How Accurate Channel Prediction Needs to be for Transmit-Beamforming With Adaptive Modulation Over Rayleigh MIMO Channels?

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

***Abstract*—Adaptive modulation improves the system throughput considerably by matching transmitter parameters to time-varying wireless fading channels. Crucial to adaptive modulation is the quality of channel state information at the transmitter. In this paper, we first present a channel predictor based on pilot symbol assisted modulation for multiple-input multiple-output Rayleigh fading channels. We then analyze the impact of the channel prediction error on the bit error rate performance of a transmit-beamformer with adaptive modulation that treats the predicted channels as perfect. Our numerical results reveal the critical value of the normalized prediction error, below which the predicted channels can be treated as perfect by the adaptive modulator; otherwise, explicit consideration of the channel imperfection must be accounted for at the transmitter.**

***Index Terms*—Adaptive modulation, channel prediction, multiantenna systems, transmit beamforming.**

I. Introduction

BY MATCHING transmitter parameters to time-varying channel conditions, adaptive modulation increases the system throughput considerably, which justifies its popularity in future high-rate wireless applications; see [4], [8], [10]–[14], [19], [26], [30], [32], and references therein. Critical to adaptive modulation is the quality of channel state information (CSI) at the transmitter, that is obtained through feedback. Due to the transmission delay and the processing delay both at the transmitter and at the receiver, the delayed CSI feedback at the transmitter becomes outdated, unless the channel variations are sufficiently slow. Taking into account the feedback delay, an effective approach in adaptive systems is to predict the channel values at future times when they will be used, and feed those predicted channels back to the transmitter [9].

Adaptive designs assuming perfect CSI perform well only when CSI imperfections induced by channel estimation errors and/or feedback delays are limited [4], [12]. For general Nakagami fading channels, the bit-error rate (BER) performance was analyzed in [4] for single-antenna systems with delayed but noiseless channel estimates. For Rayleigh fading channels, BER performance analysis was carried out in [24] and [25] for systems equipped with single transmit- and multiple-receive antennas, based on noisy predicted channels.

In this paper, we investigate an adaptive system equipped with multiple transmit and multiple receive antennas, where each information symbol is sent across multiple transmit-antennas using a transmit-beamformer. We first present a minimum-mean-square-error (MMSE) channel predictor based on pilot symbol assisted modulation (PSAM) [6], for independent and identically distributed (i.i.d.) Rayleigh fading multiple-input multiple-output (MIMO) channels. We then analyze the impact of channel prediction error on the BER performance of adaptive modulation systems based on transmit-beamforming. With an arbitrary number of transmit- and receive- antennas, we obtain a closed-form BER expression that requires integration. We also derive simple closed-form expressions, when the minimum number of transmit- and/or receive- antennas is less than or equal to two.

We establish that BER performance of the adaptive transmit-beamforming based system depends on various system parameters only through a single variable, namely the normalized channel prediction error. Our numerical results reveal that when the normalized prediction error is below a certain critical value, the predicted channels can be treated as perfect by the adaptive modulator. Otherwise, explicit consideration of the channel imperfection is required at the transmitter. Adaptive designs incorporating CSI imperfections have been pursued in [10], [26] for single-antenna systems, and in [35] for multiantenna systems, where the well-known Alamouti's space time block code [3] is incorporated into a two-dimensional, instead of the conventional one-dimensional, transmit-beamformer to benefit from diversity.

The rest of this paper is organized as follows. Section II presents the system model, and derives the MMSE channel predictor for MIMO Rayleigh fading channels based on PSAM. Section III details the adaptive multiantenna transmitter design assuming perfect CSI. Section IV analyzes the BER performance in the presence of channel prediction error. Numerical results are collected in Section V, and conclusions are drawn in Section VI.

Notation: Bold uppercase (lowercase) letters denote matrices (column vectors); $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^{\mathcal{H}}$ denote conjugate, transpose, and Hermitian transpose, respectively; $\mathrm{E}\{\cdot\}$ stands

Manuscript received September 17, 2002; revised February 24, 2003. This work was supported in part by the National Science Foundation (NSF) under Grant 0105612 and in part by the ARL/CTA under Grant DAAD19-01-2-011. The material in this paper was presented in part in the International Conference on Acoustics, Speech, and Signal Processing, Hong Kong, April 6–10, 2003. The associate editor coordinating the review of this paper and approving it for publication was Dr. H. Xu.

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).

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.2004.830842

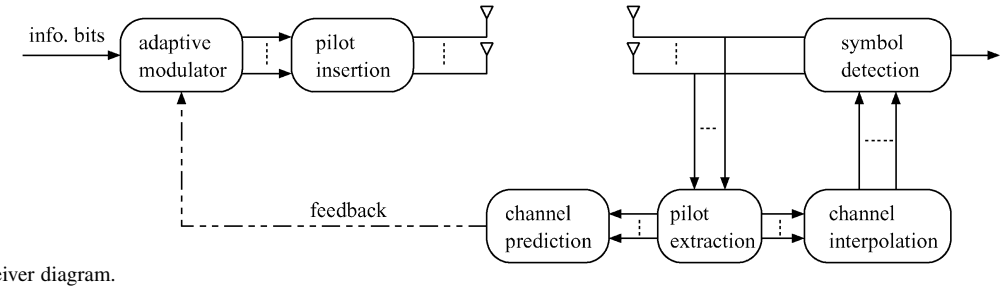

Fig. 1. The transceiver diagram.

for expectation, and $\delta(\cdot)$ for Kronecker's delta; $\mathbf{I}_K$ denotes the identity matrix of size K; $\mathbf{0}_{K\times P}$ denotes an all-zero matrix of size $K \times P$; $[\cdot]_p$ denotes the pth entry of a vector; and, $[\cdot]_{p,q}$ denotes the (p,q)th entry of a matrix. 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. System Modeling and Channel Prediction

A. System Modeling

We consider an adaptive system equipped with multiple transmit- and receive- antennas, as depicted in Fig. 1. Based on CSI obtained from the feedback channel, the transmitter optimally varies its modulation parameters, as will be detailed in Section III. To assist the receiver in performing channel estimation and symbol detection, known pilot symbols are periodically inserted at the transmitter—a technique that is known as pilot symbol assisted modulation (PSAM) [6]. At the receiver, the samples corresponding to the known pilots are extracted, based on which CSI is interpolated using optimal Wiener filtering [6]. Coherent detection is then performed for symbol demodulation.

To enable adaptive modulation, the receiver feeds the estimated CSI back to the transmitter. As is commonly assumed in, e.g., [10]–[12], we suppose that the feedback channel is error-free, which can be ensured by coding and the auto-repeat-request (ARQ) protocol. However, we account for the feedback delay, which includes not only the transmission delay, but also the processing delay both at the receiver and at the transmitter. To take into account the overall feedback delay, the receiver predicts the channels at a future time, and feeds the predicted channels back to the transmitter [9], [24].

We next specify our system model analytically. Let N_t denote the number of transmit antennas, and N_r the number of receive antennas. The spatial channels are assumed to be frequency-flat, and slowly time varying. Let $h_{\mu\nu}(n)$ denote the channel between the μth transmit- and the νth receive-antennae, at time index n. We collect channel coefficients into the $N_t \times N_r$ channel matrix $\mathbf{H}(n)$ having (μ,ν)th entry $h_{\mu\nu}(n)$. Corresponding to each receive antenna ν, we also define the channel vector

$$\mathbf{h}_\nu(n) := [h_{1\nu}(n), \ldots, h_{N_t\nu}(n)]^T.$$

Let $x_\mu(n)$ denote the transmitted symbol from the μth antenna at time n. The received signal at the νth receive antenna can then be expressed as

$$y_\nu(n) = \sum_{\mu=1}^{N_t} h_{\mu\nu}(n)x_\mu(n) + w_\nu(n), \quad \nu \in [1, N_r] \qquad (1)$$

where $w_\nu(n)$ denotes zero-mean additive Gaussian noise.

We adopt the following assumptions throughout the paper.

- AS0): The channels $\{h_{\mu\nu}(n)\}_{\mu=1,\nu=1}^{N_t,N_r}$ are independent and identically distributed (i.i.d.) with Gaussian distribution $\mathcal{CN}(0,1)$; hence, $\mathbf{H}(n) \sim \mathcal{CN}(\mathbf{0}_{N_t\times N_r}, N_r\mathbf{I}_{N_t})$.
- AS1): The channels $\{h_{\mu\nu}(n)\}_{\mu=1,\nu=1}^{N_t,N_r}$ are slowly time-varying according to Jakes' model with Doppler spread f_d; thus, we have $\mathrm{E}\{h^*_{\mu\nu}(n)h_{\mu\nu}(n')\} = J_0(2\pi f_d|n-n'|T_s), \forall \mu,\nu$, where T_s is the symbol period, and $J_0(\cdot)$ is the zeroth order Bessel function of the first kind.
- AS2): The additive Gaussian noise is white both in space and time; i.e., $\mathrm{E}\{w^*_\nu(n)w_{\nu'}(n')\} = N_0\delta(\nu-\nu')\delta(n-n')$.

Based on (1) and AS0)–AS2), we will specify our PSAM-based multiantenna transmissions, and develop the corresponding MMSE predictor for the MIMO channel.

B. MIMO Channel Prediction Based on PSAM

PSAM has well-documented merits as a fading counter-measure for continuously time-varying channels [6], [23]. Under the general framework of [9], PSAM-based channel prediction was proposed in [24] for single transmit-antenna systems. We here extend the PSAM-based channel prediction to adaptive systems with multiple transmit-antennas. We should point out that our purpose here is to provide a means of making (even imperfect) CSI available to the transmitter. Other issues such as PSAM optimality, or, optimization on PSAM parameters, are beyond the scope of this paper; for such issues, the interested reader is referred to [23].

In the single antenna case [6], the data stream is parsed into blocks of length L_b, and one known symbol is inserted per block to estimate one channel coefficient. In multiantenna transmissions, however, the number of unknown channel parameters increases, and the signals from multiple transmit-antennas are superimposed at the receiver. To estimate multiantenna channels based on PSAM, we spread the inserted known symbol in the ith block, call it $\bar{s}_\mu(i)$, by an antenna-specific signature code $\mathbf{c}_\mu := [c_\mu(0), \ldots, c_\mu(N_t-1)]^T$, over N_t symbol periods. Specifically, at time index $n = iL_\mathrm{b} + l$, where $l = 0, 1, \ldots, N_t - 1$, we transmit known symbols

$$x_\mu(n) = x_\mu(iL_\mathrm{b} + l) = \bar{s}_\mu(i)c_\mu(l)$$

as shown in Fig. 2. The signature codes of length N_t are designed to be orthogonal to each other: $\mathbf{c}^{\mathcal{H}}_\mu\mathbf{c}_{\mu'} = \delta(\mu-\mu')$. This decouples MIMO channel estimation and prediction into N_t single transmit-antenna problems. As N_t time slots are used

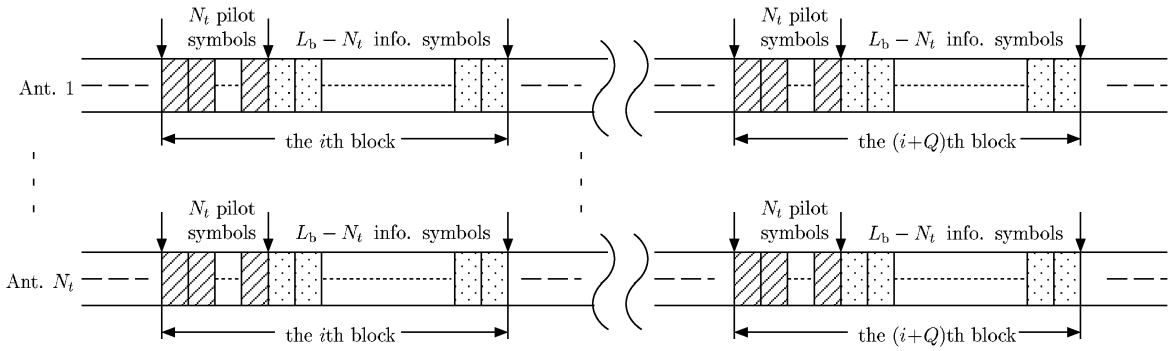

Fig. 2. The PSAM structure for multiantenna transmissions.

to transmit pilot symbols in each block of length L_b, the rate loss for acquisition of MIMO channels is N_t/L_b.

In adaptive systems, it is commonly assumed that the channels remain constant for tens to hundreds of symbols [12]. For such relatively slow varying channels, we have $h_{\mu\nu}(iL_b) = h_{\mu\nu}(iL_b + l)$, for $l = 1, \ldots, N_t - 1$. We will perform channel estimation separately on each receive antenna. To this end, at the νth receive-antenna, we collect N_t received samples corresponding to the pilots of the ith block into the $N_t \times 1$ vector: $\mathbf{y}_\nu(i) = [y_\nu(iL_b), \ldots, y_\nu(iL_b + N_t - 1)]^T$. Using (1), we have

$$\mathbf{y}_\nu(i) = \sum_{\mu=1}^{N_t} \mathbf{c}_\mu \bar{s}_\mu(i) h_{\mu\nu}(iL_b) + \mathbf{w}_\nu(i) \tag{2}$$

where $\mathbf{w}_\nu(i)$ is defined similar to $\mathbf{y}_\nu(i)$. The orthogonality among signature codes $\{\mathbf{c}_\mu\}_{\mu=1}^{N_t}$ enables separation, which allows us to acquire the channels from different antennas as

$$\begin{aligned} \bar{h}_{\mu\nu}(iL_b) &= \frac{1}{\bar{s}_\mu(i)} \mathbf{c}_\mu^{\mathcal{H}} \mathbf{y}_\nu(i) \\ &= h_{\mu\nu}(iL_b) + \tilde{w}_{\mu\nu}(iL_b) \end{aligned} \tag{3}$$

where the resulting noise is $\tilde{w}(iL_b) := \mathbf{c}_\mu^{\mathcal{H}} \mathbf{w}_\nu(i)/\bar{s}_\mu(i)$. With $E_p := |\bar{s}_\mu(i)|^2$ denoting the energy per pilot symbol, the noise $\tilde{w}_{\mu\nu}(iL_b)$ has zero mean, and variance N_0/E_p as per AS2).

Having obtained the initial channel estimates $\bar{h}_{\mu\nu}(iL_b)$ for each block, the receiver then obtains improved channel estimates $\hat{h}_{\mu\nu}(n)$, for any n, using optimal Wiener filtering that exploits the time-domain correlation present according to AS1). We refer the reader to [6] for further details on channel interpolation.

To facilitate adaptive modulation, the receiver needs also to predict channel values ahead of time. We further assume that the transmitter only adapts its parameters once per block (the modulation switching for data transmission is prepared during the pilot interval). Hence, the feedback delay is a multiple of the block duration $L_b T_s$. Suppose that at block i, the receiver predicts the MIMO channel values Q blocks ahead, using a Pth-order filter as

$$\begin{aligned} \hat{h}_{\mu\nu}((i+Q)L_b) &= \sum_{p=0}^{P} w_{\mu\nu}^*(p) \bar{h}_{\mu\nu}((i-p)L_b) \\ &= \mathbf{w}_{\mu\nu}^{\mathcal{H}} \bar{\boldsymbol{h}}_{\mu\nu}(i) \end{aligned} \tag{4}$$

where

$$\begin{aligned} \mathbf{w}_{\mu\nu} &:= [w_{\mu\nu}(0), \ldots, w_{\mu\nu}(P)]^T \\ \bar{\boldsymbol{h}}_{\mu\nu}(i) &:= [\bar{h}_{\mu\nu}(iL_b), \ldots, \bar{h}_{\mu\nu}((i-P)L_b)]^T. \end{aligned}$$

Based on AS1), (4) indicates that $\hat{h}_{\mu\nu}((i+Q)L_b)$ is also a zero-mean Gaussian random variable. We define the correlation coefficient between the true and the predicted channel as

$$\rho_{\mu\nu} := \frac{\mathrm{E}\{h_{\mu\nu}^*((i+Q)L_b)\hat{h}_{\mu\nu}((i+Q)L_b)\}}{\sqrt{\mathrm{E}\{|h_{\mu\nu}^*((i+Q)L_b)|^2\}}\sqrt{\mathrm{E}\{|\hat{h}_{\mu\nu}^*((i+Q)L_b)|^2\}}} \tag{5}$$

and the prediction error as

$$\epsilon_{\mu\nu}((i+Q)L_b) := h_{\mu\nu}((i+Q)L_b) - \hat{h}_{\mu\nu}((i+Q)L_b). \tag{6}$$

We seek the channel predictor that minimizes the MSE

$$\begin{aligned} \sigma_\epsilon^2 &:= \mathrm{E}\{|\epsilon_{\mu\nu}((i+Q)L_b)|^2\} \\ &= \mathrm{E}\{|h_{\mu\nu}((i+Q)L_b) - \mathbf{w}_{\mu\nu}^{\mathcal{H}} \bar{\boldsymbol{h}}_{\mu\nu}(i)|^2\}. \end{aligned} \tag{7}$$

The optimal solution is easily found to be

$$\begin{aligned} \mathbf{w}_{\mu\nu} = \left(\mathrm{E}\left\{\bar{\boldsymbol{h}}_{\mu\nu}(i)\bar{\boldsymbol{h}}_{\mu\nu}^{\mathcal{H}}(i)\right\}\right)^{-1} \\ \times \mathrm{E}\{h^*((i+Q)L_b)\bar{\boldsymbol{h}}_{\mu\nu}(i)\}. \end{aligned} \tag{8}$$

Based on (3) and AS0)–AS2), we obtain

$$\mathbf{w}_{\mu\nu} = \mathbf{R}^{-1}\mathbf{r} \tag{9}$$

where $\mathbf{R}$ and $\mathbf{r}$ have entries respectively given by

$$[\mathbf{R}]_{p,q} = J_0(2\pi f_d |p-q| L_b T_s) + \frac{N_0}{E_p}\delta(p-q), \tag{10}$$

$$[\mathbf{r}]_p = J_0(2\pi f_d |Q+p| L_b T_s) \tag{11}$$

with $p, q \in [0, P]$. Based on $\mathbf{w}_{\mu\nu}$ in (9), we have $\sigma_\epsilon^2 = 1 - \mathbf{r}^{\mathcal{H}}\mathbf{R}^{-1}\mathbf{r}$. Thanks to the orthogonality principle in MMSE filtering, $\hat{h}_{\mu\nu}((i+Q)L_b)$ and $\epsilon_{\mu\nu}((i+Q)L_b)$ are uncorrelated. Hence, we deduce from (6) that the predicted channel $\hat{h}_{\mu\nu}((i+Q)L_b)$ has zero mean and variance $1 - \sigma_\epsilon^2 = \mathbf{r}^{\mathcal{H}}\mathbf{R}\mathbf{r}$. Therefore, we can further simplify $\rho_{\mu\nu}$ in (5) as

$$\rho_{\mu\nu} = \sqrt{\mathbf{r}^{\mathcal{H}}\mathbf{R}^{-1}\mathbf{r}}. \tag{12}$$

Since $\mathbf{w}_{\mu\nu}$ and $\rho_{\mu\nu}$ do not depend on μ and ν, we will henceforth denote them as $\mathbf{w}$ and ρ. So, we can write $\sigma_\epsilon^2 = 1 - \rho^2$. Similar to the definition of $\mathbf{H}(n)$, we collect the predicted channels $\hat{h}_{\mu\nu}((i+Q)L_b)$ and the prediction errors $\epsilon_{\mu\nu}((i+Q)L_b)$ into matrices $\hat{\mathbf{H}}((i+Q)L_b)$ and $\boldsymbol{\Xi}((i+Q)L_b)$. Corresponding to (6), we obtain

$$\mathbf{H}((i+Q)L_b) = \hat{\mathbf{H}}((i+Q)L_b) + \boldsymbol{\Xi}((i+Q)L_b) \tag{13}$$

where

$$\begin{aligned} \hat{\mathbf{H}}((i+Q)L_b) &\sim \mathcal{CN}(\mathbf{0}_{N_t \times N_r}, N_r \rho^2 \mathbf{I}_{N_t}) \\ \boldsymbol{\Xi}((i+Q)L_b) &\sim \mathcal{CN}(\mathbf{0}_{N_t \times N_r}, N_r \sigma_\epsilon^2 \mathbf{I}_{N_t}). \end{aligned}$$

A well-documented performance measure for channel estimation and prediction is the normalized channel MSE (NMSE)

$$\text{NMSE} := \frac{\text{E}\{\|\mathbf{H}((i+Q)L_\text{b}) - \hat{\mathbf{H}}((i+Q)L_\text{b})\|_\text{F}^2\}}{\text{E}\{\|\mathbf{H}((i+Q)L_\text{b})\|_\text{F}^2\}} \quad (14)$$

where $\|\cdot\|_\text{F}$ stands for the Frobenius norm. With the MMSE channel predictor, the NMSE for (13) can be expressed as

$$\text{NMSE} = \sigma_\epsilon^2 = 1 - \rho^2 = 1 - \mathbf{r}^{\mathcal{H}}\mathbf{R}\mathbf{r}. \quad (15)$$

From (15) and (12), the correlation coefficient is uniquely determined from the NMSE as

$$\rho = \sqrt{1 - \text{NMSE}}. \quad (16)$$

As evidenced from (10), (11), and (15), the NMSE, or equivalently the correlation coefficient ρ, depends on various system parameters, including the pilot spacing L_b, the pilot signal-to-noise ratio (SNR) E_p/N_0, the Doppler spread f_d, the prediction filter order P, and the prediction range $QL_\text{b}T_s$. We will illustrate these dependencies numerically in Section V. However, we will establish in Section IV that various system parameters affect the overall BER performance of the transmit-beamforming based adaptive system, only through a *single variable*: the NMSE (or, equivalently ρ).

Three remarks are now in order.

Remark 1 (Noiseless Delayed Feedback): Suppose that we use a prediction filter with only one tap $w(0)$. Furthermore, let us assume that the channel estimate in (3) is perfect, i.e., the additive noise is absent $(N_0 = 0)$. We then obtain from (9) and (4) that $w(0) = \rho_0 := J_0(2\pi f_d Q L_\text{b} T_s)$, and

$$\hat{h}_{\mu\nu}((i+Q)L_\text{b}) = \rho_0 h_{\mu\nu}(iL_\text{b}). \quad (17)$$

And from (12), we have $\rho = |\rho_0|$. This special case corresponds to the noiseless delayed feedback in [10] and [22].

Remark 2 (Channel Mean Feedback): Notice that (13) is in the form of channel mean feedback [31], where the true channels are modeled as Gaussian distributed with nonzero mean but white, when conditioned on the feedback. Hence, the MMSE predictor for i.i.d. Rayleigh MIMO channels presented here is another realization of the general notion of "channel mean feedback," in addition to those summarized in [34].

Remark 3 (Difference From [5]): A related (but somewhat different) extension of PSAM to MIMO channels has been carried out in [5]. The difference is that [5] focuses on the receiver of a nonadaptive system, and forms an MMSE channel estimator using infinite impulse response noncausal filtering. In contrast, we are dealing with adaptive transmitters, and have to rely on finite impulse response causal channel predictors.

III. Adaptive Modulation Based on Transmit-Beamforming

Although the predicted channels $\hat{\mathbf{H}}((i+Q)L_\text{b})$ differ from the true channels $\mathbf{H}((i+Q)L_\text{b})$ as shown in (13), most existing adaptive transmitters assume the former to be perfect. For notational brevity, we will drop the time index, and denote, e.g., $\hat{\mathbf{H}}((i+Q)L_\text{b})$ by $\hat{\mathbf{H}}$.

We focus on an adaptive system based on transmit beamforming. Specifically, our adaptive transmit-beamformer first draws an information symbol $s(n)$ from a suitable constellation, that has average energy E_s, and then transmits the vector $\mathbf{u}^* s(n)$ across N_t antennas, using the beam-steering vector $\mathbf{u} := [u_1, \ldots, u_{N_t}]^T$ that we will specify soon. Assuming perfect CSI with $\hat{\mathbf{H}} = \mathbf{H}$, conventional transmit-beamforming is optimal in terms of maximizing the SNR at the receiver [7], [22], [34]. However, we should point out that transmit-beamforming is sub-optimal from a MIMO capacity point of view, since only one data stream, instead of up to $\min(N_t, N_r)$ parallel streams, is transmitted through multiple antennas [29]. Despite the potential rate loss, however, transmit-beamforming offers two major advantages: i) it allows for simple transceiver processing [15], [22], [31]; and ii) it is well suited for receivers with a small number of (e.g., one or two) antennas. Notice also that transmit-beamforming has been included in the standardization proposal for third generation (3G) wideband CDMA systems [1], [2].

Our objective in this paper is to analyze the impact of imperfect CSI at the transmitter on the overall BER performance of the adaptive transmit-beamformer. To this end, we assume[1]: *AS3): the channel estimates are error-free at the receiver.*

Assumption AS3) is reasonable since the received signals can be stored so that noncausal channel estimation (smoothing) with high accuracy can be performed [24]. This assumption is commonly adopted when designing adaptive systems; see, e.g., [4], [8], [10]–[14], [30], and [32]. Based on AS3), the received SNR at the νth receive antenna is $\gamma_\nu = |\mathbf{u}^{\mathcal{H}}\mathbf{h}_\nu|^2 E_s/N_0$. Considering maximum ratio combining (MRC) across the N_r receive antennas, we obtain the overall SNR as: $\gamma = \sum_{\nu=1}^{N_r} \gamma_\nu$. Since

$$\gamma = \sum_{\nu=1}^{N_r} |\mathbf{u}^{\mathcal{H}}\mathbf{h}_\nu|^2 E_s/N_0 = \mathbf{u}^{\mathcal{H}}\mathbf{H}\mathbf{H}^{\mathcal{H}}\mathbf{u}E_s/N_0 \quad (18)$$

each beamformed information symbol with MRC at the receiver adheres to an equivalent scalar input-output relationship

$$y(n) = h_\text{eqv}s(n) + w(n), \quad h_\text{eqv} := \sqrt{\mathbf{u}^{\mathcal{H}}\mathbf{H}\mathbf{H}^{\mathcal{H}}\mathbf{u}}. \quad (19)$$

With perfect CSI $\mathbf{H} = \hat{\mathbf{H}}$, the optimal beam-steering vector $\mathbf{u}$, that maximizes the received SNR is the eigenvector corresponding the largest eigenvalue of $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}$ [34]. To state this result formally, let the eigen decomposition of $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}$ be

$$\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}} = \mathbf{U}_\text{H}\mathbf{D}_\text{H}\mathbf{U}_\text{H}^{\mathcal{H}}, \quad \mathbf{D}_\text{H} := \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{N_t}) \quad (20)$$

where $\mathbf{U}_\text{H} := [\mathbf{u}_{\text{H},1}, \ldots, \mathbf{u}_{\text{H},\text{N}_\text{t}}]$ contains N_t eigenvectors, and $\mathbf{D}_\text{H}$ has the corresponding N_t eigenvalues on its diagonal in a nonincreasing order: $\lambda_1 \geq \lambda_2 \geq, \ldots, \geq \lambda_{N_t}$. The optimal beamforming vector is then chosen as

$$\mathbf{u} = \mathbf{u}_{\text{H},1}. \quad (21)$$

When $\mathbf{H} = \hat{\mathbf{H}}$, we have $h_\text{eqv}^2 = \lambda_1$. Therefore, the performance enhancement offered by transmit-beamforming in MIMO channels is realized through the largest eigenvalue λ_1.

Having specified the optimal beam direction, we consider next the constellation switching module of our adaptive transmitter. For simplicity and mathematical tractability, we here consider a suboptimal *constant* power transmission policy, rather than the optimal power adaptation[2] established in [8]; this is also the case in existing works [4], [24] on BER performance

[1]AS3) distinguishes our work from those in [5], [18], [21], and [27] where the impact of imperfect CSI at the receiver is analyzed for nonadaptive transmissions (i.e., with no CSI at the transmitter), on information rate [5], [18], [21] or BER performance [27], respectively.

[2]The optimal power allocation itself requires a numerical search to find the constellation switching thresholds, as evidenced in [8, eq. (52)].

analysis with imperfect CSI. We will adopt N rectangular (and square) quadrature-amplitude-modulation (QAM) constellations with size $M_i = 2^i, i = 1, N$. To facilitate transmitter adaptation, we choose those rectangular QAM's that consist of two independent pulse-amplitude-modulations (PAM), one on the In-phase branch, and the other on the Quadrature branch [33]. When the channels experience deep fades, we will allow our adaptive design to suspend data transmission (this will correspond to setting $M_0 = 0$). Since the transmitter sees an equivalent scalar channel with amplitude $h_{\rm eqv}^2 = \lambda_1$, the channel quality is determined solely by the largest eigenvalue of the MIMO channel. The adaptive modulator partitions the interval $[0, \infty)$ into $N+1$ disjoint but consecutive regions, with the boundary points denoted as $\{\alpha_i\}_{i=0}^{N+1}$, where $\alpha_0 = 0$ and $\alpha_{N+1} = \infty$. The constellation is then chosen according to

$$M = M_i, \text{ when } \lambda_1 \in [\alpha_i, \alpha_{i+1}). \tag{22}$$

The overall probability that the constellation M_i is chosen is

$$\Pr(M_i) = \Pr(\lambda_1 \in [\alpha_i, \alpha_{i+1})) = \int_{\alpha_i}^{\alpha_{i+1}} p_{\lambda_1}(\lambda_1) d\lambda_1 \tag{23}$$

where $p_{\lambda_1}(\lambda_1)$ is the probability density function (pdf) of λ_1. The overall transmission rate of such an adaptive MIMO system is, thus

$$R = \frac{L_{\rm b} - N_t}{L_{\rm b}} \sum_{i=1}^{N} \log_2(M_i) \Pr(M_i) \tag{24}$$

where the spectral efficiency loss incurred by the pilots has been considered in the numerator.

The outage probability of no transmission is

$$\mathrm{P}_{\rm out} = \Pr(M_0). \tag{25}$$

To calculate R and $\mathrm{P}_{\rm out}$, it suffices to determine the pdf of λ_1, and the boundaries $\{\alpha_i\}_{i=1}^{N}$. To specify these boundary points with perfect CSI, we rely on the approximate BER performance [12], [35]

$$\mathrm{BER}(M_i, \lambda_1) \approx 0.2 \exp(-\lambda_1 g_i E_s / N_0) \tag{26}$$

where the constellation specific constant g_i is chosen as

$$g_i = \frac{3}{2(M_i - 1)} \quad \text{for square QAM}, \tag{27}$$

$$g_i = \frac{6}{5M_i - 4} \quad \text{for rectangular QAM}. \tag{28}$$

To maintain a target BER denoted as $\mathrm{BER}_{\rm target}$, the transmitter determines the boundary points as [cf. (22), (26)]

$$\alpha_i = \frac{-\ln(5\,\mathrm{BER}_{\rm target})}{g_i E_s / N_0}, \quad i = 1, \ldots, N. \tag{29}$$

Based on (22) and (26), the average BER associated with constellation M_i is

$$\overline{\mathrm{BER}}(M_i) = \int_{\alpha_i}^{\alpha_{i+1}} \mathrm{BER}(M_i, \lambda_1) p(\lambda_1) d\lambda_1. \tag{30}$$

The overall system BER is the ratio of the number of bits in error over the total number of transmitted information bits, expressed as [4]

$$\overline{\mathrm{BER}} = \frac{\sum_{i=1}^{N} \log_2(M_i) \overline{\mathrm{BER}}(M_i)}{\sum_{i=1}^{N} \log_2(M_i) \Pr(M_i)}. \tag{31}$$

The boundary selection in (29) corresponds to an instantaneous BER constraint considered in [8], since for any realization of $\lambda_1 \in [\alpha_i, \alpha_{i+1})$, we have $\mathrm{BER}(M_i, \lambda_1) \le \mathrm{BER}_{\rm target}$. Further rate improvement is possible, by adopting an average BER constraint: $\overline{\mathrm{BER}} \le \mathrm{BER}_{\rm target}$ [8]. However, numerical search is needed to find $\{\alpha_i\}_{i=1}^{N}$ [8], which complicates the transmitter design. For this reason, [4] and [24] have also focused on the instantaneous BER constraint. Thanks to the discrete constellation choices and the instantaneous BER constraint, the average BER is actually below the target; see, e.g., BER plots in [4].

If indeed the channel prediction is perfect, the adaptive transmitter improves the transmission rate while maintaining the target BER. However, the actual BER may increase due to the imperfect channel prediction. We next throughly investigate how the BER performance is affected by the channel prediction errors.

IV. BER Performance With Imperfect CSI

For a given realization of $\hat{\mathbf{H}}$, the true channel $\mathbf{H}$ can be viewed as a Gaussian random matrix with nonzero mean and white covariance. For each $\mathbf{u} = \mathbf{u}_{\mathrm{H},1}$ specified in (21) based by $\hat{\mathbf{H}}$, the scalar channel $h_{\rm eqv}$ in (19) is now random. The average BER, averaged over all possible realizations of $\mathbf{H}$, can then be expressed as

$$\overline{\mathrm{BER}}(M_i, \hat{\mathbf{H}}) = \mathrm{E}_{\mathbf{H}}\{\mathrm{BER}(M_i, \hat{\mathbf{H}}, \mathbf{H})\} \tag{32}$$

where

$$\mathrm{BER}(M_i, \hat{\mathbf{H}}, \mathbf{H}) \approx 0.2 \exp(-h_{\rm eqv}^2 g_i E_s / N_0) \tag{33}$$

is the instantaneous BER for each realization of $\mathbf{H}$. As in (30), the average BER for each constellation is

$$\overline{\mathrm{BER}}(M_i) = \int_{\alpha_i}^{\alpha_{i+1}} \overline{\mathrm{BER}}(M_i, \hat{\mathbf{H}}) p(\lambda_1) d\lambda_1. \tag{34}$$

Substituting (34) into (31), the overall system BER can then be found in the presence of prediction errors. To calculate the overall BER, we need to find $\Pr(M_i)$ and $\overline{\mathrm{BER}}(M_i)$.

A. General Expressions

We first evaluate the average BER in (32) per feedback $\hat{\mathbf{H}}$. Define $\tilde{\mathbf{h}} := \mathbf{H}^{\mathcal{H}}\mathbf{u} = (\hat{\mathbf{H}} + \mathbf{\Xi})^{\mathcal{H}}\mathbf{u}$. Conditioned on $\hat{\mathbf{H}}$, we have $\tilde{\mathbf{h}} \sim \mathcal{CN}(\hat{\mathbf{H}}^{\mathcal{H}}\mathbf{u}, \sigma_\epsilon^2 \mathbf{I}_{N_r})$. Furthermore, with $\mathbf{u} = \mathbf{u}_{\mathrm{H},1}$, we have $\|\hat{\mathbf{H}}^{\mathcal{H}}\mathbf{u}\|^2 = \lambda_1$. Using (19), we re-express (33) as

$$\overline{\mathrm{BER}}(M_i, \hat{\mathbf{H}}) = 0.2\, \mathrm{E}_{\tilde{\mathbf{h}}}\{\exp(-\tilde{\mathbf{h}}^{\mathcal{H}}\tilde{\mathbf{h}} g_i E_s / N_0)\}. \tag{35}$$

We shall use the following identity [28, eq, (15)]

$$\mathrm{E}_{\mathbf{z}}\{\exp(-\mathbf{z}^{\mathcal{H}}\mathbf{A}\mathbf{z})\} = \frac{\exp(-\boldsymbol{\mu}^{\mathcal{H}}\mathbf{A}(\mathbf{I} + \mathbf{\Sigma}\mathbf{A})^{-1}\boldsymbol{\mu})}{\det(\mathbf{I} + \mathbf{\Sigma}\mathbf{A})} \tag{36}$$

where $\mathbf{z} \sim \mathcal{CN}(\boldsymbol{\mu}, \mathbf{\Sigma})$, and $\mathbf{A}$ is a Hermitian matrix with matching dimension. For brevity, we define

$$\phi_i = g_i E_s / N_0. \tag{37}$$

Substituting into (36) with $\mathbf{z} = \tilde{\mathbf{h}}$ and $\mathbf{A} = \phi_i \mathbf{I}_{N_r}$, we simplify (35) to

$$\overline{\mathrm{BER}}(M_i, \hat{\mathbf{H}}) = \frac{0.2}{(1 + \sigma_\epsilon^2 \phi_i)^{N_r}} \exp\left(-\frac{\lambda_1 \phi_i}{1 + \sigma_\epsilon^2 \phi_i}\right) \tag{38}$$

and consequently (34) to

$$\overline{\mathrm{BER}}(M_i) = \frac{0.2}{(1+\sigma_\epsilon^2\phi_i)^{N_r}} \times \int_{\alpha_i}^{\alpha_{i+1}} \exp\left(-\frac{\lambda_1\phi_i}{1+\sigma_\epsilon^2\phi_i}\right) p_{\lambda_1}(\lambda_1)d\lambda_1. \quad (39)$$

We next go on to specify $p_{\lambda_1}(\lambda_1)$. For ease of notation, we define two constants

$$k := \min(N_t, N_r), \quad d := |N_t - N_r| = \max(N_t, N_r) - k. \quad (40)$$

For systems with N_t transmit- and N_r receive- antennas, there are at most k nonzero eigenvalues for $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}$ [cf. (20)]. Let $\{\xi_i\}_{i=1}^k$ denote the ordered eigenvalues of the matrix $(1/\rho^2)\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}$, so that

$$\lambda_i = \rho^2\xi_i, \quad \forall i \in [1,k]. \quad (41)$$

Since each entry of $(1/\rho)\hat{\mathbf{H}}$ is Gaussian with distribution $\mathcal{CN}(0,1)$, the joint distribution of the ordered eigenvalues $\{\xi_i\}_{i=1}^k$ is [16], [29]

$$p(\xi_1,\ldots,\xi_k) = C_{k,d}\exp\left(-\sum_i \xi_i\right)\prod_i \xi_i^d \prod_{i<j}(\xi_i-\xi_j)^2, \quad \xi_1 \geq \cdots \geq \xi_k \quad (42)$$

where $C_{k,d}$ is a normalizing constant. From the joint pdf we are interested in finding the marginal pdf of ξ_1

$$p_{\xi_1}^{(k)}(\xi_1) = \int_0^{\xi_1} d\xi_2 \cdots \int_0^{\xi_{k-1}} d\xi_k p(\xi_1,\ldots,\xi_k). \quad (43)$$

Closed-form expression for $p_{\xi_1}^{(k)}(\xi_1)$ is available in, e.g., [17] and [20]. Based on $p_{\xi_1}^{(k)}(\xi_1)$, let us define the integral

$$\Psi_{\xi_1}^{(k)}(a,x) = \int_0^x e^{-(a-1)\xi_1} p_{\xi_1}^{(k)}(\xi_1)d\xi_1 \quad (44)$$

Notice that $\Psi_{\xi_1}^{(k)}(1,x) = \int_0^x p_{\xi_1}^{(k)}(\xi_1)d\xi_1$ is the cumulative distribution function (cdf) of ξ_1.

Since $\lambda_1 = \rho^2\xi_1$, we obtain from (23) and (44)

$$\Pr(M_i) = \Psi_{\xi_1}^{(k)}\left(1, \frac{\alpha_{i+1}}{\rho^2}\right) - \Psi_{\xi_1}^{(k)}\left(1, \frac{\alpha_i}{\rho^2}\right). \quad (45)$$

Similarly, we simplify (39) to

$$\overline{\mathrm{BER}}(M_i) = \frac{0.2\left[\Psi_{\xi_1}^{(k)}\left(b_i, \frac{\alpha_{i+1}}{\rho^2}\right) - \Psi_{\xi_1}^{(k)}\left(b_i, \frac{\alpha_i}{\rho^2}\right)\right]}{[1+(1-\rho^2)\phi_i]^{N_r}} \quad (46)$$

where the constant b_i is defined as

$$b_i = 1 + \frac{|\rho|^2\phi_i}{1+\sigma_\epsilon^2\phi_i} = \frac{1+\phi_i}{1+(1-\rho^2)\phi_i}. \quad (47)$$

Substituting (45) and (46) into (24) and (31), we obtain the transmission rate and the average BER. Notice that the evaluation in (45) and (46) involves integration, which is numerically plausible but involved. This motivated us to pursue simple solutions for special cases with $k = 1$ and $k = 2$. The same procedure can be extended to cases with $k > 2$, although the final expression gets increasingly complicated. To derive these simple forms we will use some preliminary results on Gamma distributions, that we summarize next.

B. Preliminaries on Gamma Distributions

Let $\Gamma(m) := \int_0^\infty t^{m-1}e^{-t}dt$ denote the Gamma function with parameter m. The corresponding *normalized* incomplete Gamma function is defined as

$$\bar{\Gamma}(m,x) = \frac{1}{\Gamma(m)}\int_0^x t^{m-1}e^{-t}dt. \quad (48)$$

When m is a positive integer, we have

$$\Gamma(m) = (m-1)!, \quad \bar{\Gamma}(m,x) = 1 - e^{-x}\sum_{j=0}^{m-1}\frac{x^j}{j!}. \quad (49)$$

A nonnegative random variable γ with mean $\bar{\gamma} := \mathrm{E}\{\gamma\}$ is Gamma distributed with parameter m, when its pdf is

$$p_\gamma(\gamma) = \left(\frac{m}{\bar{\gamma}}\right)^m \frac{\gamma^{m-1}}{\Gamma(m)}\exp\left(-m\frac{\gamma}{\bar{\gamma}}\right), \quad \gamma \geq 0. \quad (50)$$

Based on this pdf, the cdf of γ is

$$F_\gamma(x) = \int_0^x p_\gamma(\gamma)d\gamma = \bar{\Gamma}\left(m, \frac{mx}{\bar{\gamma}}\right). \quad (51)$$

From (50) and (51), we obtain the following identity that will become handy in our subsequent derivations

$$\int_0^x \gamma^{m-1}e^{-b\gamma}d\gamma = \frac{1}{b^m}\Gamma(m)\bar{\Gamma}(m,bx) \quad (52)$$

where $b \geq 0$ is a scalar.

We are now ready to present our results for special cases.

C. Multiple-Input Single-Output or Single-Input Multiple-Output

We first look at the simple case with $k = \min(N_t, N_r) = 1$, and $d = \max(N_t, N_r) - 1$. In this case, eigen-decomposition in (20) yields only one nonzero eigenvalue $\lambda_1 = \rho^2\xi_1$. The pdf of ξ_1 is deduced from (42) as

$$p_{\xi_1}^{(1)}(\xi_1) = \frac{\xi_1^d}{\Gamma(d+1)}\exp(-\xi_1), \quad \xi_1 \geq 0 \quad (53)$$

which is nothing but a Gamma distribution with parameter $m = d+1$, and mean $\mathrm{E}\{\xi_1\} = d+1$ [cf. (50)]. Substituting (53) into (44), we obtain

$$\Psi_{\xi_1}^{(1)}(a,x) = a^{-(d+1)}\bar{\Gamma}(d+1,ax). \quad (54)$$

Setting $k = 1$ in (45) and (46), we end up with closed-forms for $\Pr(M_i)$ and $\overline{\mathrm{BER}}(M_i)$, that as we mentioned earlier allow for simple evaluation of transmission rates and average BER.

D. Multiple-Input Two-Output or Two-Input Multiple-Output

Here, we consider the case with $k = 2$, and $d = \max(N_t, N_r) - 2$. From (42), the two nonzero ordered eigenvalues adhere to the bi-variate pdf

$$\begin{aligned} p(\xi_1,\xi_2) &= C_{2,d}e^{-\xi_1-\xi_2}\xi_1^d\xi_2^d(\xi_1-\xi_2)^2 \\ &= C_{2,d}e^{-\xi_1-\xi_2}\left(\xi_1^{d+2}\xi_2^d - 2\xi_1^{d+1}\xi_2^{d+1} + \xi_1^d\xi_2^{d+2}\right) \end{aligned} \quad (55)$$

where the normalizing constant is found to be

$$C_{2,d} = [\Gamma(d+2)\Gamma(d+1)]^{-1}.$$

Using (49) and (52), we find the marginal pdf of the largest eigenvalue as

$$\begin{aligned} p_{\xi_1}^{(2)}(\xi_1) &= \int_0^{\xi_1} p(\xi_1,\xi_2)d\xi_2 \\ &= e^{-\xi_1}\left[\frac{1}{(d+1)!}\xi_1^{d+2} - \frac{2}{d!}\xi_1^{d+1} + \frac{d+2}{d!}\xi_1^{d}\right] \\ &\quad - e^{-2\xi_1}\left[\sum_{j=0}^{d-2}\frac{(d-j)(d-j-1)}{(j+2)!(d+1)!}\xi_1^{d+j+2}\right. \\ &\quad \left. + \frac{1}{(d-1)!}\xi_1^{d+1} + \frac{d+2}{d!}\xi_1^{d}\right]. \end{aligned} \tag{56}$$

Substituting (56) into (44), we obtain

$$\begin{aligned} &\Psi_{\xi_1}^{(2)}(a,x) \\ &= \frac{(d+2)}{a^{d+3}}\bar{\Gamma}(d+3,ax) - \frac{2(d+1)}{a^{d+2}}\bar{\Gamma}(d+2,ax) \\ &\quad + \frac{(d+2)}{a^{d+1}}\bar{\Gamma}(d+1,ax) - \frac{d(d+1)}{(1+a)^{d+2}}\bar{\Gamma}(d+2,(1+a)x) \\ &\quad - \frac{d+2}{(1+a)^{d+1}}\bar{\Gamma}(d+1,(1+a)x) \\ &\quad - \sum_{j=0}^{d}\frac{(d-j)(d-j-1)(d+j+2)!}{(j+2)!(1+a)^{d+j+3}(d+1)!}\bar{\Gamma}(d+j+3,(1+a)x). \end{aligned} \tag{57}$$

Setting $k = 2$ in (45) and (46), we again arrive at closed-forms for $\Pr(M_i)$ and $\overline{\mathrm{BER}}(M_i)$.

E. Discussion and Links

Notice that the matrices $\hat{\mathbf{H}}\hat{\mathbf{H}}^{\mathcal{H}}$ and $\hat{\mathbf{H}}^{\mathcal{H}}\hat{\mathbf{H}}$ share the same nonzero eigenvalues. However, the BER in (46) explicitly depends on the number of receive antennas, except when the CSI is perfect ($\rho = 1$). Hence, only with perfect CSI, N_t and N_r can be exchanged without affecting the BER performance of the adaptive system. With imperfect CSI, reciprocity does not hold, but we recommend to arrange N_t and N_r so that $N_t \leq N_r$, if such a change is practically feasible.

We also wish to link our analytical results to existing BER analysis in adaptive systems. The BER performance is analyzed in [4] for single-input-single-output (SISO) systems with Nakagami-m fading channels, assuming noiseless delayed feedback. It is also carried out for multidimensional trellis coded modulation in a SIMO (single-input multiple-output) system with Rayleigh fading channels, based on noisy predicted channels [24]. Notice that in a SIMO system with Rayleigh fading channels, the square root of the overall SNR after MRC is Nakagami distributed with parameter N_r. This explains why our results in (46) and (54) for SIMO systems have similar appearance with those in [4], eq. (50)], and [24]. The approaches in [4] and [24] rely on the correlation between the predicted SNR and the actual SNR. The SNR correlation approach is suitable for SISO and SIMO systems, where all the CSI needed at the transmitter is the overall SNR to perform the constellation selection [4], [24]. However, this approach is not applicable to adaptive MIMO systems, where all predicted channel values are needed in order to select not only the optimal constellation, but also the beam-steering vector $\mathbf{u}$.

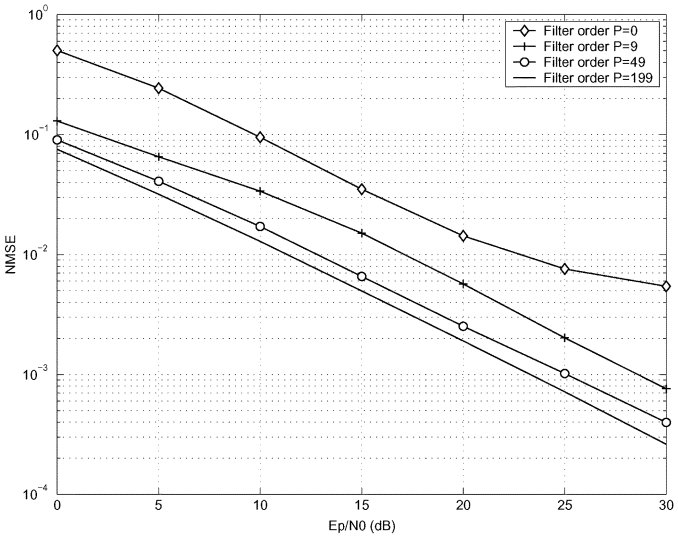

Fig. 3. NMSE versus pilot SNR E_p/N_0.

V. Numerical Results

In this section, we present numerical results for various antenna configurations (N_t, N_r). For adaptive modulation, we allow for QAM constellations with sizes $\{M_i = 2^i\}_{i=1}^8$.

Test Case 1 (Dependence of NMSE on Various System Parameters): As evidenced from (10), (11), and (15), the NMSE depends on various system parameters $(L_{\mathrm{b}}, E_p/N_0, f_dT_s, P, Q)$. Similar to [24], we consider a system with carrier frequency $f_c = 2$ GHz, and symbol rate of 400 kilo-symbols-per-second (ksps); hence, $T_s = 2.5$ μs. We consider a terminal speed of $v = 30$ m/s, which results in a Doppler spread of $f_d = 200$ Hz. The normalized Doppler spread is, thus, $f_dT_s = 5 \cdot 10^{-4}$.

We first set $Q = 2$ and $L_{\mathrm{b}} = 15$; i.e., we predict the channels $QL_{\mathrm{b}} = 30$ symbols ahead. We plot in Fig. 3 the dependence of NMSE on the pilot SNR E_p/N_0, and the prediction filter order P. As the filter order and the pilot SNR increases, the prediction accuracy improves considerably.

Next, we fix the filter order at $P = 199$. With the symbol period as unit, we plot in Fig. 4 the dependence of NMSE on the prediction distance QL_{b}, where two block lengths $L_{\mathrm{b}} = 10, 15$ are used in our MIMO PSAM. We infer from Fig. 4 that as the prediction distance increases, the prediction accuracy decreases quickly. However, to achieve a fixed level of prediction accuracy, the prediction range is considerably extended as the pilot SNR E_p/N_0 increases.

Test Case 2 (Noiseless Delayed Feedback): Here, we test the noiseless delayed feedback, that we dealt with in Remark 1. We define the overall feedback delay as $\tau := QL_{\mathrm{b}}T_s$ and, thus, $\rho = J_0(2\pi f_d\tau)$. We set $\mathrm{BER}_{\mathrm{target}} = 10^{-3}$, and plot in Fig. 5 the BER dependence on the normalized delay $f_d\tau$. First, we observe that our curves in Fig. 5 are similar to those in [4, Fig. 14]). It is concluded in [4] that the critical normalized delay is $f_d\tau = 10^{-2}$, below which the adaptive system operates satisfactorily. With $f_d\tau = 10^{-2}$, we have $\rho = 0.999$ and $\mathrm{NMSE} = 2 \cdot 10^{-3}$.

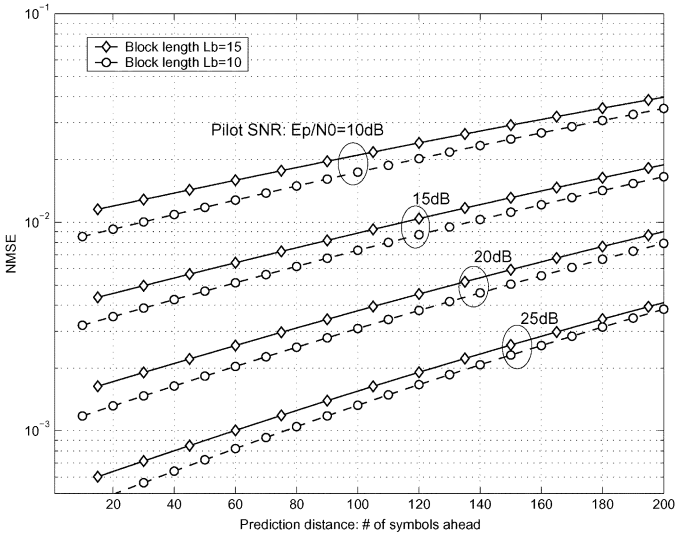

Fig. 4. NMSE versus prediction distance ($QL_{\rm b}$ symbols ahead).

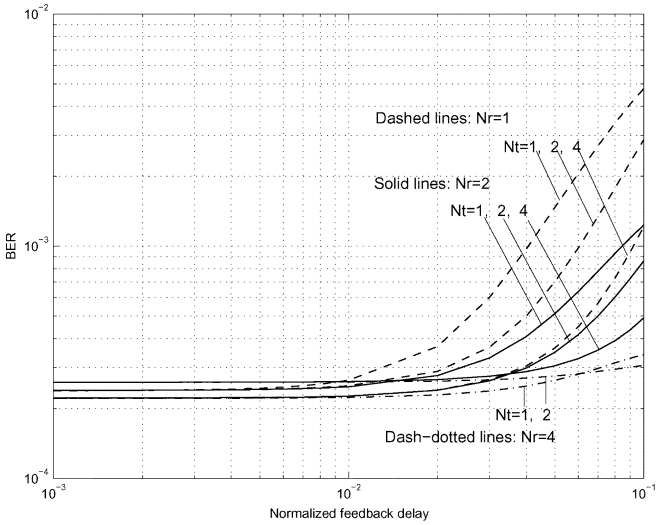

Fig. 5. Noiseless delayed feedback.

We next present general results with noisy predicted channels. Recall that the system performance depends on various system parameters only through the NMSE.

Test Case 3 (Noisy Predicted Channels): We set $\mathrm{BER}_{\rm target} = 10^{-3}$, and fix $L_{\rm b} = 15$ when computing the transmission rate using (24). With $E_s/N_0 = 10$ dB, we plot in Figs. 6 and 7 the dependence on the prediction NMSE of the system BER, and the transmission rate, respectively. Figs. 8 and 9 are the counterparts of Figs. 6 and 7, but with $E_s/N_0 = 20$ dB. Plots with other E_s/N_0 values have similar shapes. From Figs. 6 and 8, we observe that the BER's remain almost constant when $\mathrm{NMSE} < 10^{-2}$, but deteriorate quickly when $\mathrm{NMSE} > 10^{-2}$. On the other hand, the transmission rates remain nearly constant when $\mathrm{NMSE} < 10^{-1}$, and decrease quickly when $\mathrm{NMSE} > 10^{-1}$, as shown in Figs. 7 and 9. When the transmission rate drops considerably, the corresponding BER also decreases, since only signals from small size constellations will be used (in the extreme case with $\rho = 0$, no transmission is performed and, hence, the BER is zero). As the number of antennas increases, the BER performance generally becomes less sensitive to CSI imperfections. Receive-antennas prove to be more effective than transmit-antennas when dealing

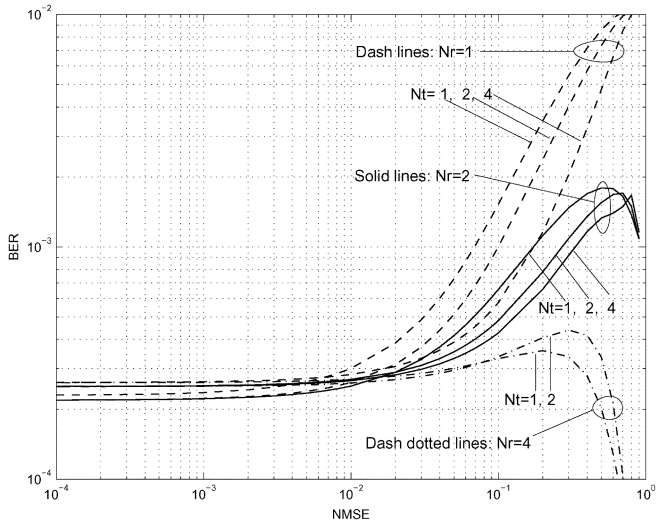

Fig. 6. BER performance at $E_s/N_0 = 10$ dB.

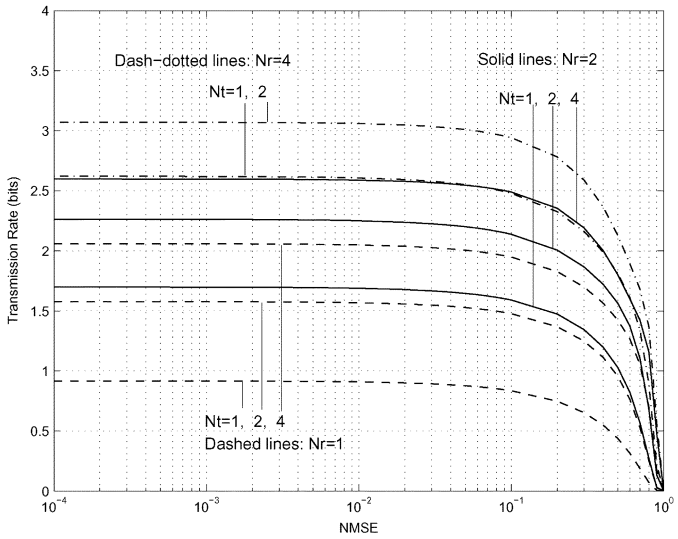

Fig. 7. Transmission rate at $E_s/N_0 = 10$ dB.

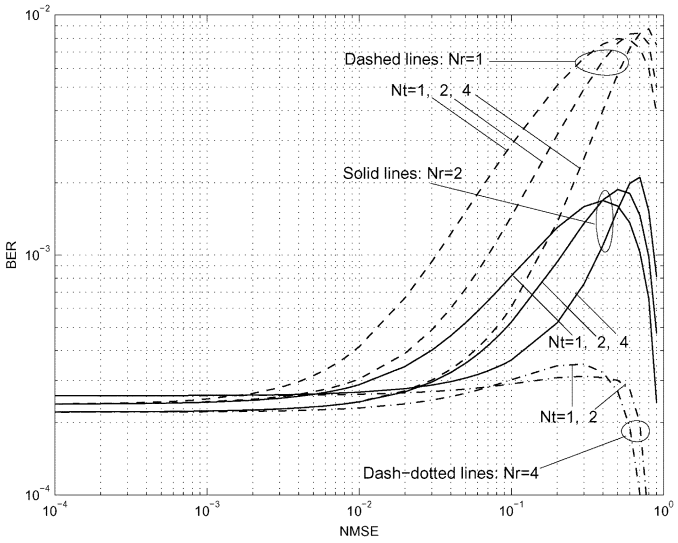

Fig. 8. BER performance at $E_s/N_0 = 20$ dB.

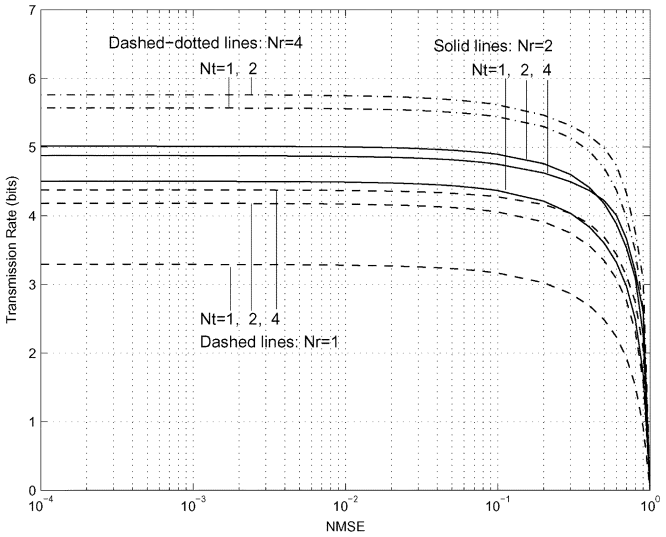

Fig. 9. Transmission rate at $E_s/N_0 = 20$ dB.

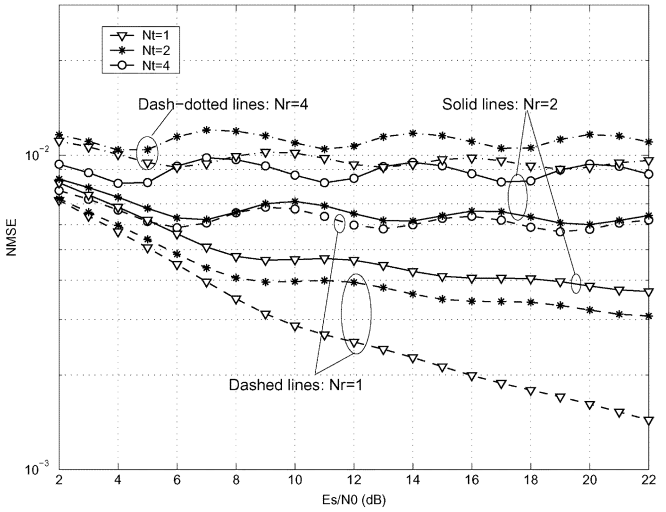

Fig. 10. The critical NMSE with $\mathrm{BER}_{\mathrm{actual}} \leq 1.1\,\mathrm{BER}_{\mathrm{ideal}}$.

with imperfect CSI at the transmitter. With four receive antennas, the BER fluctuation is essentially small.

We can now formally answer the question: *how accurate channel prediction needs to be*, so that the predicted channels can be used as if they were estimated perfectly, in a beamforming-based adaptive multiantenna system. For each E_s/N_0 and NMSE, we calculate two BERs: one is $\mathrm{BER}_{\mathrm{ideal}}$, which is obtained by assuming that the true channels coincide with predicted channels, and the other is $\mathrm{BER}_{\mathrm{actual}}$, namely the actual BER in the presence of prediction error. For each E_s/N_0, we gradually increase NMSE from 10^{-6}, and locate the first NMSE value for which

$$\mathrm{BER}_{\mathrm{actual}} = 1.1\,\mathrm{BER}_{\mathrm{ideal}}. \tag{58}$$

Those NMSE values are collected in Fig. 10. Hence, for each antenna configuration, when the actual NMSE is below the corresponding curve in Fig. 10, the transmitter is assured that the actual BER is off from the $\mathrm{BER}_{\mathrm{ideal}}$ by less than 10%. The corresponding transmission rate with NMSE values in Fig. 10 is plotted in Fig. 11.

In a nutshell, our analytical and numerical results suggest the following design guidelines for practical adaptive MIMO systems based on transmit-beamforming.

- G1: For each antenna configuration (N_t, N_r) and operating $\mathrm{SNR}(E_s/N_0)$, determine from Fig. 10. the critical NMSE value.
- G2: Determine the actual NMSE from, e.g., Figs. 3 and 4, based on various system parameters including $(L_{\mathrm{b}}, E_p/N_0, f_d T_s, P, Q)$.
- G3: If the actual NMSE is below the critical value, the adaptive transmitter can treat the predicted channels as being perfect. Otherwise, the transmitter needs to figure out the actual BER and transmission rate, as done in Figs. 6, 7, 8, and 9. Depending on the designer's judgment on whether these actual BER and rates are acceptable or not, transmitter designs without, or, with explicit CSI imperfection considered [10], [35], will be finally decided for deployment.

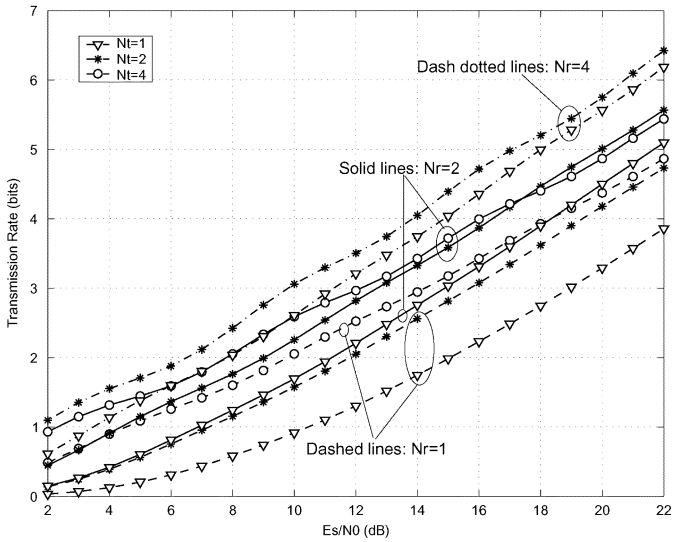

Fig. 11. The transmission rate with the desired prediction accuracy.

VI. Conclusion

In this paper, we addressed the question of how accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over MIMO Rayleigh fading channels. We first presented an MMSE channel predictor based on MIMO pilot symbol assisted modulation. We then analyzed the impact of the channel prediction error on the system's BER performance, where we obtained closed-form BER expressions. Our numerical results revealed the critical value of the normalized prediction error, below which the predicted channels can be treated as perfect by the adaptive modulator; otherwise, explicit consideration of the channel imperfection is needed at the transmitter.

Acknowledgment

The authors would like to thank Dr. H. Holm, for providing the preprint of [24].

REFERENCES

[1] Third Generation Partnership Project (3GPP), "Physical channels and mapping of transport channels onto physical channels,", 3GPP TS 25.211 V5.1.0, 2002.

[2] Third Generation Partnership Project (3GPP), "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. 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] J. Baltersee, G. Fock, and H. Meyr, "Achievable rate of MIMO channels with data-aided channel estimation and perfect interleaving," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 2358–2368, Dec. 2001.

[6] J. K. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[7] J. K. Cavers, "Single-user and multiuser adaptive maximal ratio transmission for Rayleigh channels," *IEEE Trans. Veh. Technol.*, vol. 49, no. 6, pp. 2043–2050, Nov. 2000.

[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, Sept. 2001.

[9] 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.

[10] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. Commun.*, vol. 47, pp. 844–855, June 1999.

[11] A. J. Goldsmith and S.-G. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 595–602, May 1998.

[12] ——, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, Oct. 1997.

[13] 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.

[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 Telecommunications Conf.*, vol. 2, San Antonio, TX, Nov. 25–29, 2001, pp. 1256–1261.

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

[16] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, pp. 475–501, June 1964.

[17] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hyperthesis concerning complex multivariate normal populations," *Ann. Math. Statist.*, vol. 35, pp. 1807–1810, Dec. 1964.

[18] A. Lapidoth and S. Shamai, "Fading channels: How perfect need 'perfect side information' be?," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1118–1134, May 2002.

[19] 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.

[20] M. Kang and M.-S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 418–426, Apr. 2003.

[21] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inform. Theory*, vol. 46, pp. 933–946, May 2000.

[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. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[23] S. Ohno and G. B. Giannakis, "Average-rate optimal PSAM transmissions over time-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 712–720, Oct. 2002.

[24] G. E. Oien, H. Holm, and K. J. Hole, "Impact of channel prediction on adaptive coded modulation performance in Rayleigh Fading," *IEEE Trans. Veh. Technol.*, pp. 758–769, May 2004.

[25] ——, "Adaptive coded modulation with imperfect channel state information: System design and performance analysis aspects," presented at the *IEEE Int. Symp. Advances in Wireless Commun*., Victoria, BC, Canada, Sept. 23–24, 2002.

[26] P. Ormeci, X. Liu, D. L. Goeckel, and R. D. Wesel, "Adaptive bit-interleaved coded modulation," *IEEE Trans. Commun.*, vol. 49, pp. 1572–1581, Sept. 2001.

[27] X. Tang, M. S. Alouini, and A. J. Goldsmith, "Effect of channel estimation error on M-QAM BER performance in Rayleigh fading," *IEEE Trans. Commun.*, vol. 47, pp. 1856–1864, Dec. 1999.

[28] G. Taricco and E. Biglieri, "Exact pairwise error probability of space-time codes," *IEEE Trans. Inform. Theory*, vol. 48, pp. 510–513, Feb. 2002.

[29] I. E. Telatar, Capacity of multi-antenna Gaussian channels, in Bell Lab. Tech. Memo., 1995.

[30] 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.

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

[32] W. T. Webb and R. Steele, "Variable rate QAM for mobile radio," *IEEE Trans. Commun.*, vol. 43, pp. 2223–2230, July 1995.

[33] 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.

[34] 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, Oct. 2002.

[35] S. Zhou and G. B. Giannakis, "Adaptive modulation for multi-antenna transmissions with channel mean feedback," in *Proc. IEEE Int. Conf. Communications*, 2003, pp. 2281–2285.

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, both in electrical engineering and information science. He received the Ph.D. degree in electrical engineering from the University of Minnesota (UMN), Minneapolis, in 2002.

He has been an Assistant Professor with the Department of Electrical and Computer Engineering at the University of Connecticut (UCONN), Storrs, since 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.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece in 1981. From September 1982 to July 1986, he was at the University of Southern California (USC), Los Angeles, where he received the MSc. degree in electrical engineering in 1983, the MSc. degree in mathematics in 1986, and the Ph.D. degree IN ELECTRICAL Engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia 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 at the University of Minnesota, 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- and multiuser fading communication channels, complex-field and space-time coding, multicarrier, ultra-wide band wireless communication systems, cross-layer designs, and distributed sensor networks.

Dr. Giannakis is the (co-) recipient of six paper awards from the IEEE Signal Processing (SP) and Communications Societies (1992, 1998, 2000, 2001, 2003, and 2004). He also received the SP Society's Technical Achievement Award in 2000. He served as Editor-in-Chief of IEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON ON SIGNAL PROCESSING and 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 also served as a member of the IEEE-SP Society's Board of Governors, the Editorial Board for PROCEEDINGS OF THE IEEE and the steering committee of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.