [11], on top of the beamforming scheme with quantized CSI [20], [21], [25], [30]. Probability density function (pdf) of the effective single-input single-output (SISO) channel, however, is generally unavailable as evidenced in [21], [25], and [30], where certain performance bounds are pursued. More importantly, given a total number of B feedback bits, it remains unknown how many bits should be allocated for adaptive modulation, and how many should be assigned to quantized beamforming. We will briefly study some heuristic separate design examples through Monte Carlo simulations in Section IV.

C. Upper and Lower Bounds

Before we tackle the general problem in (5), let us consider the special cases corresponding to $B = \infty$ and $B = 0$. They specify the upper and lower bounds on the achievable transmission rates that will be used to benchmark our system with a finite B. When $B = \infty$, perfect CSI can be made available to the transmitter, i.e., the transmitter knows each realization of $\mathbf{h}$. In this case, for any given power P and signal constellation M, the optimal beamforming vector is given by $\mathbf{w} = \mathbf{h}/\|\mathbf{h}\|$ [21], [22]. Consequently, (1) can be simplified to

$$y = \|\mathbf{h}\|\sqrt{P}s + \nu := z\sqrt{P}s + \nu. \tag{6}$$

Equation (6) describes a single-input single-output (SISO) system with a flat fading coefficient z. Under A2, z is Nakagami-m distributed with average energy $\mathrm{E}[z^2] = N_t\sigma_h^2$ and shape parameter $m = N_t$ [4]. With both the instantaneous value and the distribution of z known at the transmitter, the optimal adaptive modulation solutions for both continuous and discrete rates are available; see [11, eq. (23)] for the continuous-rate case, and [8, eqs. (49)–(52)] for the discrete-rate case.

When $B = 0$, no CSI is available to the transmitter and no channel-space partitioning is performed in this case. Using the same approximation as in (2), the average BER can be written as [29, eq. (16)]

$$\overline{\mathrm{BER}} \approx \frac{0.2}{1 + \frac{gP\sigma_h^2}{N_0}}. \tag{7}$$

Based on (7), we only need to identify the highest modulation level that meets the average BER requirement (c4) in (5), using either continuous or discrete rate. We reiterate that the $B = 0$ case is provided here as a lower bound; i.e., we do not advocate transmit-beamforming-only when the transmitter has no CSI available. In such cases, space–time block coding [3] and linear constellation precoding has well-documented merits [27].

We will provide comparisons between [3], [27], and the proposed scheme with limited feedback bits in Section IV.

III. Design of Transmission Modes and Fading Regions

Equation (5) provides a general formulation whose optimal solution is hard to obtain. In this paper, we resort to a simplified, albeit suboptimal, problem formulation (Section III-A). The continuous-rate system design is carried out first (Sections III-B and III-C), based on which the discrete-rate system design is derived afterwards (Section III-D). Some implementation issues are then considered (Section III-E).

A. Suboptimal Formulation

The major difficulty with the original formulation (5) is that generally the regions $\{\mathcal{R}_n\}_{n=1}^N$ will be irregular in shape, which prevents an invertible analytical expression for the $\overline{\mathrm{BER}}_n$ in (4). Notice that when N is sufficiently large, the volume of each region is relatively small, so that all channels in each region would be close to each other. Based on this observation, one could approximate $\overline{\mathrm{BER}}_n$ by

$$\overline{\mathrm{BER}}_n' = 0.2\exp\left(-P_n g_n \mathbf{w}_n^{\mathcal{H}} \cdot \frac{\mathrm{E}_{\mathbf{h}\in\mathcal{R}_n}[\mathbf{h}\mathbf{h}^{\mathcal{H}}]\mathbf{w}_n}{N_0}\right). \quad (8)$$

Inserting $\overline{\mathrm{BER}}_n'$ instead of $\overline{\mathrm{BER}}_n$ into (c4) and rearranging it, one would find a candidate approximation for b_n

$$b_n' = \log_2\left(1+\frac{1.5P_n\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n}{N_0\cdot(-\ln(5\overline{\mathrm{BER}}_0))}\right) \quad (9)$$

where $\mathbf{R}_n := \mathbf{E}_{\mathbf{h}\in\mathcal{R}_n}[\mathbf{h}\mathbf{h}^{\mathcal{H}}]$ is the channel covariance matrix locally averaged over $\mathcal{R}_n$. Unfortunately, such an approximation is only effective when N becomes very large. For other values of N, it violates the BER constraint (c4), since e^{-x} is a convex function, and consequently, $\overline{\mathrm{BER}}_n \geq \overline{\mathrm{BER}}_n'$.

To bypass this difficulty, we consider a downward scaling approach, and artificially introduce a modifying factor $\beta \geq 1$, replacing (9) with

$$\tilde{b}_n = \log_2\left(1+\frac{1.5P_n\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n}{N_0\cdot\beta(-\ln 5\overline{\mathrm{BER}}_0)}\right) \quad (10)$$

which is equivalent to setting the target BER to $5^{\beta-1}\overline{\mathrm{BER}}_0^{\beta}$. The motivation behind (10) is to employ an average BER precompensation, i.e., to preset the BER requirement to a lower level $(5^{\beta-1}\overline{\mathrm{BER}}_0^{\beta})$, so that the original BER requirement (c4) can still be met. An appropriate β is thus needed, so that the final design of transmission modes and fading regions will meet (c4). Normally, β depends on various system parameters, such as the average SNR, the number of antennas N_t, the number of feedback bits B, and the original BER constraint $\overline{\mathrm{BER}}_0$. Lacking a general expression for it, we rely on Monte Carlo simulations to determine β empirically, as will be specified in Section IV. In the ensuing analysis, we assume that β has been determined.

Replacing b_n in (5) by $\tilde{b}_n$ in (10), we end up with a simplified, albeit suboptimal, formulation as follows

$$\begin{aligned}\text{minimize } D &:= -\sum_{n=1}^N A_n\tilde{b}_n \\ &= -\sum_{n=1}^N A_n\log_2\left(1+\frac{1.5P_n\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n}{N_0\cdot\beta(-\ln 5\overline{\mathrm{BER}}_0)}\right) \\ \text{subject to (c1)}\quad &\sum_{n=1}^N A_nP_n \leq P_0 \end{aligned} \quad (11)$$

where minimizing D is equivalent to maximizing $\sum_{n=1}^N A_n\tilde{b}_n$. In this paper, we will derive system designs that optimize (11) instead of the original general problem (5). We underscore that solving (11) will only provide a suboptimal transceiver design relative to what could be achieved, had (5) been directly solved. Notice that constraints (c2) and (c3) are not included explicitly in (11), but will be later on.

As we will verify soon, our transceiver design is closely related to a vector-quantization problem. However, standard vector-quantizer design algorithms can not be applied here because of the additional average power constraint. A similar problem arises in [9], where a vector quantizer is designed with an additional constraint on the codeword index entropy. Inspired by the approach taken therein, we first introduce a deviation-cost function

$$D(C) := \inf\left\{-\sum_{n=1}^N A_n\log_2\left(1+\frac{1.5P_n\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n}{N_0\cdot\beta(-\ln 5\overline{\mathrm{BER}}_0)}\right)\Bigg| \sum_{n=1}^N A_nP_n = C\right\} \quad (12)$$

i.e., $D(C)$ is the minimum achievable deviation with a power budget C. The solution to (11) is nothing but the value of $D(C)$ evaluated at $C = P_0$. Solving (12) is an optimization problem with equality constraints. We transform it to an unconstrained optimization problem by constructing the Lagrangian

$$J(\lambda) = -\sum_{n=1}^N A_n\log_2\left(1+\frac{1.5P_n\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n}{N_0\cdot\beta(-\ln 5\overline{\mathrm{BER}}_0)}\right)+\lambda\sum_{n=1}^N A_nP_n \quad (13)$$

where the Lagrange multiplier λ represents the slope of a line supporting the deviation-cost function $D(C)$.

In Sections III-B and III-C, we first investigate system designs in the continuous-rate case. With $J(\lambda)$ as in (13), we minimize the Lagrangian for a given multiplier λ, as detailed in Section III-B. In so doing, we find an optimal operating point $(D(\lambda), C(\lambda))$ on the D-C plane [9], [10], such that $D(\lambda)$ is the minimum achievable deviation with a given power budget $C(\lambda)$. To depict the entire $D(C)$ function, we can repeat this procedure for all possible λ's. However, since we are interested only in $D(C = P_0)$, we only need to locate a suitable parameter λ^o, such that $C(\lambda^o) = P_0$, as detailed in Section III-C.

B. Finding the Optimal Pair $(D(\lambda), C(\lambda))$ for a Given λ

To find the optimal operating pair $(D(\lambda), C(\lambda))$ for a given λ, we need to minimize the Lagrangian $J(\lambda)$ in (13). To this end, joint optimization of the fading regions and the transmission modes is required. We will pursue an iterative descent algorithm that operates in two stages: 1) *we find an optimal design of transmission modes given a certain set of fading regions*; and 2) *we find an optimal design of fading regions given a certain design of the transmission modes*. Notice that here the optimality is with regard to minimizing the Lagrangian $J(\lambda)$ in (13).

1) Finding the Optimal $\{\mathcal{M}_n\}_{n=1}^N$, for a Given $\{\mathcal{R}_n\}_{n=1}^N$: Given $\{\mathcal{R}_n\}_{n=1}^N$, the probability A_n and the channel covariance matrix $\mathbf{R}_n$ can be calculated for every n. From (13), we see that beamforming can be decoupled from the power/bit loadings without loss of optimality. Furthermore, the beamforming vectors over different regions are independent. This allows us to derive the beamforming vectors separately.

Let the eigen-decomposition of the channel covariance matrix be denoted as $\mathbf{R}_n = \mathbf{U}_n\mathbf{\Sigma}_n\mathbf{U}_n^{\mathcal{H}}$, where $\mathbf{\Sigma}_n := \mathrm{diag}(\sigma_{n1}, \ldots, \sigma_{nN_t})$ is the diagonal matrix containing eigen-values in a nonincreasing order $\sigma_{n1} \geq \ldots \geq \sigma_{nN_t}$, and $\mathbf{U}_n$ is the unitary matrix containing the corresponding eigen-vectors; i.e., $\mathbf{U}_n := (\mathbf{u}_{n1}, \ldots, \mathbf{u}_{nN_t})$. We see from (11) that the optimal beamforming vector $\mathbf{w}_n^{\mathrm{opt}}$ maximizing $\tilde{b}_n$ in region $\mathcal{R}_n$, is actually the vector that maximizes $\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n$, subject to the constraint (c2). Consequently

$$\mathbf{w}_n^{\mathrm{opt}} = \arg \max_{\mathbf{w}_n^{\mathcal{H}}\mathbf{w}_n=1} \mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n = \mathbf{u}_{n1} \tag{14}$$

and the achieved maximum value of $\mathbf{w}_n^{\mathcal{H}}\mathbf{R}_n\mathbf{w}_n$ is σ_{n1}. In other words, for each region, the optimal beam points along the principal eigen-vector of $\mathbf{R}_n$, which is also known as eigen-beamforming [22], [24], [28]. With optimal eigen-beamforming over each region, minimizing $J(\lambda)$ becomes a standard optimization problem, and the power loading follows the well-known water-filling principle

$$P_n = \left(\frac{1}{\lambda \ln 2} - \frac{N_0 \cdot \beta(-\ln 5\overline{\mathrm{BER}_0})}{1.5\sigma_{n1}}\right)^+ \quad \forall n \tag{15}$$

where $(\cdot)^+$ is the operator defined as $(x)^+ := \max(x, 0)$. Bit-loading solutions can then be determined directly by inserting (15) into (10). Equations (14), (15), and (10) collectively determine the optimal transmission modes $\{\mathcal{M}_n\}_{n=1}^N$ for a given set of fading regions $\{\mathcal{R}_n\}_{n=1}^N$.

2) Finding the Optimal $\{\mathcal{R}_n\}_{n=1}^N$, for a Given $\{\mathcal{M}_n\}_{n=1}^N$: Given fading regions, it is relatively easy to find the optimal design of transmission modes. The converse, however, is more difficult. To deal with this difficulty, let us consider

$$\tilde{J}(\lambda) := \sum_{n=1}^{N} A_n \mathrm{E}_{\mathbf{h}\in\mathcal{R}_n}\left[-\log_2\left(1+\frac{1.5P_n\left|\mathbf{w}_n^{\mathcal{H}}\mathbf{h}\right|^2}{N_0 \cdot \beta(-\ln 5\overline{\mathrm{BER}_0})}\right)\right] + \lambda\sum_{n=1}^{N} A_n P_n \tag{16}$$

which is an upper bound to $J(\lambda)$, due to the convexity of the function $-\log_2(\cdot)$. Minimizing an upperbound of the objective function is a commonly used technique when direct optimization of the objective function itself is difficult. Therefore, in this subsection, we try to minimize the deviation upperbound $\tilde{J}(\lambda)$ instead of $J(\lambda)$ itself. The validity of the approximations will be confirmed numerically.

Optimization of $\tilde{J}(\lambda)$ relates to a vector-quantizer design problem. Consider a size-N vector quantizer Q, which accepts the N_t-dimensional complex channel vectors $\{\mathbf{h}\}$ as inputs, and maps them to a codebook comprising N codewords, which in our case correspond to the transmission modes $\{\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_N\}$. Associated with this vector quantizer Q is a partition of the input-vector space $\mathcal{R}$ into N cells, which are exactly $\{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_N\}$ in our case. Furthermore, this equivalent vector quantizer relies on the function

$$d(\mathbf{h}, \mathcal{M}_n) := -\log_2\left(1+\frac{1.5P_n\left|\mathbf{w}_n^{\mathcal{H}}\mathbf{h}\right|^2}{N_0 \cdot (-\ln 5\overline{\mathrm{BER}_0})}\right) + \lambda P_n \tag{17}$$

to measure the distortion between the input vector $\mathbf{h}$ and the codeword $\mathcal{M}_n$. The objective $\tilde{J}(\lambda)$ can thus be interpreted as the overall average distortion measure associated with Q

$$\tilde{J}(\lambda) = \sum_{n=1}^{N} A_n \mathrm{E}_{\mathbf{h}\in\mathcal{R}_n}\left[d(\mathbf{h}, \mathcal{M}_n)\right] = \mathrm{E}_{\mathbf{h}}\left[d(\mathbf{h}, \mathcal{M}_n)\right]. \tag{18}$$

Designing a vector quantizer amounts to finding a codebook along with a partition rule that minimize the overall average distortion measure (18). Two necessary conditions of optimality prove to be essential for the design of a vector quantizer. First, necessary for the optimality of the codebook, is the so-called *centroid condition*, which specifies that for each region, the optimal codeword should be chosen to minimize the distortion measure averaged over that region. Eigen-beamforming [(14)] together with power waterfilling [(15)] correspond to the centroid condition in our case.

Second, necessary for the optimality of the partition is the *nearest-neighbor rule*, which dictates all input vectors closer (in the smaller distortion measure sense) to the codeword $\mathcal{M}_i$ than to any other codeword, be assigned to region $\mathcal{R}_i$. The nearest neighbor rule, partitions the channel space as follows

$$\mathbf{h} \in \mathcal{R}_i \quad \text{iff} \quad d(\mathbf{h}, \mathcal{M}_i) \leq d(\mathbf{h}, \mathcal{M}_j) \quad \forall i, j = 1, \ldots, N. \tag{19}$$

3) Iterative Descent Algorithm: Generally, it is impossible to optimize transmission modes and fading regions jointly. In the following, we present an iterative descent algorithm that couples these two subtasks repeatedly to enable the final design of the transmission modes and the fading regions.

Iterative descent algorithm

s1) Initialize with any valid set of $\{\mathcal{M}_n, \mathcal{R}_n\}_{n=1}^N$, and calculate $J(\lambda)$ in (13).

s2) For the current transmission modes, find the optimal fading regions minimizing $J(\lambda)$.

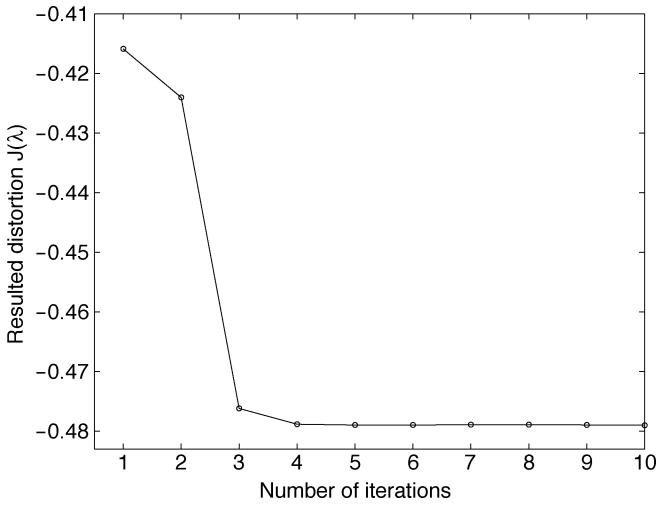


Fig. 2. Convergence of the iterative descent algorithm.

s3) Exit with the current design of transmission modes and fading regions if $J(\lambda)$ converges.
s4) For the current fading regions, find the optimal transmission modes minimizing $\tilde{J}(\lambda)$.
s5) Loop back to s2).

This iterative descent algorithm is similar to the generalized Lloyd algorithm, which is popular in the vector-quantization literature [9], [10]. For a typical vector quantizer, steps s2) and s4) have the same objective $\tilde{J}(\lambda)$. This implies that the overall distortion is reduced or at least remains unchanged after each iteration, and proves that the generalized Lloyd algorithm is guaranteed to converge in a finite number of iterations [10]. The difference here is that $\tilde{J}(\lambda)$ is an upper bound to $J(\lambda)$. Due to the slight objective mismatch in steps s2) and s4), there is no guarantee that our iterative descent algorithm converges in finite number of iterations. To assess convergence, we will rely on simulations. Fig. 2 shows one realization of the achieved distortion $J(\lambda)$ after each iteration, from which it can be verified that the iterative descent algorithm converges fast, typically within five iterations.

Upon completion of this iterative descent algorithm, we have available a joint design of the transmission modes, and the fading regions. Meanwhile, we have obtained an optimal operating point $(D(\lambda), C(\lambda))$ on the D-C plane, meaning that a continuous-rate throughput of $-D(\lambda)$ is achieved with a power budget $C(\lambda)$.

C. Bisection Search to Locate λ^o

In Section III-B, we minimized the Lagrangian $J(\lambda)$ using an iterative descent algorithm for a given parameter λ. A continuous-rate design of fading regions and transmission modes was thus obtained. In the meantime, we obtained a deviation $D(\lambda)$ and a power cost $C(\lambda)$, that determine an optimal pair $(D(\lambda), C(\lambda))$ on the D-C plane. Our goal in this subsection is to quickly find the optimal pair with parameter λ^o, such that $C(\lambda^o) = P_0$.

It can be easily seen that $D(C)$ is monotonically decreasing with C. Furthermore, $D(C)$ is convex. To establish the latter, suppose we have scheme 1 operating at point (D_1, C_1) on the D-C plane, to achieve a minimum deviation of D_1 with power budget C_1; and scheme 2 operating at point (D_2, C_2), to achieve a minimum deviation of D_2 with power budget C_2. When the given power budget is $(C_1 + C_2)/2$, a deviation of $(D_1 + D_2)/2$ can be guaranteed by a balanced time sharing between schemes 1 and 2. Therefore, the optimal scheme with power budget $(C_1 + C_2)/2$ must achieve a minimum deviation less than $(D_1 + D_2)/2$, i.e., $D((C_1 + C_2)/2) \leq (D_1 + D_2)/2$. Since schemes 1 and 2 are chosen arbitrarily, the function $D(C)$ is convex. Motivated by the monotonicity and convexity of $D(C)$, we briefly describe in the following a bisection search algorithm [18] to locate λ^o.

1) Choose a proper Lagrange multiplier λ_{low} and apply the iterative descent algorithm, to obtain $C_{\text{low}} = C(\lambda_{\text{low}})$, and $D_{\text{low}} = D(\lambda_{\text{low}})$. Make sure $C_{\text{low}} < P_0$.
2) Choose a proper Lagrange multiplier λ_{high} and apply the iterative descent algorithm, to obtain $C_{\text{high}} = C(\lambda_{\text{high}})$, and $D_{\text{high}} = D(\lambda_{\text{high}})$. Make sure $C_{\text{high}} > P_0$.
3) Update the Lagrange multiplier using: $\lambda_{\text{new}} = (D_{\text{high}} - D_{\text{low}})/(C_{\text{high}} - C_{\text{low}})$.
4) Apply the iterative descent algorithm to obtain $D_{\text{new}} = D(\lambda_{\text{new}})$, and $C_{\text{new}} = C(\lambda_{\text{new}})$.
5) If $|C_{\text{new}} - P_0| \leq \epsilon$ with $\epsilon \to 0$, then exit; If $C_{\text{new}} > P_0$, then set $\lambda_{\text{high}} = \lambda_{\text{new}}$, $C_{\text{high}} = C_{\text{new}}$, $D_{\text{high}} = D_{\text{new}}$; Else, set $\lambda_{\text{low}} = \lambda_{\text{new}}$, $C_{\text{low}} = C_{\text{new}}$, $D_{\text{low}} = D_{\text{new}}$.
6) Loop back to step 3).

Basically, we start the search knowing that the desired λ^o lies inside the interval $(\lambda_{\text{low}}, \lambda_{\text{high}})$. We proceed by generating a λ_{new} inside $(\lambda_{\text{low}}, \lambda_{\text{high}})$, and test in which of the subintervals, $(\lambda_{\text{low}}, \lambda_{\text{new}})$ or $(\lambda_{\text{new}}, \lambda_{\text{high}})$, the desired λ^o lies. With the interval $(\lambda_{\text{low}}, \lambda_{\text{high}})$ properly updated, this procedure is applied iteratively to yield the final result λ^o.

In summary, to maximize the data throughput in (11) with continuous-rate system designs, we adopt a nested iterative approach, as shown in the flow chart of Fig. 3. Two loops operate jointly to yield final designs of the transmission modes and the fading regions. Specifically, the bisection search algorithm operates in the outer loop to locate a proper parameter λ^o, while the iterative descent algorithm operates in the inner loop to actually find the transmission modes and the fading regions. We stress that the entire design procedure, although complicated, will be performed offline before the system operation starts. In the operational mode, the receiver only needs to assign the estimated channel to the region it belongs to, as in (19). This complexity increases only linearly with N. Since N is of moderate value in practice, real-time operation is certainly affordable.

D. The Discrete-Rate Case

Up till now, we have been focusing on the continuous-rate scenario. Although system designs for the continuous-rate scenario provide intuition and useful guidelines, the associated nonsquare M-QAMs require high implementation complexity [11]. For this reason, practical system designs typically opt for discrete-rate signal constellations.

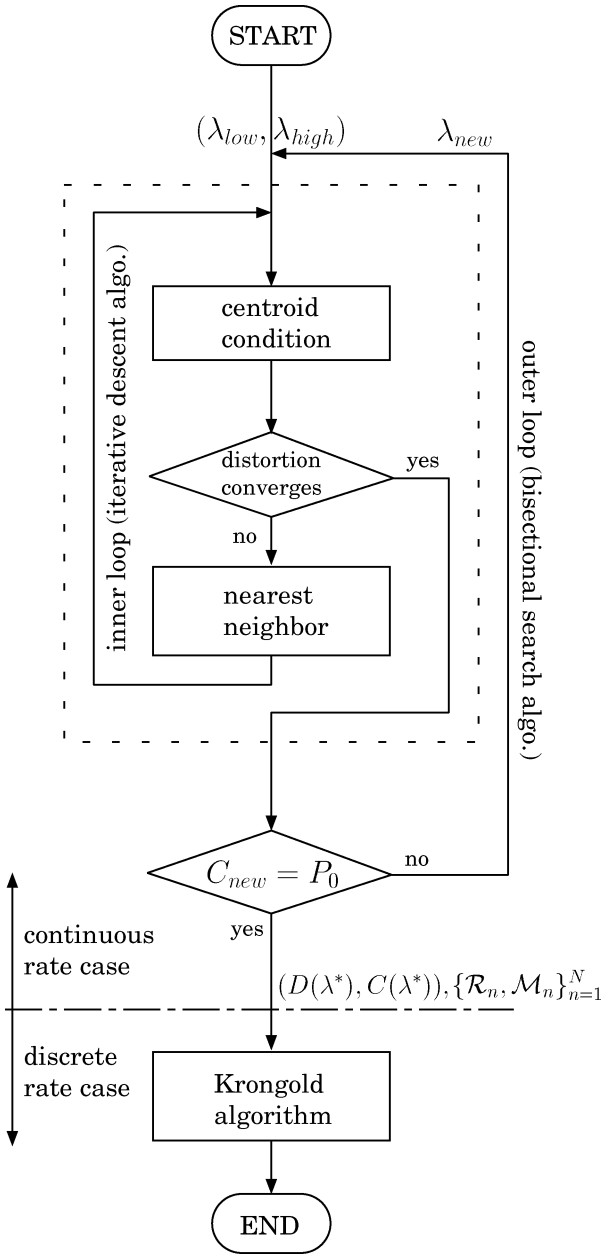


Fig. 3. Algorithm flow chart.

In the discrete-rate case, however, even directly optimizing the suboptimal objective in (11) does not lead to a meaningful solution. Considering the design resulting from the nested iterative algorithm for the continuous-rate case, we see that it is the design of transmission modes that violates the discrete-rate constraint. This suggests retaining the design of the fading regions from the continuous-rate case, and continuing to pursue the design of transmission modes that meet the discrete-rate requirement. In other words, we wish to

$$\begin{aligned} &\max . \sum_{n=1}^{N} A_n \log_2\left(1+\frac{1.5P_n\sigma_{n1}}{N_0\cdot\beta(-\ln 5\overline{\mathrm{BER}}_0)}\right) \\ &\text{subject to (c1)} \sum_{n=1}^{N} A_n P_n \le P_0 \text{ and} \\ &\qquad \text{(c3)}\ \tilde{b}_n \in \{0,2,4,\ldots\} \quad \forall n \end{aligned} \tag{20}$$

where, different from (11), A_n's are already known here. If all regions are chosen with the same probability, then the problem formulation in (20) is identical (up to a multiplicative constant) to the power and bit loading in multicarrier OFDM systems with the integer bit constraint [18], where all OFDM carriers appear with the same probability 1. It can be easily verified that the algorithm in [18] can be extended to our case with nonuniform A_n across different transmission modes. We refer the readers to [18] for the algorithm development.

We underscore that with β chosen to meet (c4) in the continuous-rate case, the power- and bit-loading solutions in the discrete-rate case will also guarantee (c4), since the product $P_n g_n$ remains the same in both cases, based on (10), leading to the same actual BER in (4).

The resulting power/bit loadings, together with the eigenbeamforming in (14), constitute a practical design of transmission modes under the discrete-rate constraint. Combined with the design of fading regions inherited from the continuous-rate scenario, the discrete-rate design of fading regions and transmission modes is, therefore, complete.

E. Implementation Issues

1) Training-Set-Based Implementation: Notice that the iterative descent algorithm requires exact specification of the fading regions' geometry, which is not easy in practice. Furthermore, computation of $\{A_n, \mathbf{R}_n\}_{n=1}^N$ involves multiple integrals over regions of arbitrary shapes, and is generally intractable analytically. For practical implementation purposes, we will use a training set. Specifically, let $\mathcal{H}_s := \{\mathbf{h}_{ks}\}_{k=1}^K$ be a training set, where each $\mathbf{h}_{ks}$ is a sample vector generated according to the true distribution $p(\mathbf{h})$, where $K \gg N$ is the size of the training set. Instead of the true channel distribution $p(\mathbf{h})$, we use the sample distribution based on the training set

$$p_s(\mathbf{h}) := \frac{1}{K}\sum_{k=1}^{K} \delta(\mathbf{h}-\mathbf{h}_{ks}) \tag{21}$$

where $\delta(\cdot)$ is the Dirac delta function. This is actually a Monte Carlo method to evaluate the integrals involved, and has become a standard approach in vector-quantizer designs. The strong law of large numbers ensures that as the training-set size K increases, the sample distribution will converge with probability 1 to the actual distribution [10, p. 364]. Specifying region geometries boils down to classifying the training set, and the complicated integrals are replaced by simple sums. Furthermore, we can see from (21) that the transceiver design applies equally well for correlated and even non-Gaussian channels, as long as a statistical description of the channel vectors is available to generate the training set. We stress that the training set is required only in the transceiver-design phase, but is not necessary during the data-transmission phase.

2) Outage Region: We find that the implementation of our suboptimal algorithm can be improved by explicitly setting one region to be an outage region, over which no data transmission occurs. Specifically, we introduce an auxiliary region

$$\mathcal{M}_N := \{P_N = 0, b_N = 0, \mathbf{w}_N \text{ arbitrary}\}. \tag{22}$$

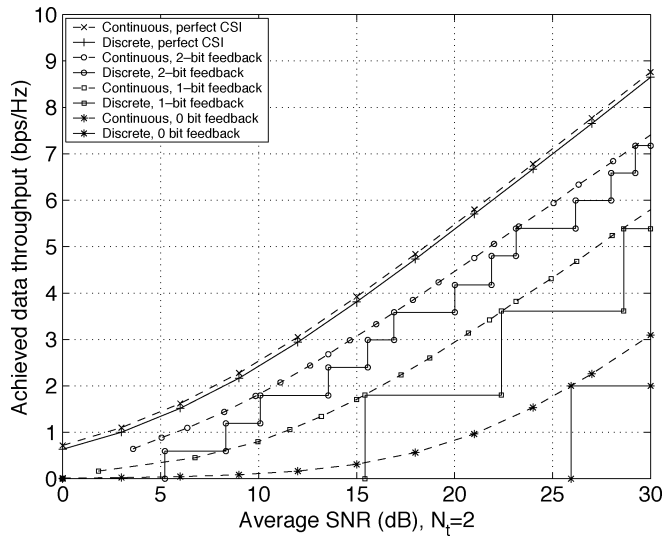


Fig. 4. Achievable data throughput, $N_t = 2$. $\beta = 4.0$, $\rho_0 = 0.33$ (1 b), $\beta = 1.90$, $\rho_0 = 0.66$ (2 b, average SNR >= 10 dB), $\beta = 1.94$, $\rho_0 = 0.66$ (2 b, average SNR< 10 dB).

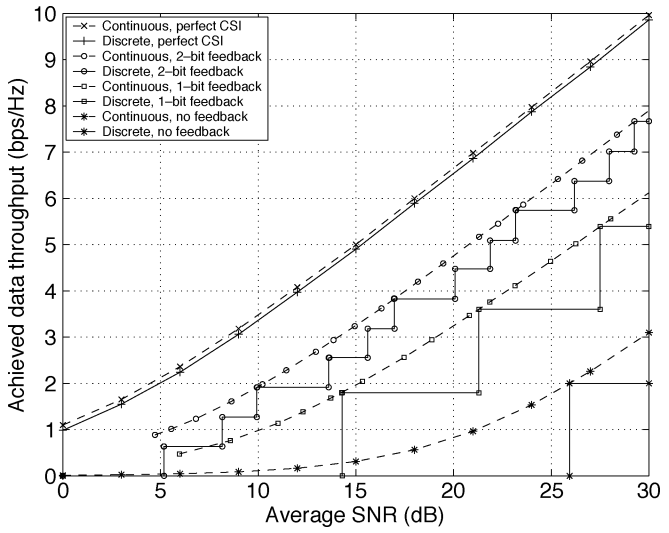


Fig. 5. Achievable data throughput, $N_t = 4$. $\beta = 3.2$, $\rho_0 = 0.33$ (1 b), $\beta = 1.9$, $\rho_0 = 0.66$ (2 b).

This outage region is well motivated from the ideal case with perfect channel knowledge. Even with perfect CSI, data transmissions have to be suspended for channels with small amplitudes, in order to improve the overall spectral efficiency [8], [11]. This insight turns out to be very useful, as we find that introducing the outage region $\mathcal{M}_n$ leads to a higher data throughput relative to the alternative implementation without specifying the outage region. The reason is that the required BER precompensation factor β is now smaller because "bad channels" in the outage region are excluded from the average BER (which is usually dominated by the worst channel realizations).

Taking into account the outage region (22), we modify our nearest-neighbor rule in (19) as

$$\begin{aligned} &\mathbf{h} \in \mathcal{R}_i \quad \text{iff} \quad \left|\mathbf{w}_i^{\mathcal{H}}\mathbf{h}\right|^2 > \rho_0 \text{ and} \\ &\quad d(\mathbf{h}, \mathcal{M}_i) \le d(\mathbf{h}, \mathcal{M}_j), \ \forall i, j = 1, \ldots, N-1 \\ &\mathbf{h} \in \mathcal{R}_N \quad \text{iff} \quad \left|\mathbf{w}_i^{\mathcal{H}}\mathbf{h}\right|^2 \le \rho_0, \ \forall i = 1, \ldots, N-1 \end{aligned} \tag{23}$$

where ρ_0 is a tuning parameter. We use empirical values for ρ_0 in our numerical tests in Section IV. Our nested algorithm is then implemented with a mandate $\mathcal{M}_N$ in (22), and a modified nearest-neighbor rule in (23).

IV. Numerical Results

We set $\sigma_h^2 = 1$, $\overline{\text{BER}}_0 = 10^{-3}$, $N_0 = 0.1$, and define the transmit SNR as P_0/N_0. If $\sigma_h^2 \neq 1$, we just need to absorb σ_h^2 into the transmission power P_0; the codebook design in Section III depends only on the average received SNR $P_0\sigma_h^2/N_0$. The training set size is chosen as 25 600. In each test case, the parameters β, ρ_0 are selected empirically to meet the BER constraint (c4), and will be reported along with the plots.

Test Case 1—Achieved Data Throughput: For $N_t = 2$ and 4, respectively, Figs. 4 and 5 display the simulated data throughput with a variable number of feedback bits. The throughput upper and lower bounds from Section II-C are also plotted. Comparing the data throughput achieved by continuous-rate designs with their corresponding upper and lower bounds, we see that a limited number of feedback bits can considerably improve the overall system throughput. The same holds for discrete-rate designs. The improvement is most significant for the first feedback bit, while the extra data throughput improvement decreases as the number of feedback bits increases. A tradeoff between the achievable data throughput and the number of feedback bits therefore emerges. In practice, only a small number of feedback bits (say three or four per block) might be favored.

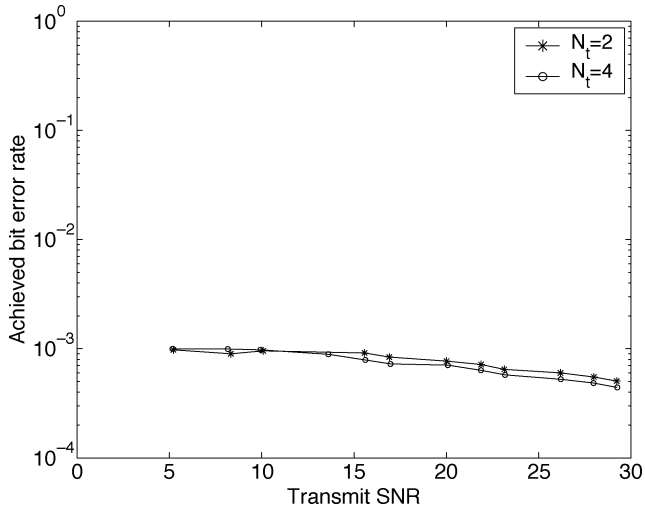


Fig. 6. Achieved BER.

Due to the discrete-rate constraint, an SNR-dependent data throughput gap always exists between a discrete design and its continuous counterpart. However, this throughput gap shrinks very quickly as the number of feedback bits B increases (see, e.g., Fig. 4). When B is large enough, the gap stabilizes and becomes uniform across the entire SNR range. This justifies our investigation of the continuous-rate case despite its practical limitations. The exact average BER performance achieved by the discrete-rate design is displayed in Fig. 6, which demonstrates that the BER constraint is indeed satisfied.

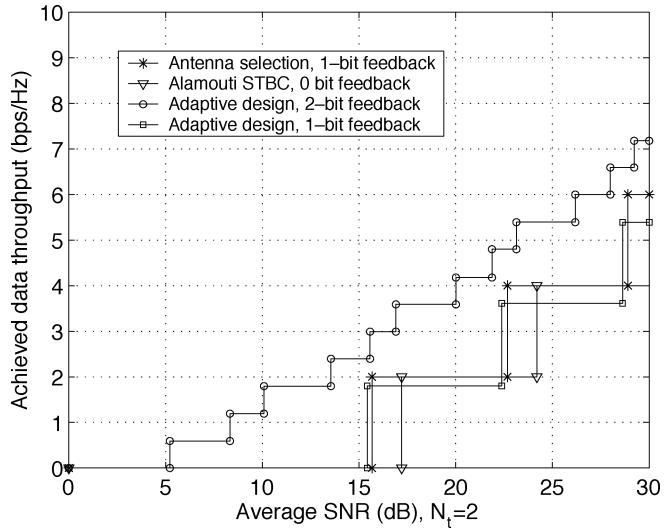


Fig. 7. Comparison with existing schemes, $N_t = 2$.

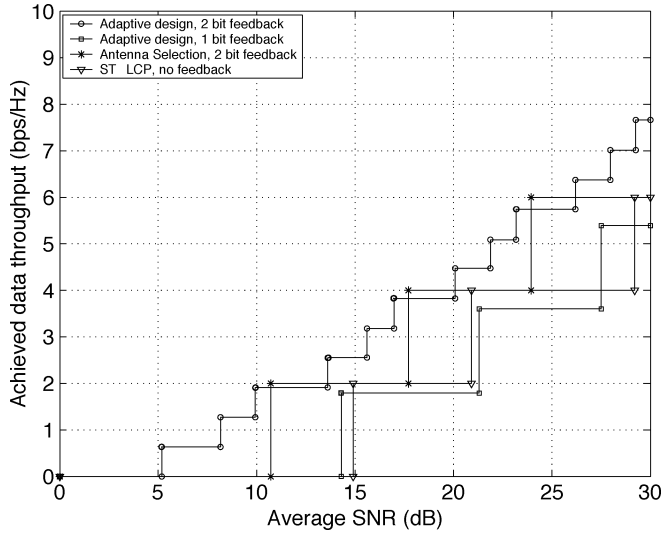


Fig. 8. Comparison with existing schemes, $N_t = 4$.

We next carry out comparisons with certain existing transmission strategies in terms of the achieved data throughput under the same average BER constraint, where only discrete-rate designs are considered. Fig. 7 with $N_t = 2$ and Fig. 8 with $N_t = 4$ compare the proposed adaptive design with the space–time linear constellation precoding (ST-LCP) scheme of [27], which requires no feedback bit, and with the strongest antenna-selection scheme [7], which requires two feedback bits. As we see, when only one feedback bit is available, the adaptive design normally does not provide much gain over the STBC. However, when two or more bits are used, the adaptive design typically outperforms nonadaptive alternatives. Furthermore, with the same number of feedback bits, the proposed adaptive scheme either achieves comparable data throughput $(N_t = 2)$, or outperforms the antenna-selection scheme $(N_t = 4)$.

For comparison purposes, we also study some heuristic design examples of separate modulation and beamforming adaptation, as a systematic design approach is not available yet. Specifically, we investigate a multiple-input, single-output (MISO) system example with $N_t = 2$ transmit antennas and $B = 2$ feedback bits per block. Suppose one feedback bit a_1 is used to index modulation types, while the other feedback bit a_2 is used to select the stronger antenna from the two. For this special case, the pdf of the equivalent SISO channel can be written in a closed form by using order statistics. For the two candidate modulation types indexed by a_1, we consider three alternatives: 1) no transmission and 4-QAM modulation; 2) no transmission and 16-QAM modulation; and 3) no transmission and 64-QAM modulation. For example, for the last alternative, $a_1 = 0$ may stand for no transmission, while $a_1 = 1$ may signal 64-QAM modulation. The modulation-adaptation policy on top of the equivalent SISO channel is then obtained through numerical search, following the design approach in [11]. The achieved data throughput for each of the three alternatives is plotted in Fig. 9, and compared with the joint-adaptation scheme. We see that the joint-adaptation schemes generally provide higher data throughput, although at the cost of higher design complexity.

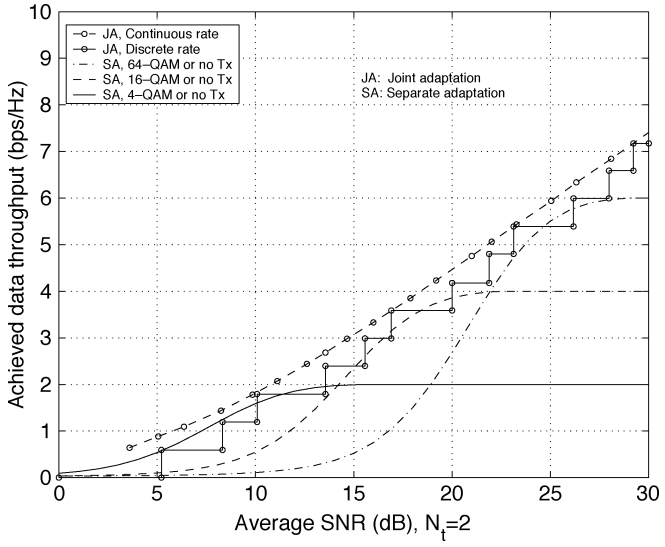


Fig. 9. Comparison with separate adaptations, $N_t = 2, B = 2$.

Notice that the separate design could also make SNR-dependent choices: it can use alternative 1 in the low-SNR range, alternative 2 in the medium-SNR range, and alternative 3 in the high-SNR range. In such a scenario, the joint design has a performance edge over the separate design only in the range of medium-to-high SNRs, and becomes inferior at low SNR (due to the approximations made in the development).

Test Case 2—Typical Channel Space Partition: We finally consider a two-input single-output (TISO) system with $N = 16$ and SNR $= 20$ dB. The channel space in such a case is four-dimensional with two complex channel coefficients, which is impossible to plot directly. We therefore consider plotting only a subspace of the entire channel space. We choose the x-axis to be the relative phase defined as $\theta := \measuredangle h_1 - \measuredangle h_2$ ranging from $-\pi$ to π, and the y-axis to be the relative amplitude defined as $\phi := \tan^{-1}(|h_1|/|h_2|)$ ranging from 0 to $\pi/2$. Fig. 10 shows the $(N-1)$ normal regions partitioning the entire subspace. For every region, the small block inside represents the beamforming vector. Generally, each region has an irregular shape. This corroborates our usage of the channel sample distribution based on the training set, as discussed in Section III-E.1.

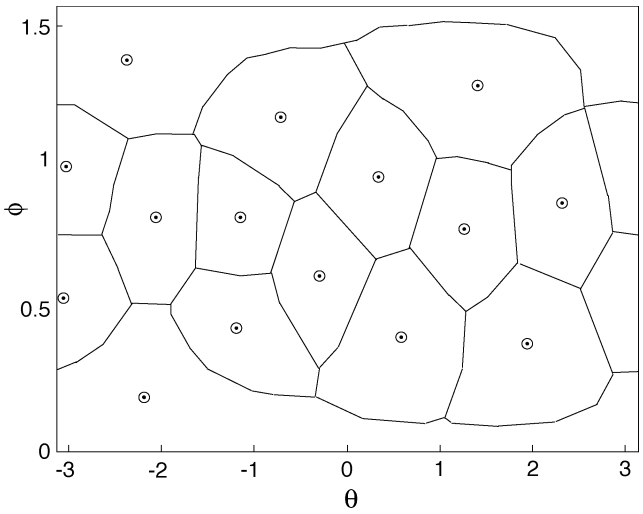


Fig. 10. A display of the normal regions: $N_t = 2$, $N = 16$.

V. Conclusions

In this paper, we investigated a multiantenna system with adaptive modulation and transmit beamforming based on rate-limited feedback. Subject to BER and transmit-power constraints, our adaptive system design was carried out by jointly adjusting the transmit power, the signal constellation, the beamforming direction and the feedback strategy according to the available finite number of feedback bits. For the continuous-rate case, a nested iterative approach was developed to yield a suboptimal yet practical system design. A discrete-rate system was designed based on the continuous-rate system. Numerical results demonstrated increased data throughput confirming that a finite number of feedback bits can improve the overall performance considerably.

References

[1] *Physical Channels and Mapping of Transport Channels Onto Physical Channels*, 3GPP TS 25.211 V5.1.0, 2002.

[2] *Physical Layer Procedures*, 3GPP TS 25.214 V5.1.0, 2002.

[3] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[4] M.-S. Alouini and A. J. Goldsmith, "Adaptive modulation over Nakagami fading channels," *J. Wireless Commun.*, vol. 13, pp. 119–143, May 2000.

[5] S. Bhashyam, A. Sabharwal, and B. A. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Trans. Commun.*, vol. 50, pp. 785–798, May 2002.

[6] R. Blum, "MIMO with limited feedback of channel state information," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, vol. 4, Hong Kong, China, Apr. 2003, pp. 89–92.

[7] R. Blum and J. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, pp. 322–324, Aug. 2002.

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

[9] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, pp. 31–42, Jan. 1989.

[10] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.

[11] 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.

[12] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA and OFDM Systems*. New York: Wiley, 2002.

[13] K. J. Hole, H. Holm, and G. E. Øien, "Adaptive multidimensional coded modulation over flat fading channels," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 1153–1158, Jul. 2000.

[14] 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 Telecommun. Conf.*, vol. 2, San Antonio, TX, Nov. 2001, pp. 1256–1261.

[15] S. A. Jafar and A. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. IEEE Int. Symp. Inf. Theory*, Washington, DC, Jun. 2001, p. 321.

[16] G. Jöngren and M. Skoglund, "Utilizing quantized feedback information in orthogonal space-time block coding," in *Proc. Global Telecommun. Conf.*, vol. 2, San Francisco, CA, Nov.–Dec. 2000, pp. 995–999.

[17] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inf. Theory*, vol. 48, pp. 611–627, Mar. 2002.

[18] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Trans. Commun.*, vol. 48, pp. 23–27, Jan. 2000.

[19] V. K. N. Lau, Y. Liu, and T.-A. Chen, "On the design of MIMO block-fading channels with feedback-link capacity constraint," *IEEE Trans. Commun.*, vol. 52, pp. 62–70, Jan. 2004.

[20] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2735–2747, Oct. 2003.

[21] K. Mukkavilli, A. Sabharwal, E. Erkip, and B. A. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2562–2579, Oct. 2003.

[22] 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. Sel. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[23] I. Telatar, "Capacity of multi-antenna Gaussian channels," *Bell Labs Tech Memo.*, 1995.

[24] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[25] P. Xia and G. B. Giannakis, "Design and analysis of transmit beamforming based on limited-rate feedback," in *Proc. IEEE Veh. Technol. Conf.*, Los Angeles, CA, Sep. 2004.

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

[27] Y. Xin, Z. Wang, and G. B. Giannakis, "Space-time diversity systems based on linear constellation precoding," *IEEE Trans. Wireless Commun.*, vol. 2, pp. 294–309, Mar. 2003.

[28] S. Zhou and G. B. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Signal Process.*, vol. 50, pp. 2599–2613, Oct. 2002.

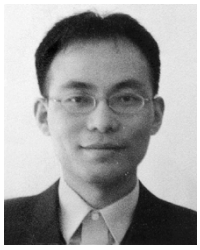
[29] ——, "Adaptive modulation for multi-antenna transmissions with channel mean feedback," *IEEE Trans. Wireless Commun.*, to be published.

[30] S. Zhou, Z. Wang, and G. B. Giannakis, "Quantifying the power loss when transmit-beamforming relies on finite rate feedback," *IEEE Trans. Wireless Commun.*, to be published.

[31] P. Xia, S. Zhou, and G. B. Giannakis, "Achieving the Welch bound with difference sets," *IEEE Trans. Inf. Theory*, to be published.

Pengfei Xia (S'03) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 1997 and 2000, respectively. He is now working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis.

His research interest lies in the area of communications and signal processing, including multi-input multi-output wireless communications, space–time coding, diversity techniques, multicarrier transmissions, adaptive modulation and coding, quantization, rate-distortion theory, and turbo codes.



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

He joined the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, as an Assistant Professor in 2003. His research interests lie in 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. He is currently an Associate Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.

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

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 with the University of Minnesota, Minneapolis, as a Professor in the Department of Electrical and Computer Engineering, and 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. Current research focuses on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, complex-field and space–time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and distributed sensor networks.

Dr. Giannakis is the corecipient of six Best Paper Awards from the IEEE Signal Processing (SP) Society (1992, 1998, 2000, 2001, 2003, and 2004). He also received the Society's Technical Achievement Award in 2000. He has served as Editor in Chief for the IEEE SIGNAL PROCESSING L ETTERS, 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 Election Committee. He has served on the IEEE-SP Society's Board of Governors, and served as a member of the Editorial Board for the PROCEEDINGS OF THE IEEE and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.