Efficient Signaling Schemes for Wideband Space-time Wireless Channels Using Channel State Information*

Eko N. Onggosanusi, Barry D. Van Veen, and Akbar M. Sayeed

Department of Electrical and Computer Engineering
University of Wisconsin–Madison
Madison, WI 53706
onggo@cae.wisc.edu, [vanveen,akbar]@engr.wisc.edu

Abstract

An orthogonal decomposition of a general wideband space-time multipath channel is derived assuming antenna arrays at both the transmitter and receiver. Knowledge of channel state information is assumed at both the transmitter and receiver. The decomposition provides a framework for efficiently managing the degrees of freedom in the space-time channel to optimize any combination of bit-error rate and throughput in single-user or multiuser applications. The decomposition is used to derive efficient signaling schemes and receiver structures for a variety of scenarios. For a fixed throughput system, we investigate a power allocation scheme that minimizes the effective bit-error rate. In addition, a strategy to maximize the throughput under a worst-case bit-error rate constraint is proposed. For multiuser applications, we propose a signaling scheme that achieves orthogonality among users by exploiting the temporal channel modes which are common to all users.

Keywords: Multi-antenna systems, wireless communications, diversity, high data rate, multiple-access communications.

*This research is supported in part by NSF Grant No. ECS-9979448.

1 Introduction

Use of antenna arrays at both base stations and mobile handsets is envisioned in future wireless communication systems [1]. It is well-known that availability of multiple antennas at the transmitter and receiver significantly increases link capacity. Consequently, exploitation of spatio-temporal diversity has emerged as a key technology in state-of-the-art systems. For example, an antenna array is required at the base station in the third generation WCDMA (wideband code-division multiple access) standard [2].

Most existing space-time techniques assume channel state information (CSI) at the receiver (see, e.g., [3]). However, availability of CSI at the transmitter can be exploited to attain improved performance as demonstrated by some recent works (see, e.g., [4, 5]). Channel information may be obtained at the transmitter via several means. In time-division duplexing (TDD), the uplink and downlink channels are reciprocal so the transmitter can estimate the channel using pilot and/or data symbols transmitted by the receiver. In frequency-division duplexing (FDD), a feedback channel may be used to relay channel information estimated by the receiver back to the transmitter.

In this paper, we investigate efficient signaling schemes and receiver designs that exploit CSI at both the transmitter and receiver. This work builds on our earlier results in [6], in which minimum bit-error rate (BER) signaling schemes are derived for single-user system given different types of channel information. When perfect CSI is available, the BER-optimal scheme can be viewed as a generalization of selection diversity via adaptive frequency hopping with spatial beamforming. This scheme uses only one channel dimension in a particular symbol duration. The results in [6] are extended in this paper to exploit all channel dimensions. An orthogonal decomposition of the general L-path wideband space-time channel is derived assuming P-transmit and Q-receive antennas. The channel modes are shown to be outer products of discrete-time sinusoids (representing temporal dimensions) and spatial beamformers (representing spatial dimensions). The decomposition provides a framework for efficiently managing the degrees of freedom in the space-time channel to optimize BER and/or throughput in single-user and multiuser applications.

We consider a fixed phase modulation scheme for concreteness. The throughput of the system is determined by the number of distinct data streams. Our results show that there is a trade-off between throughput and BER – higher throughput generally results in a higher BER. Using only the most dominant channel mode to transmit a single data stream

represents one (minimum BER) extreme. On the other hand, transmitting distinct data streams on all modes represents the other (maximum throughput) extreme. For a given fixed throughput, we derive a power allocation scheme that minimizes the BER under a total power constraint. Given a worst-case BER constraint, we propose a strategy for maximizing the average throughput under a total power constraint.

We also propose a multiuser scheme that employs different sinusoids or temporal dimensions for different users to attain perfect user separation without any processing at the receiver. Furthermore, the signaling and receiver design for each user do not require the channel or signaling information of other users. This scheme is analogous to methods employed in multiuser multicarrier (OFDMA) systems (see, e.g., [7]).

Prior work closely related to the results in this paper is found in [5], in which the singular value decomposition (SVD) of the overall space-time channel is exploited for maximum throughput application in narrow band systems. The SVD in [5] does not admit a closed-form expression. In this paper, we obtain a closed-form SVD for each channel coupling all the transmit antennas to a particular receive antenna. These closed-form SVDs then yield a lower dimensional eigendecomposition for the channel after coherent combining across all receive antennas. We demonstrate that our approach requires lower computation complexity. Another related idea is adaptive multicarrier modulation for temporal channels [8], which is an extension of the original adaptive modulation approach proposed in [4]. Basically, adaptive multicarrier modulation utilizes CSI at the transmitter to dynamically choose the modulation scheme for each group of sub-bands.

The rest of the paper is organized as follows. The space-time channel model and orthogonal decomposition are given in Sections 2 and 3, respectively. Signaling and receiver designs for single-user systems that trade throughput for BER are outlined in Section 4 and 5, respectively. We discuss multiuser system design within our framework in Section 6, followed by the concluding remarks in Section 7.

Superscript T, H, and $*$ indicate matrix transpose, matrix conjugate transpose, and complex conjugation, respectively. Uppercase boldface denotes a matrix while lowercase boldface indicates a vector. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix. $\mathbf{x} \sim \mathcal{N}_C[\mathbf{m}, \mathbf{R}]$ denotes a complex Gaussian vector $\mathbf{x}$ with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. Expectation is denoted as E[.] and the Euclidean norm of vector $\mathbf{x}$ is denoted as $\|\mathbf{x}\|$. The symbol $\otimes$ denotes Kronecker product and vec($\mathbf{A}$) is formed by stacking the columns of matrix $\mathbf{A}$ into a vector

[9, 10]. $\mathbf{e}_n$ is the column vector $[\ \mathbf{0}^T\ 1\ \mathbf{0}^T]^T$ with 1 located at the n-th row. The diagonal matrix generated by the vector $\mathbf{v}$ is denoted by $\mathbf{diag}\{\mathbf{v}\}$. The i-th largest eigenvalue and the corresponding eigenvector of matrix $\mathbf{A}$ is denoted by $\lambda_i[\mathbf{A}]$ and $\mathbf{ev}_i[\mathbf{A}]$, respectively, and we assume $\lambda_i[\mathbf{A}] \geq \lambda_{i+1}[\mathbf{A}]$. In this paper, since $\mathbf{A}$ is always Hermitian symmetric and non-negative definite, $\lambda_i[\mathbf{A}] \geq 0$. The following assumptions are used:

1. The maximum transit time across the array is small compared to the inverse bandwidth of the signal. We also assume that the channel coefficients corresponding to different paths and antennas are not completely correlated.

2. Perfect estimates of CSI are available at the transmitter and receiver.

2 Wideband Space-time Channel Model

Consider a single-user system with P transmit and Q receive antennas. We assume that each component of the transmitted signal $\mathbf{x}(t) = [\ x_1(t)\ ,\ x_2(t)\ ,\cdots,\ x_P(t)\ ]^T$ has duration T and two-sided essential bandwidth B. The transmitted signal $\mathbf{x}(t)$ propagates over an L-path P-input, Q-output fading channel with delays $\{\tau_1, \tau_2, \cdots, \tau_L\}$. Without loss of generality, assume $\tau_1 < \tau_2 < \cdots < \tau_L$. We consider a typical wideband multipath channel where $\tau_L \ll T$. Hence, intersymbol interference is negligible and the baseband signal at the receiver $\mathbf{r}(t) \in \mathbb{C}^Q$ over one symbol interval $0 \leq t < T$ can be written as

$$\begin{aligned} \mathbf{r}(t) &= \begin{bmatrix} r_1(t) \\ \vdots \\ r_Q(t) \end{bmatrix} = \sum_{l=1}^{L} \begin{bmatrix} \mathbf{h}_{l1}^T \\ \vdots \\ \mathbf{h}_{lQ}^T \end{bmatrix} \mathbf{x}(t-\tau_l) + \mathbf{n}(t)\ , \\ \mathbf{h}_{lq}^T &= \begin{bmatrix} h_{lq1} & \cdots & h_{lqP} \end{bmatrix}\ . \end{aligned} \tag{1}$$

The channel coefficient h_{lqp} corresponds to the l-th path between the q-th receive and p-th transmit antenna. The additive noise vector $\mathbf{n}(t)$ is assumed to be zero mean circularly symmetric complex Gaussian with $E[\mathbf{n}(t)\mathbf{n}^H(t')] = \sigma^2\delta(t-t')\mathbf{I}_Q$.

The signal $r_q(t)$ at the q-th receive antenna in (1) can be written as

$$\begin{aligned} r_q(t) &= \sum_{l=1}^{L} \mathbf{h}_{lq}^T \mathbf{x}(t-\tau_l) + n_q(t) \\ &= \begin{bmatrix} \mathbf{x}^T(t-\tau_1) & \cdots & \mathbf{x}^T(t-\tau_L) \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{h}_{1q} \\ \vdots \\ \mathbf{h}_{Lq} \end{bmatrix}}_{\mathbf{h}_q} + n_q(t)\ . \end{aligned} \tag{2}$$

We assume for a single transmitted data stream b that the p-th antenna waveform $x_p(t)$ has the following form:

$$x_p(t) = \sqrt{\rho}\, b \sum_{i=0}^{N-1} s_p[i]\omega(t - i/B)\ , \quad 0 \le t < T \ , \tag{3}$$

where $\omega(t)$ is the (unit-energy) chip waveform of duration $1/B$, ρ is the transmit power, and $N = TB$. Here $s_p[i]$, $i = 0, 1, \cdots, N-1$ represents the signature sequence associated with the p-th antenna. Since $x_p(t)$ is essentially bandlimited to B, assuming τ_l to be an integer multiple of $1/B$ suffices [11]. That is, $\tau_l = d_l/B$, $d_l \in \{0, 1, \cdots, N-1\}$ and $d_L \ll N$. We sample $r_q(t)$ at the rate $\frac{1}{B}$ to enable discrete-time processing without loss of information. Let

$$\begin{aligned} \mathbf{r}_q &\stackrel{def}{=} [r_q(0), r_q(1/B), \cdots, r_q((N-1)/B)]^T \ , \\ \mathbf{s}_p &\stackrel{def}{=} [s_p[0], s_p[1], \cdots, s_p[N-1]]^T \ , \\ \mathbf{S} &\stackrel{def}{=} [\ \mathbf{s}_1\ \cdots\ \mathbf{s}_P\ ] \ . \end{aligned} \tag{4}$$

Hence, $\mathbf{r}_q$ contains samples of the received signal at the q-th antenna over one symbol duration, while $\mathbf{S}$ is an $N \times P$ matrix containing the signature codes from all transmit antennas. Now, define $\mathbf{\Delta}_{d_l} \in \mathbb{C}^{N\times N}$ as the time-shift matrix corresponding to the path delay d_l. We assume the delay is cyclic, so that $\mathbf{\Delta}_{d_l}$ is circulant. For example, when $d_l = 1$ and $N = 3$, $\mathbf{\Delta}_{d_l} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$. While the actual delay corresponds to a linear shift, negligible error is introduced by this assumption for sufficiently large N provided that $d_L \ll N$. Furthermore, if a *chip-level* cyclic prefix is introduced, the cyclic shift is exact [11]. By defining $\mathbf{\Delta} \stackrel{def}{=} [\mathbf{\Delta}_{d_1}\ \cdots\ \mathbf{\Delta}_{d_L}] \in \mathbb{C}^{N\times NL}$, $\mathbf{s} = \mathrm{vec}\,(\mathbf{S}) \in \mathbb{C}^{NP}$, $\mathbf{H}_q \stackrel{def}{=} [\mathbf{h}_{1q}\ \cdots \mathbf{h}_{Lq}] \in \mathbb{C}^{P\times L}$, and $\mathbf{h}_q = \mathrm{vec}\,(\mathbf{H}_q)$, we have from (2) – (3)

$$\mathbf{r}_q = \sqrt{\rho}\, b\mathbf{\Delta}\,(\mathbf{I}_L \otimes \mathbf{S})\,\mathbf{h}_q + \mathbf{n}_q\ .$$

Defining $\mathcal{H}_q = \mathbf{\Delta}\left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right)$ and applying the identity $\mathrm{vec}(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{X})$ [9] twice, we get

$$\mathbf{r}_q = \sqrt{\rho}\, b\, \mathcal{H}_q\mathbf{s} + \mathbf{n}_q \tag{5}$$

where $\mathbf{n}_q \sim \mathcal{N}_C\,[\mathbf{0}, \sigma^2\mathbf{I}_N]$.

3 Space-Time Channel Decomposition

The overall space-time channel may be represented as $\mathcal{H} \stackrel{def}{=} \left[\begin{array}{ccc} \mathcal{H}_1^T & \cdots & \mathcal{H}_Q^T \end{array} \right]^T \in \mathbb{C}^{NQ \times NP}$. The number of available space-time dimensions, N_{dim}, is precisely the rank of $\mathcal{H}$. Since we assume that the channel coefficients $\{h_{lqp}\}$ are not perfectly correlated (assumption 2 in Section 1), $N_{dim} = N \times \min(P, Q)$ w.p.1. Our goal is to design transceivers that access all N_{dim} degrees of freedom in a way that different channel modes do not interfere with one another. This goal can be accomplished when a singular value decomposition (SVD) of $\mathcal{H}$ is available, analogous to [5]. This SVD has to be computed numerically since for $Q > 1$, a closed-form SVD for $\mathcal{H}$ can not be obtained. Numerical computation is prohibitive in practice since N is usually large for wideband applications (≥ 32).

In this paper, instead of using an SVD for $\mathcal{H}$, we derive a closed-form SVD for $\mathcal{H}_q$ and show that via appropriate signaling and receiver designs, all N_{dim} degrees of freedom can be accessed via non-interfering modes. The circulant structure of $\mathbf{\Delta}_{d_l}$ is exploited to obtain a closed-form SVD for the q-th receive antenna space-time channel matrix $\mathcal{H}_q$ in (5).

Theorem 1 *Define*

$$\mathbf{c}_n = \frac{1}{\sqrt{N}} \left[\begin{array}{cccc} 1 & e^{-j2\pi n/N} & \cdots & e^{-j2\pi(N-1)n/N} \end{array} \right]^T \in \mathbb{C}^N \ , \tag{6}$$

$$\mathbf{g}_{n,q} = \frac{1}{\sqrt{N}} \left[\begin{array}{cccc} \sum_{l=1}^L h_{lq1} e^{-j2\pi n d_l/N} & \cdots & \sum_{l=1}^L h_{lqP} e^{-j2\pi n d_l/N} \end{array} \right]^H \in \mathbb{C}^P \ . \tag{7}$$

Then, $\mathcal{H}_q \in \mathbb{C}^{N \times NP}$ admits the following SVD*:*

$$\mathcal{H}_q = \sum_{n=0}^{N-1} \sigma_{n,q} \mathbf{c}_n \mathbf{v}_{n,q}^H \ , \tag{8}$$

$$\sigma_{n,q} = \|\mathbf{g}_{n,q}\| \quad , \quad \mathbf{v}_{n,q} = \frac{\mathbf{g}_{n,q}}{\|\mathbf{g}_{n,q}\|} \otimes \mathbf{c}_n \tag{9}$$

Proof: Appendix A.

Notice that the left singular vectors $\{\mathbf{c}_n\}_{n=0}^{N-1}$ are independent of q since they are associated with temporal channel characteristics. Also, notice that the p-th row of $\mathbf{g}_{n,q}$ is the complex conjugate of the frequency response of the channel between the q-th receive and p-th transmit antenna at frequency $\frac{2\pi}{N}n$.

Consider implementing the maximum likelihood or maximum ratio combining (MRC) receiver for the data b. We may decompose the MRC receiver into two stages: front-end

matched filtering with *only* the channel coefficients and combining across all received antennas, as denoted by $\tilde{\mathbf{r}} \stackrel{def}{=} \sum_{q=1}^{Q} \mathcal{H}_q^H \mathbf{r}_q$, followed by matched filtering with the signature code $\mathbf{s}$. We shall study the NP-dimensional vector $\tilde{\mathbf{r}}$ as it contains the channel coefficients. Substitute for $\mathbf{r}_q$ to obtain

$$\tilde{\mathbf{r}} = \sqrt{\rho}\, b \left(\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q \right) \mathbf{s} + \sum_{q=1}^{Q} \mathcal{H}_q^H \mathbf{n}_q \tag{10}$$

Now apply Theorem 1 and the identity $(\mathbf{X}_1 \otimes \mathbf{X}_2)(\mathbf{Y}_1 \otimes \mathbf{Y}_2) = \mathbf{X}_1\mathbf{Y}_1 \otimes \mathbf{X}_2\mathbf{Y}_2$ to show

$$\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q = \sum_{n=0}^{N-1} \mathbf{\Gamma}_n \otimes \mathbf{c}_n \mathbf{c}_n^H \in \mathbb{C}^{PN \times PN} \tag{11}$$

where

$$\mathbf{\Gamma}_n = \sum_{q=1}^{Q} \mathbf{g}_{n,q} \mathbf{g}_{n,q}^H \in \mathbb{C}^{P \times P} \tag{12}$$

is the overall spatial mode matrix at frequency $\frac{2\pi}{N} n$. Since the rank of $\mathbf{\Gamma}_n$ in (12) is $\min(P, Q)$, it is clear from (11) that $\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q$ is of rank $N_{dim} = N \times \min(P, Q)$ w.p.1. Also, $\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q = \mathcal{H}^H \mathcal{H}$ is the Grammian of $\mathcal{H}$, hence $\text{rank}(\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q) = \text{rank}(\mathcal{H})$ [10]. It is easy to verify that the N_{dim} eigenvectors of $\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q$ are

$$\begin{aligned} &\mathbf{w}^{((i-1)N+n)} \otimes \mathbf{c}_n \quad , \quad n \in \{0, 1, \cdots, N-1\} \quad , \quad i \in \{1, 2, \cdots, \min(P, Q)\} \\ &\mathbf{w}^{((i-1)N+n)} = \mathbf{ev}_i \left[\mathbf{\Gamma}_n \right] \quad . \end{aligned} \tag{13}$$

with the corresponding eigenvalues $\gamma_{(i-1)N+n+1} = \lambda_i \left[\mathbf{\Gamma}_n \right]$. These N_{dim} eigenmodes represent all the available *non-interfering space-time sub-channels* for any MRC based receiver. While a single data stream was assumed to motivate the MRC receiver structure, in general N_{dim} non-interfering data streams can be transmitted in parallel using the N_{dim} eigenmodes of $\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q$. Each of these streams can be demodulated using the MRC receiver described above. This issue will be discussed in Section 5.

Notice that to compute all the eigenvectors in (13), the most costly operation is finding the eigenvectors of N matrices, each with the size of $P \times P$. This is much less complex than computing the singular modes of the $NQ \times NP$ matrix $\mathcal{H}$ since N is typically large (≥ 32) and P is typically small (currently 2) in practice.

4 Minimum BER Single-User System

Minimum BER is obtained by transmitting only a single data stream via the most dominant sub-channel [6]. In this case, we choose

$$\mathbf{s} = \mathbf{w} \otimes \mathbf{c}_{\overline{n}} \tag{14}$$

where

$$\mathbf{w} = \mathbf{ev}_1 \left[\mathbf{\Gamma}_{\overline{n}}\right] \quad , \quad \overline{n} = \arg \max_{n=0,\cdots,N-1} \lambda_1 \left[\mathbf{\Gamma}_n\right] \quad . \tag{15}$$

That is, $\overline{n}$ is the frequency of the spatial mode matrix $\mathbf{\Gamma}_n$ with the largest dominant eigenvalue and $\mathbf{w}$ is the corresponding dominant eigenvector. Notice that only *one* dimension is used in any one symbol duration to achieve minimum BER. This signaling scheme can be implemented as shown in Figure 1(a) with $n = \overline{n}$.

For receiver design, we assume BPSK modulation ($b \in \{\pm 1\}$), although extension to other modulation schemes is straightforward. The MRC receiver computes $\tilde{\mathbf{r}}$ in (10) and then correlates $\tilde{\mathbf{r}}$ with $\mathbf{s} = \mathbf{w} \otimes \mathbf{c}_{\overline{n}}$ to obtain the decision statistic Z for b. To simplify receiver complexity, we exploit Theorem 1 as follows. It is easy to show using the identity $(\mathbf{X}_1 \otimes \mathbf{X}_2)(\mathbf{Y}_1 \otimes \mathbf{Y}_2) = \mathbf{X}_1\mathbf{Y}_1 \otimes \mathbf{X}_2\mathbf{Y}_2$ and the orthogonality of $\{\mathbf{c}_n\}_{n=0}^{N-1}$ that

$$\mathcal{H}_q \left(\mathbf{w} \otimes \mathbf{c}_{\overline{n}}\right) = \left(\mathbf{g}_{\overline{n},q}^H \mathbf{w}\right) \mathbf{c}_{\overline{n}} \tag{16}$$

Hence, Z can be written as follows:

$$Z = \left(\mathbf{w} \otimes \mathbf{c}_{\overline{n}}\right)^H \sum_{q=1}^{Q} \mathcal{H}_q^H \mathbf{r}_q = \sum_{q=1}^{Q} \left(\mathbf{w}^H \mathbf{g}_{\overline{n},q}\right) \mathbf{c}_{\overline{n}}^H \mathbf{r}_q \; . \tag{17}$$

This can be implemented as shown in Figure 1(b) with $n = \overline{n}$. Note that $\{\mathbf{g}_{n,q}\}$, which represents the CSI, and the transmit beamformer $\mathbf{w}$ need to be known at the receiver. For BPSK modulation, the maximum likelihood detector is $\hat{b} = \text{sgn}\left(\text{Re}\{Z\}\right)$. In this case, [6][1]

$$BER_{min} = \mathcal{Q}\left(\sqrt{\frac{2\rho}{\sigma^2} \mathbf{w}^H \left(\sum_{q=1}^{Q} \mathbf{g}_{\overline{n},q} \mathbf{g}_{\overline{n},q}^H\right) \mathbf{w}}\right) \; . \tag{18}$$

Choosing $\mathbf{s} = \mathbf{w} \otimes \mathbf{c}_{\overline{n}}$ with $\mathbf{w}$ defined in (15) maximizes the argument of $\mathcal{Q}(\sqrt{\cdot})$ and therefore the BER is minimized. Note that the effective SNR is $\frac{\rho}{\sigma^2}\lambda_1\left[\mathbf{\Gamma}_{\overline{n}}\right]$, so $\lambda_1\left[\mathbf{\Gamma}_{\overline{n}}\right]$ is the gain of the dominant sub-channel.

[1] $\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} \, du$.

We assume the availability of $\{\mathbf{g}_{n,q}\}$, $\overline{n}$, and $\mathbf{w}$ at both the transmitter and receiver. In practice, this availability depends on the system constraints. For TDD systems, $\{\mathbf{g}_{n,q}\}$, $\overline{n}$, and $\mathbf{w}$ can be computed at both the transmitter and receiver by exploiting reciprocity. Alternatively, $\overline{n}$ and $\mathbf{w}$ can be computed at the transmitter and signaled to the receiver. $\{\mathbf{g}_{n,q}\}$ can always be estimated at the receiver for front-end processing. In FDD systems, reciprocity does not hold, hence CSI needs to be signaled to the transmitter via a feedback channel. In order to reduce feedback overhead, the receiver may compute $\overline{n}$ and $\mathbf{w}$ and feed them back to the transmitter, rather than feeding back CSI.

5 High Throughput Single-User Systems

Since we assume a fixed modulation scheme, the throughput of the system is determined by the number of streams transmitted simultaneously. To transmit M data streams via the channel, we choose a transmitted signal of the form

$$\sum_{m=1}^{M} \sqrt{\rho_m}\, b_m\, \mathbf{s}^{(m)}\ , \tag{19}$$

where the signature codes $\{\mathbf{s}^{(m)}\}$ are chosen to be a subset of the eigenvectors of $\sum_{q=1}^{Q} \mathcal{H}_q^H \mathcal{H}_q$ given in (13). Define the *relative throughput* of a system as the total system throughput relative to the throughput of a one-stream system with the same modulation scheme. The relative throughput M is bounded by N_{dim}, the maximum number of parallel sub-channels. Without loss of generality, we assume that $\sigma^2 = 1$ in this section. That is, ρ_m is the transmit power for the m-th stream normalized with respect to the noise variance σ^2. At the receiver, different streams are easily separated due to the orthogonality of $\{\mathbf{s}^{(m)}\}$. The transmitter and receiver for this maximum throughput scheme may be implemented as shown in Figure 1(a) and 1(b), respectively, for each data stream with n and $\mathbf{w}$ chosen accordingly.

Throughput is maximized by using all the N_{dim} spatio-temporal dimensions for data transmission as discussed above. However, for a fixed total transmitted power, this comes at the expense of BER since the power has to be distributed between N_{dim} streams. Hence, there is a trade-off between throughput and BER. One may trade throughput for lower BER by choosing to transmit with $M < N_{dim}$ data streams. Since perfect CSI is available at the transmitter, the sub-channel gains (eigenvalues) $\{\gamma_m\}_{m=1}^{N_{dim}}$ defined in Section 3 can be

determined. Without loss of generality, assume that

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_M \geq \gamma_{M+1} \geq \cdots \geq \gamma_{N_{dim}} > 0 \ . \tag{20}$$

where the non-zero gain assumption is achieved w.p.1 by Assumption 2 in Section 1. Clearly, the most power-efficient way to achieve a relative throughput of M is to use the M sub-channels with the highest gains. We define the *effective* BER of an M-stream system as:

$$BER_{eff}^{(M)} \stackrel{def}{=} \frac{1}{M} \sum_{m=1}^{M} BER\left(\rho_m \gamma_m\right) \ . \tag{21}$$

where $\rho_m \gamma_m$ is the received SNR corresponding to the m-th stream. $BER\left(\rho_m \gamma_m\right)$ is a strictly decreasing function of the received SNR, and depends on the chosen modulation scheme. The effective BER reflects the average system performance across M sub-channels. The transmit power allocated for all streams $\{\rho_m\}_{m=1}^{M}$ is assumed to satisfy the constraint $\sum_{m=1}^{M} \rho_m = \rho_{TOT}$. In general, ρ_m is chosen based on the sub-channel gains $\{\gamma_m\}_{m=1}^{M}$. Note that to achieve relative throughput of M, we require $\rho_m > 0$ for all $m \in \{1, 2, \cdots, M\}$.

We now investigate a power allocation scheme that minimizes the effective BER for a fixed throughput (M) in Section 5.1. This scheme further leads to a strategy to maximize the instantaneous throughput for a given worst-case BER requirement, as discussed in Section 5.2.

5.1 Fixed Throughput Criterion

We assume BPSK or QPSK modulation, so that $BER(\rho_m \gamma_m) = \mathcal{Q}\left(\sqrt{2\rho_m \gamma_m}\right)$. The results presented below can be extended to other modulation schemes. For a given relative throughput of M, we choose $\{\rho_m\}_{m=1}^{M}$ to minimize $BER_{eff}^{(M)}$. Since the effective BER reflects the average performance over all sub-channels, it may result in some sub-channels with extremely high and some with extremely low received SNR $\rho_m \gamma_m$ being used for data transmission. This effect is more pronounced when the total transmit power ρ_{TOT} is low. To prevent this, we impose a worst-case SNR constraint resulting in the following optimization problem:

$$\{\overline{\rho}_m\}_{m=1}^{M} = \arg \min_{\rho_1, \cdots, \rho_M} \sum_{m=1}^{M} \mathcal{Q}\left(\sqrt{2\rho_m \gamma_m}\right) \tag{22}$$

$$\text{s.t.} \quad \sum_{m=1}^{M} \rho_m = \rho_{TOT} \ , \tag{23}$$

$$\rho_m \gamma_m \geq c_m \ , \quad m = 1, \cdots, M \ . \tag{24}$$

The constant c_m is chosen such that $\mathcal{Q}\left(\sqrt{2c_m}\right)$ is the worst-case BER for sub-channel m.

The above optimization problem can be solved via the Kuhn-Tucker conditions [12]. It is shown in Appendix B that:

1. A solution exists if and only if

$$\rho_{TOT} \geq \rho_{co,M} \quad , \quad \rho_{co,M} = \sum_{m=1}^{M} \frac{c_m}{\gamma_m} \; . \tag{25}$$

   where $\rho_{co,M}$ denotes the *cut-off* transmit power for a relative throughput of M. When (25) is met, the solution is unique and characterized by:

$$\overline{\rho}_m = \max\left(\frac{c_m}{\gamma_m}, \rho_m\right) \quad , \tag{26}$$

   where ρ_m satisfies

$$\sqrt{\frac{\gamma_m}{\rho_m}} \times \exp\left(-\rho_m \gamma_m\right) = \overline{\mu} \; , \quad m = 1, \cdots, M \tag{27}$$

   and $\overline{\mu}$ is chosen such that $\sum_{m=1}^{M} \overline{\rho}_m = \rho_{TOT}$.

2. For a given total transmit power ρ_{TOT} and sub-channel SNR values $\{\gamma_m\}_{m=1}^{N_{dim}}$, the minimum effective BER power allocation results in

$$BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)} \; . \tag{28}$$

   provided that $\{c_m\}_{m=1}^{M}$ is a constant or decreasing sequence.

Hence, to maintain a relative throughput of M for different channel realizations, the total power ρ_{TOT} may need to be adjusted accordingly. The property represented by (28) demonstrates the trade-off between BER and throughput. That is, higher throughput results in higher effective BER. Note that although this property is intuitively pleasing, it is generally not true for arbitrary power allocation schemes.

The solution of (27) must be obtained numerically. The derivation in Appendix B suggests an iterative procedure to obtain the solution of (26) and (27). Starting with an arbitrary $\overline{\mu}$, (27) is solved numerically for $\{\overline{\rho}_m\}_{m=1}^{M}$. A unique solution is guaranteed for any value of $\overline{\mu}$. If $\sum_{m=1}^{M} \overline{\rho}_m > \rho_{TOT}$, $\overline{\mu}$ is increased to reduce each $\overline{\rho}_m$. Similarly, if $\sum_{m=1}^{M} \overline{\rho}_m < \rho_{TOT}$, μ is lowered to increase each $\overline{\rho}_m$. This procedure is repeated until $\sum_{m=1}^{M} \rho_m = \rho_{TOT}$ is satisfied within a prescribed numerical tolerance.

An approximate solution of (22) can be obtained by replacing the exact BER for each sub-channel in (22) with its Chernoff bound. In this case, we minimize the following upper bound on the effective BER:

$$BER_{eff}^{(M)} \leq \frac{1}{2M} \sum_{m=1}^{M} \exp\left(-\rho_m \gamma_m\right) \ .$$

This approximation is accurate for sufficiently large ρ_{TOT}. Again, using Kuhn-Tucker conditions as in Appendix B, we obtain the following closed-form solution for ρ_m, $m = 1, \cdots, M$ assuming (25) holds:

$$\tilde{\rho}_m = \frac{\max\left(c_m, \log \gamma_m - \tilde{\mu}\right)}{\gamma_m} \tag{29}$$

where $\tilde{\mu}$ is chosen to satisfy the power constraint $\sum_{m=1}^{M} \tilde{\rho}_m = \rho_{TOT}$. We term this solution the *Chernoff-based* power allocation. Analogous to the exact solution, it can be shown that $BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)}$ holds in this case as well.

When ρ_{TOT} is sufficiently large, it is easy to see that the worst-case sub-channel BER constraint in (24) is not active. In this case, the exact solution to (22) is $\overline{\rho}_m = \rho_m$ for all m, where ρ_m is given in (27). The Chernoff-based solution is simply $\tilde{\rho}_m = \frac{\log \gamma_m - \tilde{\mu}}{\gamma_m}$.

Another simple sub-optimal power allocation scheme that satisfies the constraints in (23) and (24) assuming (25) holds can be obtained as follows:

$$\hat{\rho}_m = \frac{c_m}{\gamma_m} + \frac{1}{M}\left(\rho_{TOT} - \rho_{co,M}\right) \ . \tag{30}$$

That is, after satisfying the minimum SNR constraint in each sub-channel, the remaining power is distributed equally for all sub-channels. We term this scheme *uniform* power allocation. It is easy to see that $BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)}$ holds for uniform power allocation.

5.2 Maximum Throughput Criterion

We now consider an adaptive throughput scheme where the instantaneous relative throughput is maximized subject to constraints (23) and (24). Let the set of "allowable" relative throughput be $\mathcal{M}$ with $0 \in \mathcal{M}$. This is intended to allow "no-transmission" when the channel undergoes such deep fades that the BER requirement cannot be achieved for a given ρ_{TOT}. The solution of this problem is simply choosing the largest M such that

$$\rho_{co,M} \leq \rho_{TOT} \tag{31}$$

still holds for each channel realization.

Note that the maximum throughput criterion is not coupled to any sub-channel power allocation scheme but only requires $\rho_m \geq \frac{c_m}{\gamma_m}$. Thus, one may use the minimum effective BER or uniform allocation scheme described above to choose the ρ_m.

5.3 Examples

For all the examples, we consider a $P = Q = 2$ and $N = 16$ system. There are $L = 3$ paths with $d_l \in \{0, 1, 2\}$. The channel coefficients $\{h_{lqp}\}$ are assumed IID and $\mathcal{N}_C[0, 1/QL]$ (Rayleigh fading). To illustrate the fixed throughput criterion, we consider $M = 32$. The system is required to achieve the same worst-case BER of ε on each sub-channel. Hence,

$$c_m = \left(\mathcal{Q}^{-1}(\varepsilon)\right)^2 /2 \ , \quad m = 1, \cdots, M$$

where $\mathcal{Q}^{-1}(x)$ is the inverse of $\mathcal{Q}(x)$.

We first compare the minimum effective BER power allocation (based on the exact and Chernoff-bounded effective BER) to uniform power allocation. One channel realization and $\varepsilon = 10^{-2}$ are used . The resulting sorted sub-channel SNR values and the cut-off transmit power $\rho_{co,M}$ as a function of M are depicted in Figure 2(a) and 2(b), respectively. For $M = 32$, which implies $\rho_{TOT} \geq 33.54$ dB is required to satisfy the worst-case BER constraint. The allocation of power and resulting BER across sub-channels for $\rho_{TOT} = 33.6$ and 37 dB are depicted in Figure 3(a)–(d). From this example, we make the following observations:

1. The Chernoff-based solution is virtually identical to the exact minimum effective BER solution. The difference between the minimum effective BER and uniform power allocation schemes is small when ρ_{TOT} is close to the minimum value 33.54 dB. The difference is more pronounced when excess power is available.

2. Both the minimum BER and uniform power allocation schemes allocate relatively more power to sub-channels with low gain. However, as evident from Figure 3 (b) and (d), the *received* SNR $\rho_m \gamma_m$ is largest in the sub-channels with the largest channel gain γ_m. Notice that the minimum effective BER solution tends to allocate more power to sub-chanels with low channel gain than does the uniform scheme, which results in lower effective BER.

3. The worst-case sub-channel BER constraint in (24) is active only for low ρ_{TOT}. This is evident from Figure 3 (c) and (d). The sub-channel BER values for $\rho_{TOT} = 37$ dB fall below 10^{-2}, which indicate that $\rho_m \gamma_m > c_m$ for all m.

The effective BER for different M is displayed in Figure 4(a) as a function of ρ_{TOT} using the same channel realization and exact minimum effective BER solution. A comparison to the effective BER obtained using Chernoff bound and uniform power allocations is shown in Figure 4(b). Observe that the loss of performance due to uniform power allocation compared to the minimum effective BER solution is more pronounced as M increases. Also, the Chernoff approximation introduces negligible performance loss.

To demonstrate the notion of adaptive throughput introduced in Section 5.2, we assume $\varepsilon = 10^{-2}$ and 10^{-4} worst-case BER requirements and use the Chernoff-bound power allocation scheme. Four different sets of allowable relative throughputs are used:

$$\begin{array}{ll} \mathcal{M}_1 = \{0, 2, 8, 32\} & (4 \text{ levels}) \\ \mathcal{M}_2 = \{0, 2, 4, 8, 16, 32\} & (6 \text{ levels}) \\ \mathcal{M}_3 = \{0, 1, 2, 4, 6, 8, \cdots, 30, 32\} & (18 \text{ levels}) \\ \mathcal{M}_4 = \{0, 1, 2, 3, \cdots, 31, 32\} & (33 \text{ levels}) \ . \end{array}$$

The resulting average relative throughput and average BER_{eff} are depicted in Figure 5 (a)–(d). The averages were computed over 500 channel realizations. Observe that larger sets result in better average throughput for any ρ_{TOT} and the resulting BER profiles are closer to the worst-case requirement. With small sets excess power tends to reduce the effective BER rather than increase the number of channels, while with the larger sets increases in ρ_{TOT} tend to increase the number of channels, rather than reduce average BER_{eff}. Also, decreasing the worst-case BER results in a decrease in throughput as more power is needed to achieve a certain throughput.

6 Multiuser Systems

For a general multistream system with K active users, the sampled signal at the q-th receive antenna $\mathbf{r}_q$ in (5) can be written as

$$\mathbf{r}_q = \sum_{k=0}^{K-1} \mathcal{H}_q^{(k)} \sum_{m=1}^{M_k} \sqrt{\rho_{m,k}}\, b_{m,k} \mathbf{s}^{(m,k)} + \mathbf{n}_q \ , \tag{32}$$

where $\mathcal{H}_q^{(k)} = \mathbf{\Delta}^{(k)} \left(\mathbf{H}_q^{(k)T} \otimes \mathbf{I}_N \right)$ and the index k denotes the k-th user. Here the k-th user transmits M_k data streams. Note that in general the signal transmitted by different users

see different channels. It is easy to show that the SVD of $\mathcal{H}_q^{(k)} \in \mathbb{C}^{N\times NP}$ in Theorem 1 can be written as

$$\mathcal{H}_q^{(k)} = \sum_{n=0}^{N-1} \mathbf{c}_n \left(\mathbf{g}_{n,q}^{(k)} \otimes \mathbf{c}_n\right)^H \ , \tag{33}$$

where the p-th row of $\mathbf{g}_{n,q}^{(k)}$ is the complex conjugate of frequency response of the k-th user channel between the q-th receive and p-th transmit antenna at frequency $\frac{2\pi}{N}n$.

In this section, we demonstrate that up to N users can be accommodated without resulting in multi-access interference. This is possible by exploiting the left singular vectors $\{\mathbf{c}_n\}_{n=0}^{N-1}$ which are independent of the channels of different users. Choose $\mathbf{s}^{(m,k)} = \mathbf{w}^{(m,\pi(k))} \otimes \mathbf{c}_{\pi(k)}$, where $\{\pi(k)\}_{k=0}^{K-1}$ can be any arbitrary permutation of $\{0, 1, \cdots, K-1\}$. Here, $\pi(k)$ represents the frequency assignment for different users. For simplicity, we choose $\pi(k) = k$. Analogous to (16), we have from (32) and (33)

$$\mathbf{r}_q = \sum_{k=0}^{K-1} \mathbf{c}_k \sum_{m=1}^{M_k} \sqrt{\rho_{m,k}}\, b_{m,k} \left(\mathbf{g}_{k,q}^{(k)H} \mathbf{w}^{(m,k)}\right) + \mathbf{n}_q \ . \tag{34}$$

Since $\{\mathbf{c}_k\}_{k=0}^{K-1}$ are orthogonal, perfect user separation can be achieved without using a decorrelating detector at the receiver. It is apparent that the separation of users is achieved using only the N orthogonal temporal dimensions. In general, spatial dimensions can not be used for separating users because different users generally have different channel coefficients. After this temporal frequency assignment, each user has $\min(P, Q)$ different spatial dimensions. These available spatial dimensions may be used to increase throughput and/or minimize BER by deriving the appropriate spatial beamformers.

Suppose the maximum number of users $K = N$ is desired. The transmitter and receiver structure in Figure 1 may be employed for each user (or each user's data stream) with $\mathbf{c}_n = \mathbf{c}_k$, $\mathbf{w} = \mathbf{w}^{(m,k)}$, and $\mathbf{g}_{n,q} = \mathbf{g}_{k,q}^{(k)}$. Denote the test statistic for data stream m of user k' as $Z^{(m,k')}$, which can be written as

$$Z^{(m,k')} = \sum_{m=1}^{M_{k'}} \sqrt{\rho_{m,k'}}\, b_{m,k'} \mathbf{w}^{(m,k)H} \left(\sum_{q=1}^{Q} \mathbf{g}_{k',q}^{(k')} \mathbf{g}_{k',q}^{(k')H}\right) \mathbf{w}^{(m,k')} + \mathbf{w}^{(m,k')H} \sum_{q=1}^{Q} \mathbf{g}_{k',q}^{(k')} \mathbf{c}_{k'}^H \mathbf{n}_q \ .$$

The spatial beamformer $\mathbf{w}^{(m,k')}$ may be chosen to optimize the k'-th user's BER, throughput, or a combination between the two as discussed in Section 4 and 5.

If the number of users $K < N$, then temporal dimensions can also be used to increase each user's throughput and/or minimize BER. In this case, user k is assigned to $\min(P, Q) \times N_{T,k}$

dimensions, where $N_{T,k}$ is the number of temporal dimensions (frequencies) for user k and $\sum_{k=0}^{K-1} N_{T,k} \leq N$.

As noted, user separation in this framework is a result of the channel's temporal eigenstructure. The temporal eigenstructure is independent of the channel realization and hence, is common to all users. This implies that imperfect CSI at the transmitter and/or receiver will not affect multiuser separation. Each user, however, will incur some performance loss due to imperfection in the spatial beamformer if perfect CSI is not available. This also implies that multiuser separation can be achieved without the availability of CSI at the transmitter. This is analogous to the use of sinusoids for multiuser separation in OFDMA systems.

The perfect multiuser separation property also implies that the signaling and receiver design for each user do not require any channel or signaling information of other users. This greatly simplifies system design. For example, each user's total transmit power $\rho_{TOT}^{(k)}$ may be independently adjusted. Thus, CSI at the transmitter can be utilized to adjust the amount of transmitted power $\{\rho_{1,k}, \cdots, \rho_{M_k,k}\}_{k=0}^{K-1}$ to achieve a certain target BER or SNR for each user.

7 Conclusions

An orthogonal decomposition of a general space-time multi-antenna multipath channel is derived and utilized to design efficient signaling strategies and the corresponding receiver structures. The decomposition explicitly characterizes $N \times \min(P, Q)$ available non-interfering spatio-temporal dimensions in the channel. The time bandwidth product N represents the number of available temporal dimensions. The number of available spatial dimensions is $\min(P, Q)$, where P and Q are the number of transmit and receive antennas, respectively. This decomposition provides a framework to jointly address system design for minimum BER, maximum throughput, multiuser applications, as well as the combination of the three, provided channel state information (CSI) is available at the transmitter. For a fixed throughput system, a power allocation scheme that minimizes the instantaneous effective BER of a multistream transmission is derived. In addition, a strategy to maximize the system throughput given a worst-case BER requirement and is proposed. For multiuser applications, analogous to OFDMA systems, the proposed scheme possesses a perfect multiuser separation property as a result of the common temporal eigenstructure across all channels.

Appendix A: Proof of Theorem 1

Since $\mathbf{\Delta}_{d_l}$ is circulant, $\mathbf{\Delta}_{d_l} = \mathbf{F}\mathbf{\Lambda}_{d_l}\mathbf{F}^H$, where $\mathbf{F} = \left[\begin{array}{cccc} \mathbf{c}_0 & \mathbf{c}_1 & \cdots & \mathbf{c}_{N-1} \end{array} \right] \in \mathbb{C}^{N\times N}$ is the N-DFT matrix and [10]

$$\mathbf{\Lambda}_{d_l} = \mathbf{diag}\left\{\mathbf{F}\mathbf{e}_{N+1-d_l}\right\} = \frac{1}{\sqrt{N}}\,\mathbf{diag}\left\{1, e^{j2\pi d_l/N}, \cdots, e^{j2\pi(N-1)d_l/N}\right\} .$$

Let $\mathbf{D}_{qp} \stackrel{def}{=} \sum_{l=1}^{L} h_{lqp}\mathbf{\Lambda}_{d_l}$. Then, it can be shown that

$$\begin{aligned} \mathcal{H}_q &= \mathbf{\Delta}\left(\mathbf{H}_q^T \otimes \mathbf{I}_N\right) = \left[\begin{array}{ccc} \sum_{l=1}^{L} h_{lq1}\mathbf{\Delta}_{d_l} & \cdots & \sum_{l=1}^{L} h_{lqP}\mathbf{\Delta}_{d_l} \end{array} \right] \\ &= \mathbf{F}\left[\begin{array}{ccc} \mathbf{D}_{q1} & \cdots & \mathbf{D}_{qP} \end{array} \right] (\mathbf{I}_P \otimes \mathbf{F})^H . \end{aligned} \tag{35}$$

Using the same argument as in [6], it can be shown that

$$\left[\begin{array}{ccc} \mathbf{D}_{q1} & \cdots & \mathbf{D}_{qP} \end{array} \right] = \left[\begin{array}{ccc} \|\mathbf{g}_{0,q}\| & & \\ & \ddots & \\ & & \|\mathbf{g}_{N-1,q}\| \end{array} \right] \left[\begin{array}{ccc} \mathbf{g}_{0,q}^H/\|\mathbf{g}_{0,q}\| & & \\ & \ddots & \\ & & \mathbf{g}_{N-1,q}^H/\|\mathbf{g}_{N-1,q}\| \end{array} \right] \mathbf{\Pi}_{(P,N)}^H , \tag{36}$$

where $\mathbf{g}_{n,q}$ is given in (7) and $\mathbf{\Pi}_{(P,N)} = \left[\begin{array}{cccc} \mathbf{I}_P \otimes \mathbf{e}_1 & \mathbf{I}_P \otimes \mathbf{e}_2 & \cdots & \mathbf{I}_P \otimes \mathbf{e}_N \end{array} \right] \in \mathbb{C}^P$ is the (P, N) unitary permutation matrix [9]. Hence,

$$(\mathbf{I}_P \otimes \mathbf{F})\,\mathbf{\Pi}_{(P,N)} = \left[\begin{array}{cccc} \mathbf{I}_P \otimes \mathbf{F}\mathbf{e}_1 & \cdots & \mathbf{I}_P \otimes \mathbf{F}\mathbf{e}_N \end{array} \right] = \left[\begin{array}{cccc} \mathbf{I}_P \otimes \mathbf{c}_0 & \cdots & \mathbf{I}_P \otimes \mathbf{c}_{N-1} \end{array} \right] , \tag{37}$$

where $\mathbf{c}_n$ is defined in (6) and the first equality follows from the identity

$$(\mathbf{X}_1 \otimes \mathbf{X}_2)\,(\mathbf{Y}_1 \otimes \mathbf{Y}_2) = \mathbf{X}_1\mathbf{Y}_1 \otimes \mathbf{X}_2\mathbf{Y}_2 , \tag{38}$$

with appropriate dimensions. The proof is completed by combining (35), (36), and (37).

Appendix B: Minimum BER_{eff} Solution

In this appendix, we denote the power parameter as ρ_m and the optimal solution as $\overline{\rho}_m$. The same holds for the Lagrange multipliers μ and ν_m (defined later). The superscript (M) for $\rho_m^{(M)}$ will be suppressed when the context is clear.

Claim 1:

If (25) is not satisfied, the constraints in (23) and (24) are not feasible. If (25) is satisfied, the constraints in (23) and (24) define a compact set in $\{\rho_m\}_{m=1}^M$. This ensures the existence

of a minimizer $\{\overline{\rho}_m\}_{m=1}^M$. Hence, (25) is a necessary and sufficient condition for the existence of a minimizer.

Notice that the cost function in (22) and the inequality constraints in (24) are convex. The equality constraint in (23) is affine linear. Hence, the optimization problem is convex and the Kuhn-Tucker conditions are necessary and sufficient conditions to find the minimizer $\{\overline{\rho}_m\}_{m=1}^M$. The Lagrangian of this optimization problem is

$$\mathcal{L} = \sum_{m=1}^{M} \mathcal{Q}\left(\sqrt{2\rho_m\gamma_m}\right) + \sum_{m=1}^{M} \nu_m \left(\frac{c_m}{\gamma_m} - \rho_m\right) + \mu\left(\sum_{m=1}^{M} \rho_m - \rho_{TOT}\right) \tag{39}$$

where $\{\nu_m\}$ and μ are the Lagrange multipliers. The optimizers $\{\overline{\rho}_m\}_{m=1}^M$, $\{\overline{\nu}_m\}_{m=1}^M$, and $\overline{\mu}$ must satisfy the following conditions:

- For $\{\overline{\rho}_m\}_{m=1}^M$, $\{\overline{\nu}_m\}_{m=1}^M$, and $\overline{\mu}$,
$$\frac{\partial \mathcal{L}}{\partial \rho_m} = 0 \ , \quad m = 1, \cdots, M \ . \tag{40}$$
- Constraints (23) and (24).
- $\overline{\nu}_m \geq 0, \quad m = 1, \cdots, M.$
- $\overline{\nu}_m \left(\frac{c_m}{\rho_m} - \overline{\rho}_m\right) = 0, \quad m = 1, \cdots, M.$

The condition in (40) is equivalent to

$$\sqrt{\frac{\gamma_m}{\overline{\rho}_m}}\, e^{-\overline{\rho}_m\gamma_m} = \overline{\mu} - \overline{\nu}_m$$

It will be demonstrated later in the uniqueness argument that the parameter $\overline{\mu}$ can be chosen to satisfy the power constraint (23).

For a given $\overline{\mu}$, the optimizers $\overline{\rho}_m$ and $\overline{\nu}_m$ that satisfy four conditions above can be chosen for each m as follows. Consider the equation

$$\sqrt{\frac{\gamma_m}{\rho_m}}\, e^{-\rho_m\gamma_m} = \overline{\mu} \ . \tag{41}$$

If there exists a $\rho_m \geq \frac{c_m}{\gamma_m}$ that solves (41), choose $\overline{\nu}_m = 0$ and $\overline{\rho}_m = \rho_m$. This happens when $\overline{\mu}$ is sufficiently small, which is the case when ρ_{TOT} is sufficiently large. Otherwise, choose $\overline{\rho}_m = \frac{c_m}{\gamma_m}$ and $\overline{\nu}_m > 0$ such that $\gamma_m e^{-c_m} = \sqrt{c_m}\,(\overline{\mu} - \overline{\nu}_m)$. Such $\overline{\nu}_m > 0$ exists because the left hand side of (41) is a decreasing function of ρ_m. Hence, the minimizer $\{\overline{\rho}_m\}$ is

$$\overline{\rho}_m = \max\left(\frac{c_m}{\gamma_m}, \rho_m\right)$$

where ρ_m is the solution of (41).

Now, we argue the uniqueness of the minimizer. Define a function $F(\rho)$ for a fixed $\gamma > 0$ as follows:

$$F(\rho) \stackrel{def}{=} \sqrt{\frac{\gamma}{\rho}} \exp\left(-\rho\gamma\right) \ . \tag{42}$$

Hence, (26) can be written as $F(\overline{\rho}_m) = \overline{\mu}$, $m = 1, \cdots, M$. Notice that $F(\rho)$ is a strictly decreasing function of ρ. Hence, given γ_m, $\overline{\rho}_m$ and $\overline{\mu}$ are one-to-one related and inversely proportional. Also, $F(0) = \infty$ and $F(\infty) = 0$. Since $F(\rho)$ is continuous and strictly decreasing, $F(\rho)$ takes on all values on $[0, \infty)$. Hence, there exists a solution to $F(\rho) = \mu$ for any $\mu \in [0, \infty)$. Note that solution is unique due to one-to-one relation between ρ and μ.

As demonstrated in the Claim 2 derivation below, the power constraint $\sum_{m=1}^{M} \rho_m = \rho_{TOT}$ is satisfied by tuning the parameter μ to obtain $\overline{\mu}$. Note that in the equations $F(\rho_m) = \mu$, $m = 1, \cdots, M$, increasing/decreasing μ results in simultaneous decrease/increase in all $\{\rho_m\}_{m=1}^{M}$ since μ is common to all m. Hence, it is easy to see that there is a one-to-one correspondence between $\overline{\mu}$ and ρ_{TOT}. Combining this fact and the uniqueness argument in the previous paragraph, it can be inferred that there is only one combination of $\{\rho_m\}_{m=1}^{M}$ with μ that satisfies (26) and (27). This establishes the uniqueness of the minimum effective BER solution.

Claim 2: $BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)}$

Given a channel realization with N_{dim} sub-channel SNR values in (20), let $BER_{eff}^{(M)}$ and $BER_{eff}^{(M+1)}$ represent the effective BER defined in (21) for systems with relative throughput of M and $M+1$, respectively. Let the corresponding sub-channel power allocation be $\left\{\rho_m^{(M)}\right\}_{m=1}^{M}$ and $\left\{\rho_m^{(M+1)}\right\}_{m=1}^{M+1}$, respectively, with the same total power constraint ρ_{TOT}.

First, we show that sufficient conditions for $BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)}$ to hold are:

$$1. \quad \rho_m^{(M)} \geq \rho_m^{(M+1)} \ , \ \ m = 1, \cdots, M \tag{43}$$

$$2. \quad \rho_m^{(M)}\gamma_m \geq \rho_{m+1}^{(M)}\gamma_{m+1} \ , \ \ m = 1, \cdots, M-1 \quad \text{for any } M. \tag{44}$$

Since $BER\left(\rho_m\gamma_m\right)$ is a strictly decreasing function of $\rho_m\gamma_m$, the following can be obtained from (43) and (44):

$$\begin{aligned} BER_{eff}^{(M)} &= \frac{1}{M}\sum_{m=1}^{M} BER\left(\rho_m^{(M)}\gamma_m\right) \\ &= \frac{1}{M+1}\left(\sum_{m=1}^{M} BER\left(\rho_m^{(M)}\gamma_m\right) + \frac{1}{M}\sum_{m=1}^{M} BER\left(\rho_m^{(M)}\gamma_m\right)\right) \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{M+1}\left(\sum_{m=1}^{M} BER\left(\rho_m^{(M+1)}\gamma_m\right) + \frac{1}{M}\sum_{m=1}^{M} BER\left(\rho_m^{(M+1)}\gamma_m\right)\right) \\
&\leq \frac{1}{M+1}\left(\sum_{m=1}^{M} BER\left(\rho_m^{(M+1)}\gamma_m\right) + \frac{1}{M}\times M \times BER\left(\rho_{M+1}^{(M+1)}\gamma_{M+1}\right)\right) \\
&= \frac{1}{M+1}\sum_{m=1}^{M+1} BER\left(\rho_m^{(M+1)}\gamma_m\right) = BER_{eff}^{(M+1)}
\end{aligned}
$$

Condition 1 states that for a given total power ρ_{TOT} and sub-channel gains $\{\gamma_m\}_{m=1}^{N_{dim}}$, adding one new sub-channel (sub-channel $M+1$) does not increase the amount of power allocated to any existing sub-channel, which is intuitively pleasing. Condition 2 requires that the total received SNR for a sub-channel is proportional to the corresponding sub-channel gain for a given M, ρ_{TOT}, and $\{\gamma_m\}_{m=1}^{M}$ assuming (20).

Next, we argue that the exact minimum effective BER solution satisfies Condition 1 and 2 above. To show that Condition 1 is satisfied, observe the optimality conditions for minimizing $BER_{eff}^{(M)}$ and $BER_{eff}^{(M+1)}$ without the inequality constraints:

$$
\left.\begin{array}{l} F(\rho_m^{(M)}) = \overline{\mu}^{(M)}\ ,\ \ m = 1,\cdots,M \\ \sum_{m=1}^{M}\rho_m^{(M)} = \rho_{TOT} \end{array}\right\} \tag{45}
$$

$$
\left.\begin{array}{l} F(\rho_m^{(M+1)}) = \overline{\mu}^{(M+1)}\ ,\ \ m = 1,\cdots,M+1 \\ \sum_{m=1}^{M+1}\rho_m^{(M+1)} = \rho_{TOT} \end{array}\right\} \tag{46}
$$

Assume that for a given N_{dim} sub-channel gains $\{\gamma_m\}_{m=1}^{N_{dim}}$, the solution of (45) has been found. By choosing $\overline{\mu}^{(M+1)} = \overline{\mu}^{(M)}$ and $\rho_m^{(M+1)} = \rho_m^{(M)}$, for $m = 1,\cdots,M$, the first optimality condition in (46) is satisfied, but

$$
\sum_{m=1}^{M+1}\rho_m^{(M+1)} = \rho_{TOT} + \rho_m^{(M+1)} \geq \rho_{TOT}\ .
$$

Consider choosing $\mu^{(M+1)} = \mu^{(M)} + \delta$, $\delta > 0$. Since $\rho_m^{(M+1)}$ and $\mu^{(M+1)}$ are one-to-one and inversely proportional, this results in $\rho_m^{(M+1)} < \rho_m^{(M)}$ for *all* $m = 1,\cdots,M$. By choosing an appropriate δ, the power constraint in (46) can be satisfied, and the solution of (46) is obtained. Applying the inequality constraints, $\overline{\rho}_m^{(M)} = \max\left(\frac{c_m}{\gamma_m}, \rho_m^{(M)}\right)$, hence we have $\overline{\rho}_m^{(M+1)} \leq \overline{\rho}_m^{(M)}$. This demonstrates that Condition 1 is satisfied.

To demonstrate that Condition 2 is satisfied, we show that for a given $\{\gamma_m\}_{m=1}^{M}$, $\{\rho_m\gamma_m\}_{m=1}^{M}$ is a decreasing sequence assuming $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_M$. This is trivially satisfied for m where $\rho_m < \frac{c_m}{\gamma_m}$ since $\overline{\rho}_m\gamma_m = c_m$ as long as c_m is the same for all m or a decreasing sequence.

When $\gamma_m > \frac{c_m}{\gamma_m}$, consider the following version of (41):

$$\frac{\exp\left(-\rho_m\gamma_m\right)}{\sqrt{\rho_m\gamma_m}} = \frac{\overline{\mu}}{\gamma_m}\ , \tag{47}$$

Note that for a given *fixed* $\{\gamma_m\}_{m=1}^{M}$ and ρ_{TOT}, $\overline{\mu}$ is also fixed. Let $x = \rho_m\gamma_m$ and $\gamma = \gamma_m$. Since $\rho_m > 0$, we have $x > 0$. Also, $\mu > 0$ as argued above. To show that Condition 2 holds, it suffices to show that x is an increasing function of γ for a fixed $\overline{\mu}$. Differentiating both sides of (47) after the change of variables, we obtain

$$\frac{dx}{d\gamma} = \frac{2\overline{\mu}e^x x^{3/2}}{\gamma^2(2x+1)} > 0\ .$$

Hence, for a fixed $\overline{\mu}$, x is an increasing function of γ. This demonstrates that Condition 2 is satisfied. Since Condition 1 and 2 in (43) and (44) are satisfied, we conclude that $BER_{eff}^{(M)} \leq BER_{eff}^{(M+1)}$.

References

[1] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, pp. 23–27, February 1998.

[2] *http://www.3gpp.org.*

[3] A. Naguib, N. Seshadri, and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Mag.*, pp. 76–92, May 2000.

[4] A. J. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1218–1230, October 1997.

[5] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, pp. 357–366, March 1998.

[6] E. N. Onggosanusi, B. D. V. Veen, and A. M. Sayeed, "Optimal antenna diversity signaling for wideband systems utilizing channel side information," *submitted to IEEE Trans. Commun.*, October 2000.

[7] Z. Wang and G. B. Giannakis, "Wireless multicarrier commmunications," *IEEE Signal Processing Mag.*, pp. 29–48, May 2000.

[8] T. Keller and L. Hanzo, "Adaptive multicarrier modulation: A convenient framewrok for time-frequency processing in wireless communications," *Proc. IEEE*, vol. 88, pp. 611–640, May 2000.

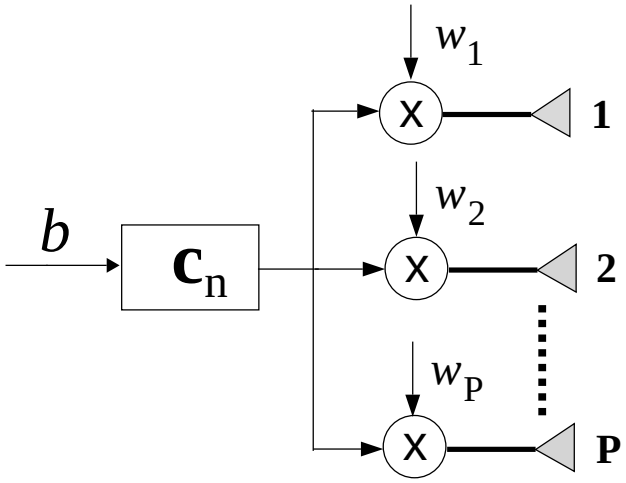
[9] J. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. Circ. System*, vol. CAS-25, pp. 772–781, Sept. 1978.

[10] G. H. Golub and C. F. V. Loan, *Matrix Computations.* John Hopkins, 1996.

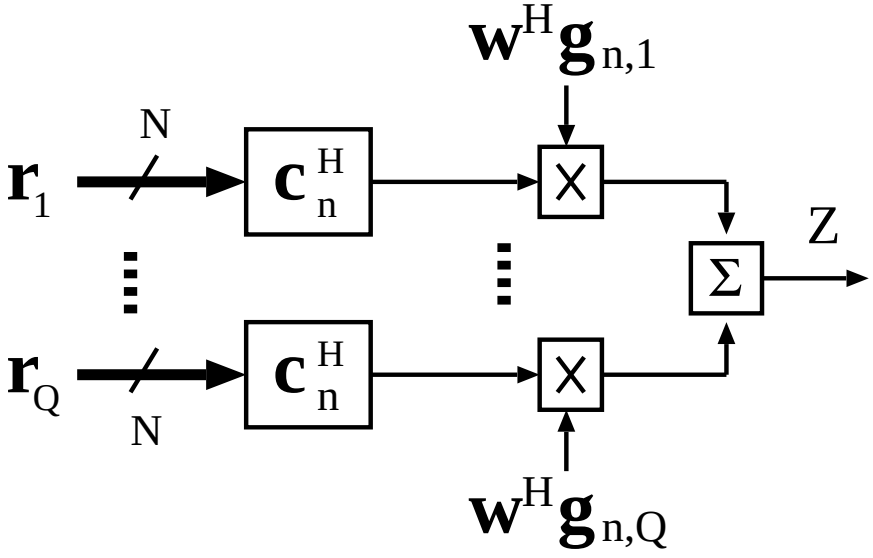
[11] J. G. Proakis, *Digital Communications.* New York: McGraw Hill, 3rd ed., 1995.

[12] S. Boyd, *EE364 Notes.* Stanford University, 1999.

List of Figures

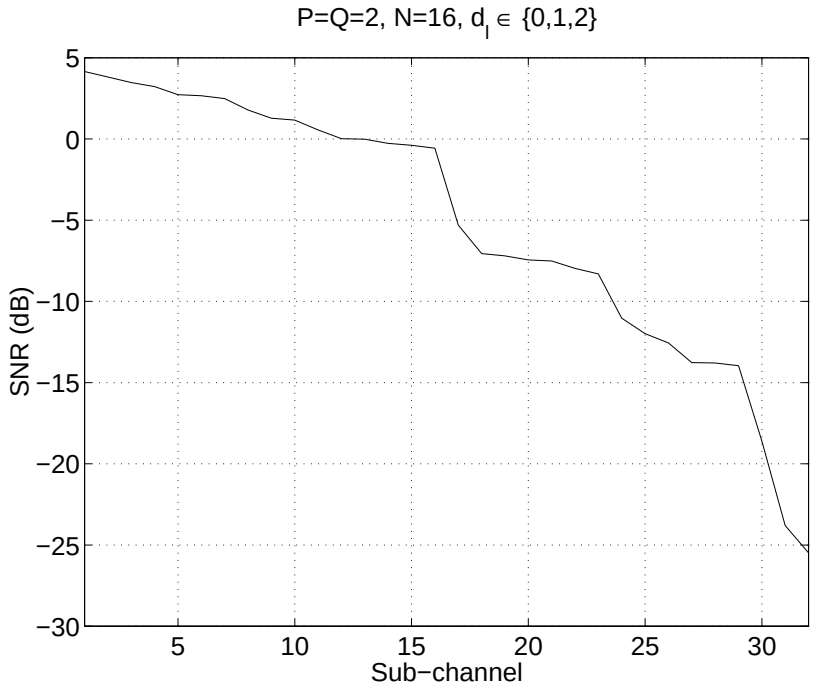


(a)

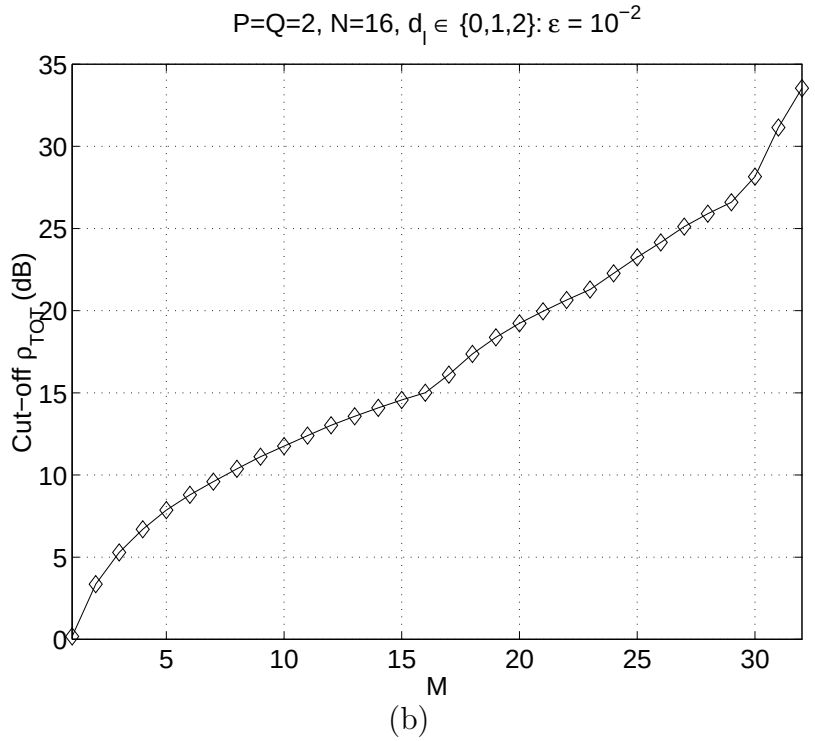


(b)

Figure 1: Single-user minimum BER system for BPSK (a) transmitter (b) receiver

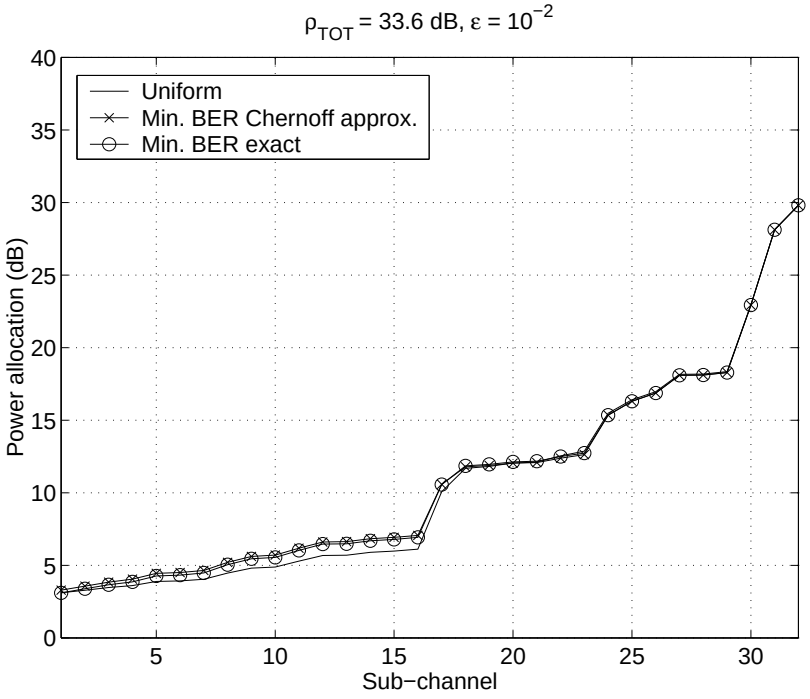


(a)

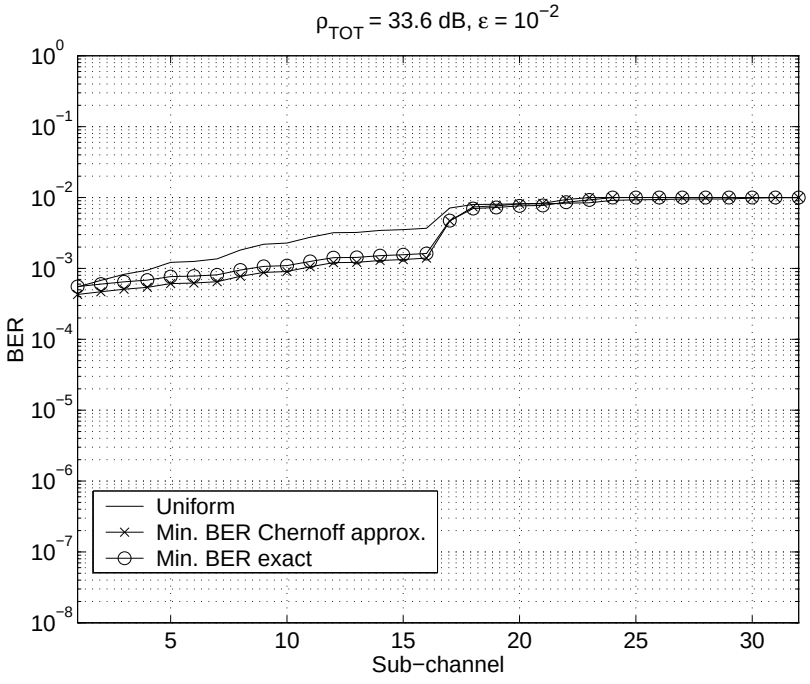


(b)

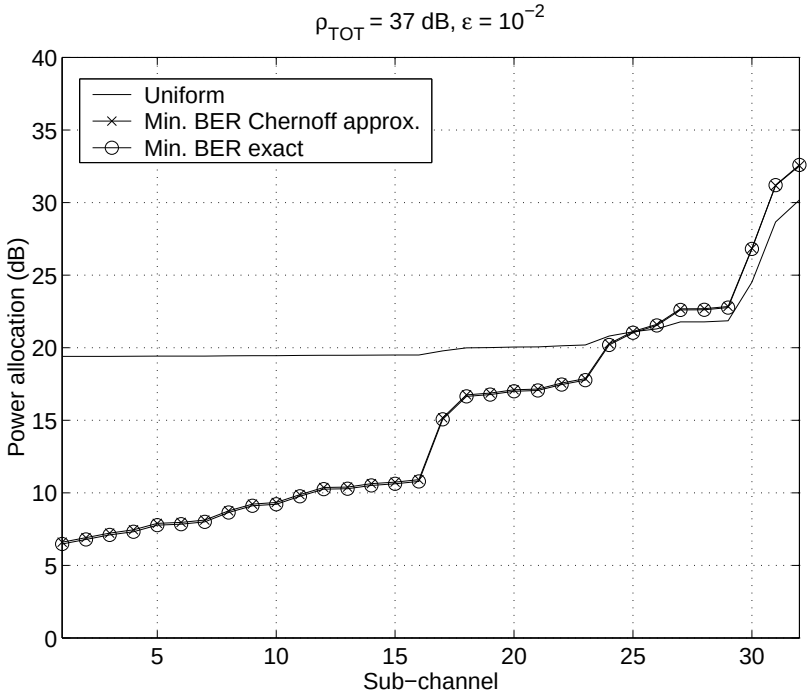
Figure 2: (a) A channel realization for $N = 16$, $P = Q = 2$, $L = 3$ ($d_l \in \{0, 1, 2\}$) (b) Cut-off power $\rho_{co,M}$ as a function of M to achieve feasibility condition (25).


ρ_{TOT} = 33.6 dB, $\varepsilon = 10^{-2}$
Uniform
Min. BER Chernoff approx.
Min. BER exact
Power allocation (dB)
Sub−channel
0
5
10
15
20
25
30
35
40

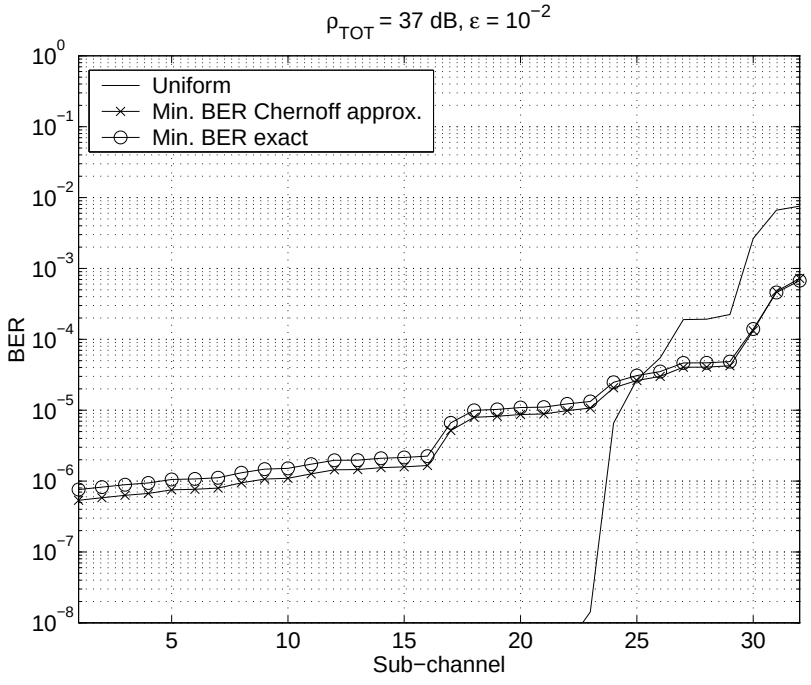
(a)


ρ_{TOT} = 33.6 dB, $\varepsilon = 10^{-2}$
Uniform
Min. BER Chernoff approx.
Min. BER exact
BER
Sub−channel
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}

(b)

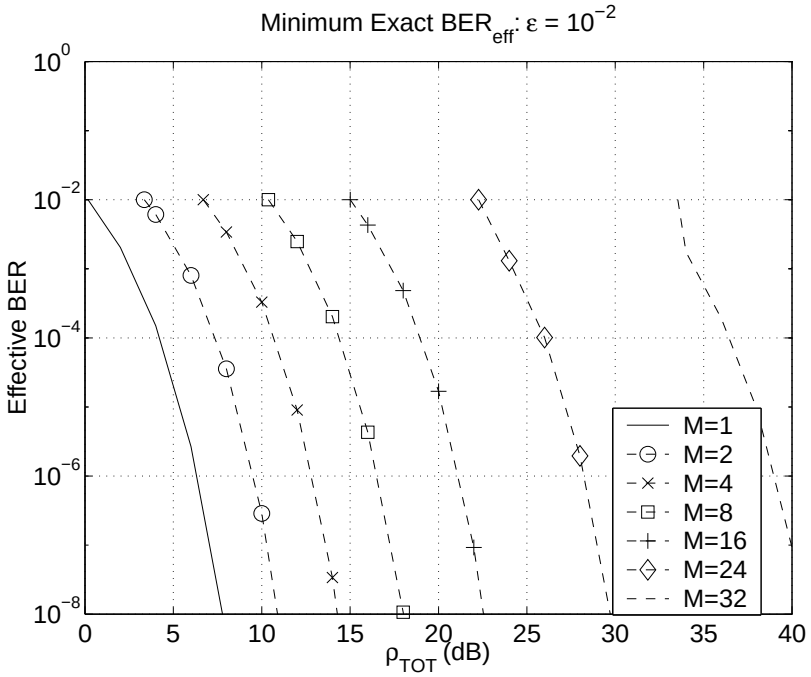


(c)

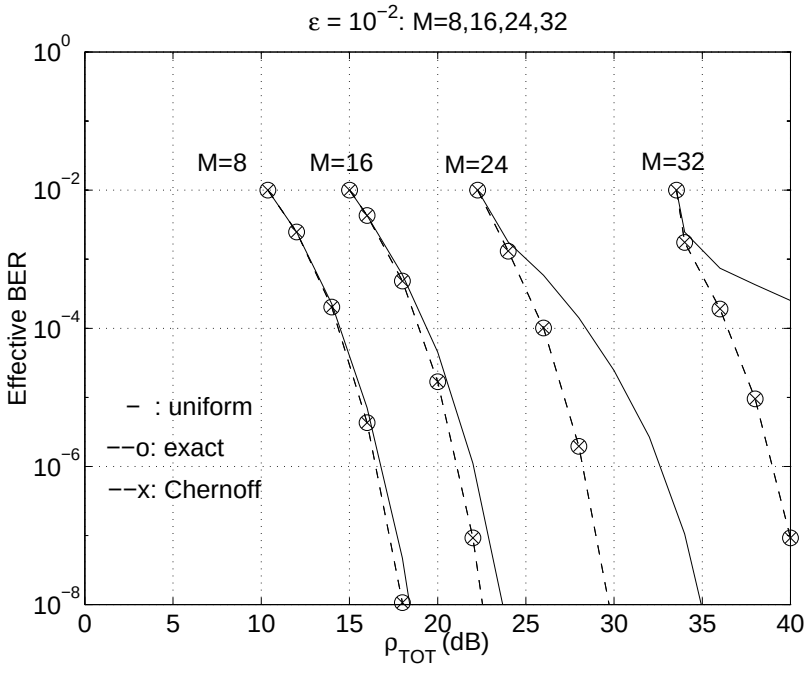


(d)

Figure 3: Different power allocation schemes and the resulting sub-channel BER for $\rho_{TOT} =$ 33.6 dB ((a)–(b)) and $\rho_{TOT} = 37$ dB ((c)–(d)).

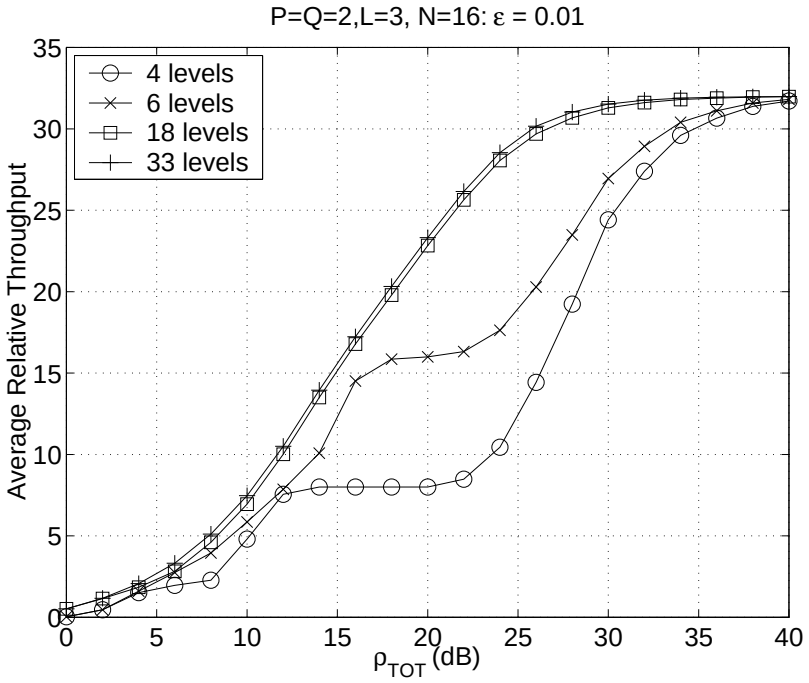


(a)

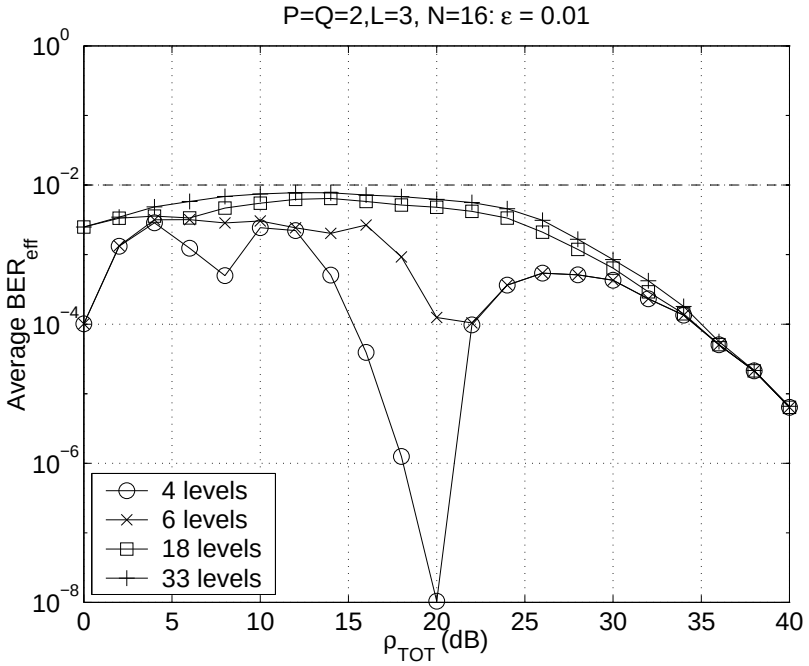


(b)

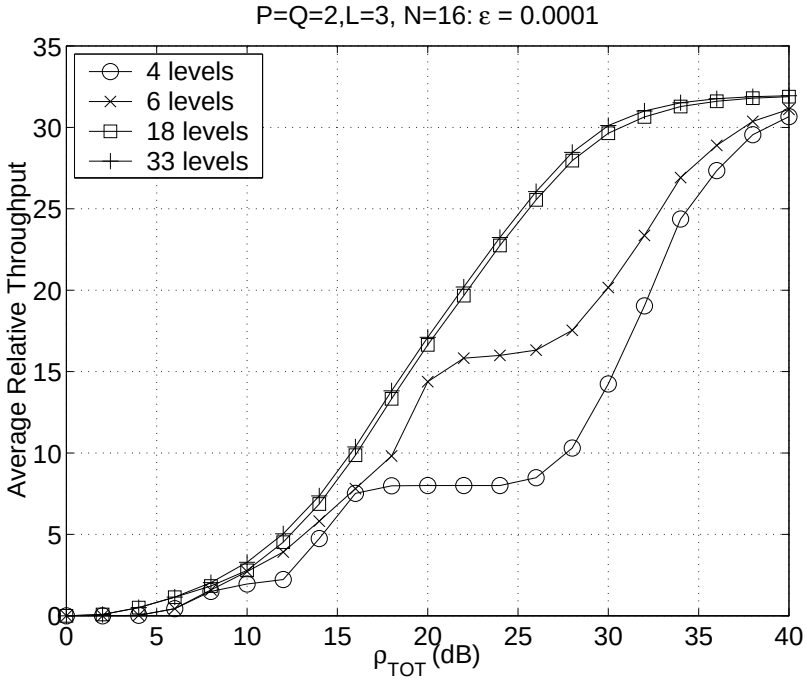
Figure 4: $BER_{eff}^{(M)}$ for different M. (a) Minimum $BER_{eff}^{(M)}$ (b) Comparison of power allocation schemes.


P=Q=2,L=3, N=16: ε = 0.01
4 levels
6 levels
18 levels
33 levels
Average Relative Throughput
ρTOT (dB)
0
5
10
15
20
25
30
35
40

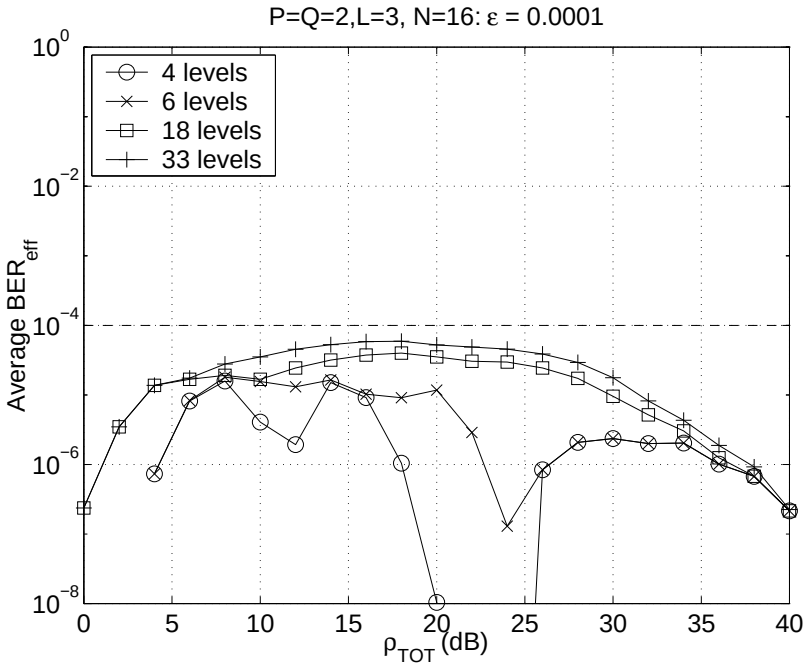
(a)


P=Q=2,L=3, N=16: ε = 0.01
Average BEReff
ρTOT (dB)
4 levels
6 levels
18 levels
33 levels

(b)



(c)



(d)

Figure 5: Adaptive throughput performance. We choose $N = 16$, $P = Q = 2$, $L = 3$ ($d_l \in \{0, 1, 2\}$). (a) Average M for $\varepsilon = 10^{-2}$ (b) Average BER_{eff} for $\varepsilon = 10^{-2}$ (c) Average M for $\varepsilon = 10^{-4}$ (d) Average BER_{eff} for $\varepsilon = 10^{-4}$