DRDO–IISc Programme on
Advanced Research in Mathematical
Engineering

Exponential Diversity Achieving Spatio-Temporal Power Allocation Scheme for Fading Channels

(TR-PME-2004-08)

by

K. Premkumar, Vinod Sharma, and Arun Rangarajan
Department of ECE, Indian Institute of Science, Bangalore, 560012, India
vinod@ece.iisc.ernet.in

April 20th, 2004

Indian Institute of Science
Bangalore 560 012

Exponential Diversity Achieving Spatio-Temporal Power Allocation Scheme for Fading Channels

K. Premkumar, Vinod Sharma, and Arun Rangarajan

April 20th, 2004

ABSTRACT

In this paper, we analyze optimum (in space and time) adaptive power transmission policies for fading channels when the channel state information (CSI) at the transmitter (CSIT) and receiver (CSIR) is available. The transmitter has a long-term (time) average power constraint. There can be multiple antennas at the transmitter and receiver. The channel experiences Rayleigh fading. We consider *beamforming* and *space-time coded* systems with perfect/imperfect CSIT and CSIR. We also consider an inner convolutional code (with beamforming or an outer space-time code). The performance measure is the bit error rate (BER). We show that in both coded and uncoded systems, our power allocation policy provides exponential diversity gain if perfect CSIT is available. We also show that, if the quality of CSIT degrades then the exponential diversity is retained only in the low SNR region but we get only polynomial diversity in the high SNR region.

Keywords: Power allocation, fading channels, CSIT, CSIR, convolutional codes, beamforming, space-time codes.

1 Introduction

Wireless channels are essential to provide ubiquitous connectivity to the users. However, due to multipath fading, low bandwidth and broadcast nature, providing Quality of Service (QoS) to users in such channels has been a challenge. Currently, a lot of effort is being invested into increasing the capacity and reducing the BER of wireless channels. A significant gain can be achieved by using multiple antennas at the transmitter and at the receiver (MIMO) (see [8], [9]). In this paper, we consider the BER of MIMO systems with Rayleigh fading (our methods can be used with other fading distribution also).

In an AWGN channel the BER decreases (for various error correcting codes and modulation schemes) exponentially with the signal-to-noise ratio (SNR). However, the average BER in a Rayleigh fading channel decreases only *linearly* with SNR. This indicates the severe degradation in performance caused by Rayleigh fading. This degradation can be partially mitigated by a fully interleaved convolution code of *minimum distance*, d_{min} (a fully interleaved convolution code with soft decision decoding can be viewed as one of providing time diversity with BER decreasing as $SNR^{-d_{min}}$). However, this is done at the cost of reduced information rate, increased receiver complexity and the increased decoding delay. Alternatively, the degradation in BER performance due to fading can be partially mitigated by using multiple antennas at the transmitter and the receiver. For example, if n_t antennas are used at the transmitter and n_r at the receiver, one can achieve BER decay at a rate $SNR^{-n_t n_r}$ (called a *diversity order* of $n_t n_r$) ([8], [9]). This diversity order can be achieved with Space-Time codes, even if the transmitter does not have any knowledge of the channel ([8], [9]). If the transmitter also has CSI, even if not exact, it can be exploited to obtain further reduction in BER ([8], [9] [7]). But, the diversity order remains $n_t n_r$. It is generally believed to be the maximum diversity order one can achieve for a Rayleigh fading channel (although if a convolutional code with minimum distance d_{min} and interleaving is used, one can get the diversity order of $d_{min} n_t n_r$).

In all the works cited above and in the references therein, the power allocation is among the different transmit antennas (space-only power allocation). This is apparently due to the fact that the power allocation in time does not yield much improvement in capacity for a Rayleigh fading channel ([4]). However, it is shown in [1] and [11] that power allocation in time can provide significant reduction in BER. The results in [1] and [11] are for single input single output (SISO) system. A close look at the results in [1] and [11] reveals that even for an SISO system with Rayleigh fading, the power allocation policies obtained in [1] and [11] provide *exponential diversity* order i.e., the BER decreases exponentially with SNR, as in an AWGN channel.

This paper takes a close look at the exponential diversity order aspect of the policies provided in [1] and [11]. We also study the corresponding policies for the MIMO systems. For perfect CSIT case, our power allocation policy provides an exponential diversity order, i.e., the BER, $P_b \leq \alpha e^{-f(n_t,n_r)}$, where α is a positive constant, and $f > 0$ is an increasing function of n_t and n_r. We have provided the lower bounds on the exponential diversity order we achieve.

We also study the more realistic scenario of partial CSIT at the transmitter. We obtain a very interesting result that as the quality of CSIT degrades, the exponential diversity is observed only in the low SNR regime but we get only polynomial diversity (as in previous studies) for high SNR case. However, our policies still provide significant improvements in BER as compared to previous studies ([8], [9], [7]).

In our earlier work [12], we studied the bounds on the *beamforming* approach in MISO systems with perfect CSIT (and perfect CSIR). The present work is an extension of [12]. We provide the BER performance bounds for the perfect and imperfect CSIT cases for SISO, MISO and MIMO systems here. Furthermore, the power allocation policy considered in [12] is an extension of [7], whereas here, we report that the power allocation policy of perfect CSIT case can be extended to the imperfect CSIT case (without any degradation in BER performance as compared to [12]).

The rest of the paper is organized as follows. In section 2, we explain the perfect CSIT system model and the optimization problem. We consider SISO, MISO, and MIMO systems. In section 3, we study the system with imperfect CSIT. In Section 4, we study the case of imperfect CSIR. We study the BER performance of Space-Time Coded system with our power allocation policy in Section 5. Section 6 concludes this study.

2 System With Perfect CSIT

We consider a single user narrowband (*flat fading*) communication system employing n_t transmit antennas and n_r receive antennas. We describe the channel between i^{th} receive antenna and j^{th} transmit antenna by a complex Gaussian random variable h_{ij}. Thus, the $n_r \times n_t$ matrix, $\mathbf{H} = [h_{ij}]$ represents the channel. We assume independent Rayleigh fading on each of the diversity branches. Also, we assume the fading to vary independently from one symbol to another i.e., $\{\mathbf{H}_k : k \in \mathcal{Z}^+\}$ is an i.i.d. process[1]. The additive noise, N, is temporally and spatially white with mean zero and follows a multivariate complex circular Gaussian distribution, $N \sim \mathcal{N}_C(0, \sigma^2 \mathrm{I}_{n_r})$, i.e., $\mathrm{E}[NN^T] = 0$ and $\mathrm{E}[NN^\dagger] = \sigma^2 \mathrm{I}_{n_r}$, where $\dagger$ denotes Hermitian. Coherent signaling is assumed. We assume perfect CSIR (the imperfect CSIR case is studied in Sec. 4).

In this section, we consider the perfect CSIT case, where the channel $\mathbf{H}$ is known to the transmitter. In this case, the optimum (in the SNR or BER sense) power allocation in space is achieved by *beamforming* [8]. We call this as *Space-Only Power Allocation* (SOPA). It is to be noted that in SOPA, the total power transmitted in a symbol duration is constant. On the other hand, we use the CSIT, $\mathbf{H}$ to compute the power $P(.)$ to be allocated in each symbol duration while satisfying the beamforming power allocation. We call this as *Space-Time Power Allocation* (STPA). Thus, in STPA, the total power transmitted in a symbol duration varies from one symbol to another, but the long term (time) average of transmit power is fixed.

We derive the optimal power allocation policy here. The output of the matched filter is sampled at symbol duration and the received complex signal vector at time k is given by

$$Y_k \quad = \quad \sqrt{P(\gamma)}\, \mathbf{H}_k\ W_k\ x_k + N_k, \quad k = 0, 1, 2, \ldots, \tag{1}$$

where x_k is the transmitted symbol, W_k is the input beamforming weight vector, $P(\gamma)$ is the transmit power, and $\gamma = \|\mathbf{H}_k W_k\|^2 \mathrm{E}|x|^2/\sigma^2$ is the SNR of the channel. We define $s = \mathrm{E}|x|^2/\sigma^2$ as the average SNR per branch. Because of the i.i.d. assumptions of $\mathbf{H}_k$ and N_k, the optimal power at time k does not need to depend upon previous decisions and hence we can drop the time index.

The maximum likelihood (ML) detection of x given Y corresponds to minimizing the following metric

$$\arg\min_x \left| x - \frac{(\mathbf{H}W)^\dagger Y}{\sqrt{P(\gamma)}\ ||\ \mathbf{H}W\ ||^2} \right|^2 \quad = \quad \arg\min_x\ |x - \hat{x}|^2 \tag{2}$$

where

$$\hat{x} \quad \triangleq \quad \frac{(\mathbf{H}W)^\dagger Y}{\sqrt{P(\gamma)} \|\mathbf{H}W\|^2} \quad \sim \quad \mathcal{N}_C\left(x^*, \frac{\sigma^2}{P(\gamma)\ ||\ \mathbf{H}W\ ||^2}\right).$$

Thus, the performance of the MIMO system (SISO and MISO systems are special cases) can be interpreted as the output of an AWGN channel with SNR $= P(\gamma)\ s\ \|\mathbf{H}W\|^2$. The transmit weight vector W is chosen to maximize the output SNR subject to the average transmit power constraint. The optimum

[1]Please refer to Appendix-4 for a note on the i.i.d. fading assumption.

W is the eigenvector of $\mathbf{H}^\dagger\mathbf{H}$ corresponding to its largest eigenvalue, λ ([8]) and $\|\mathbf{H}W\|^2 = \lambda$. For a given γ $(= s\lambda)$, the BER is given by

$$P_{b|\gamma} \;\; = \;\; c \cdot Q\left(\sqrt{g \cdot \; \gamma \; P(\gamma)}\right) \tag{3}$$

where $Q(.)$ denotes the Q-function, and c and g are constants. Typically, the value of g is related to the minimum distance in the constellation, and c is related to the number of constellation points that achieve this minimum distance. For $BPSK$, $c = 1$ and $g = 2$, and for $QPSK$ with Gray coding, $c = 1$ and $g = 2$ gives the exact BER. For $M - ary\ QAM$ an accurate (for high SNR) and useful approximation of the BER is obtained ([13]) by setting $c = 4\frac{\sqrt{M}-1}{\sqrt{M}\cdot\log_2 M}$ and $g = \frac{3}{M-1}\log_2 M$. For $M - PSK$ and high SNR, an approximate BER can be obtained with $c = 2/\log_2 M$ and $g = 2\sin^2(\pi/M)\log_2 M$ ([13]).

The optimization problem of minimizing the BER given in Eqn. 3 subject to the average transmit power constraint is

$$\min_{E_\gamma[P(\gamma)] \leq \overline{P}} \int_0^\infty c \cdot \mathrm{Q}\left(\sqrt{g \cdot \gamma P(\gamma)}\right) f_\gamma(\gamma) d\gamma \tag{4}$$

where f_γ is the pdf of γ.

In the following, we will take $\overline{P} = 1$. The solution of this problem will provide the optimal (in the BER sense) power allocation policy in space and time. Applying Lagrange's method to the above problem, we get the following family of unconstrained optimization problems parameterized by a multiplier $\mu > 0$,

$$\min_P \int_0^\infty \left(c \cdot \mathrm{Q}\left(\sqrt{g \cdot \gamma P(\gamma)}\right) + \mu\left[P(\gamma) - 1\right]\right) f_\gamma(\gamma)\ d\gamma.$$

The above unconstrained minimization corresponds to minimizing $c\cdot\mathrm{Q}\left(\sqrt{g \cdot \gamma P}\right) + \mu P$ for any given γ. The Q function is complicated, but can be replaced by a tight upper bound $\frac{1}{2}e^{-g\cdot\gamma P/2}$. Since $\frac{c}{2}e^{-g\cdot\gamma P/2} + \mu P$ is a convex function of P, the optimum P can be obtained by differentiating it w.r.t. P and equating to zero. This provides the optimum solution

$$P(\gamma) \;\; = \;\; \begin{cases} \frac{2/g}{\gamma}\ln\left(\frac{\gamma}{\gamma_0}\right) & \text{for } \gamma \geq \gamma_0 \\ 0 & \text{for } \gamma < \gamma_0. \end{cases} \tag{5}$$

where γ_0 is found by solving,

$$\int_{\gamma_0}^\infty P(\gamma) f_\gamma(\gamma) d\gamma \;\; = \;\; 1. \tag{6}$$

It should be noted here that γ_0 depends on the average SNR per branch, s and the modulation scheme. Also, the transmitter should have the complete knowledge of the fading statistics to compute γ_0 and the instantaneous fade values to compute $P(.)$.

The average BER is given by,

$$P_b \;\; = \;\; \frac{1}{2}F_\gamma\left(\gamma_0\right) + \int_{\gamma_0}^\infty c \cdot \mathrm{Q}\left(\sqrt{g \cdot \gamma P\left(\gamma\right)}\right) f_\gamma(\gamma)\ d\gamma \tag{7}$$

where F_γ is the c.d.f. of f_γ.

The SNR $\gamma = s\lambda$ (where $\lambda = \|\mathbf{H}W\|^2$, the largest eigenvalue of $\mathbf{H}^\dagger\mathbf{H}$) and its distribution varies for the SISO, MISO and MIMO cases. It is possible to compute Eqn. 7 exactly in each of these cases (either by numerical computation or via computing the integral by a Monte-Carlo method). However, in the following sections, we will show that the BER decays exponentially by providing explicit upper bounds. We study the cutoff γ_0/s and the bounds on the BER for SISO, MISO and MIMO cases separately. We consider $BPSK$ modulation scheme throughout the paper (although, as mentioned above, by choosing c and g appropriately, we can consider other modulation schemes).

2.1 SISO

We consider a system with single transmit antenna and single receive antenna (i.e., $n_t = n_r = 1$). The channel, $\mathbf{H}$ is a complex Gaussian random variable with mean 0 and variance 1. Here, the beamforming element $W = \mathbf{H}^*/|\mathbf{H}|$. Thus, the SNR of the channel is given by $\gamma = s|\mathbf{H}|^2 \sim \exp(1/s)$. The cutoff γ_0 is computed from

$$1 = \int_{\gamma_0}^{\infty} \frac{1}{\gamma}\ln\left(\frac{\gamma}{\gamma_0}\right)\frac{1}{s}e^{-\gamma/s}\, d\gamma$$

Taking $x = \gamma/\gamma_0$, we obtain

$$s = \int_{x=1}^{\infty} \frac{\ln(x)\, e^{-(\gamma_0/s)x}}{x}\, dx$$

From [5], p. 573, Eqn. 4.362 (2)[2],

$$\int_1^{\infty} \frac{e^{-\mu x}\ln(2x-1)}{x}\, dx = \frac{1}{2}[\mathrm{Ei}(-\mu/2)]^2$$

where Re $\mu > 0$. Since,

$$\begin{aligned}
\int_{x=1}^{\infty} \frac{e^{-\frac{\gamma_0}{s}x}\ln(x)}{x}\, dx &< \int_{x=1}^{\infty} \frac{e^{-\frac{\gamma_0}{s}x}\ln(2x-1)}{x}\, dx, \\
s &< 0.5\left[\mathrm{Ei}\left(-\frac{\gamma_0/s}{2}\right)\right]^2 \\
&= 0.5\left[-\mathrm{E}_1\left(\frac{\gamma_0/s}{2}\right)\right]^2
\end{aligned}$$

Using Theorem 2 of [2],

$$s \leq 0.5\left[\ln\left(1 - e^{-\frac{\gamma_0/s}{2}}\right)\right]^2$$

[2]The exponential integral $\mathrm{Ei}(-x) = -\int_x^\infty \frac{e^{-t}}{t}\, dt$ and the function $\mathrm{E}_1(x) = \int_x^\infty \frac{e^{-t}}{t}\, dt$, where $x > 0$.

Thus, using $-\ln(1-x) \leq kx$ (see Appendix-1)

$$\frac{\gamma_0}{s} \leq 2ke^{-\sqrt{2s}}. \tag{8}$$

Therefore, from (7), using the upper bound on Q-function,

$$\begin{aligned}
P_b &\leq 0.5\left[1 - e^{-\gamma_0/s} + \int_{\gamma_0/s}^{\infty} e^{-\ln(\frac{s\beta}{\gamma_0})} \cdot e^{-\beta}\, d\beta\right] \\
&= 0.5\left[1 - e^{-\gamma_0/s} + \frac{\gamma_0}{s}\int_{\gamma_0/s}^{\infty} \frac{e^{-\beta}}{\beta}\, d\beta\right] \\
&\leq 0.5\left[1 - e^{-\gamma_0/s} - \frac{\gamma_0}{s}\ln\left(1 - e^{-\gamma_0/s}\right)\right] \qquad \text{(Using Theorem 2 of [2])} \\
&\leq 0.5\left[\frac{\gamma_0}{s} + \frac{\gamma_0}{s}ke^{-\gamma_0/s}\right] \\
&\leq 0.5\frac{\gamma_0}{s}\left[1 + k\right].
\end{aligned}$$

Therefore, from Eqn. 8

$$P_b \leq ke^{-\sqrt{2s}}\left[1 + k\right]. \tag{9}$$

This shows exponential diversity. From Fig. 1, we see for large s, the bound is quite tight and hence we obtain correct diversity order.

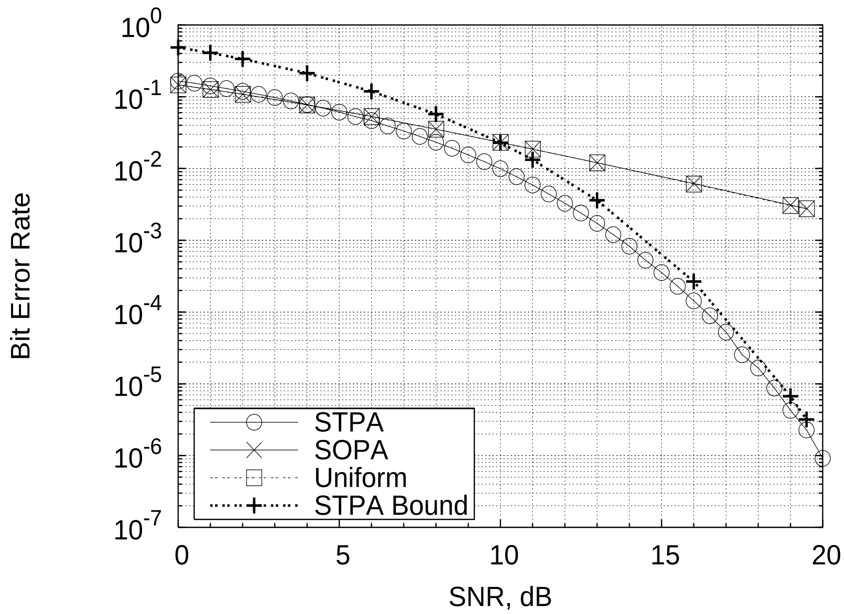

Figure 1: BER versus average SNR for a SISO system with perfect CSIT.

Let us compare the performance of the above STPA scheme with SOPA (which for SISO is same as no power control). From [10], p. 818, for SOPA we obtain, $P_b = \frac{1}{2}\left(1 - \sqrt{\frac{s}{1+s}}\right)$ which tends to $\frac{1}{2s}$ as $s \to \infty$.

In Fig. 1, we plot the upper bound given in Eqn. 9 along with P_b for STPA and SOPA. One observes that atleast for high SNR, the upper bound is quite tight. It is clear from Fig. 1 that the fall in BER with SNR is exponential in the case of STPA and linear in the case of SOPA. Also, the SOPA case provides the same performance as uniform power allocation. It is also evident from Fig. 1 that as SNR, s increases STPA provides significant performance as compared to SOPA. For example, to achieve a BER of 10^{-2}, STPA requires an SNR of 10 dB whereas SOPA requires an SNR of 16 dB, thus providing a power saving of 6 dB. For a stringent BER requirement, this power saving by STPA is substantial over SOPA.

2.2 MISO

We consider a system with $n_t > 1$ transmit antennas and $n_r = 1$ receive antenna. The channel in this case is a $1 \times n_t$ vector, $\mathbf{H} \sim \mathcal{N}_C(0, I_{n_t})$. Here, the beamforming vector is $W = \mathbf{H}^\dagger / \|\mathbf{H}\|$, and the SNR of the channel is given by $\gamma = s\|\mathbf{H}W\|^2 = s\|\mathbf{H}\|^2 \sim \text{Erlang}_{n_t}(1/s)$. The cutoff γ_0 is given by

$$\begin{aligned} 1 &= \int_{\gamma_0}^{\infty} \frac{1}{\gamma} \ln\left(\frac{\gamma}{\gamma_0}\right) \frac{1}{s} \frac{e^{-\gamma/s}\left(\frac{\gamma}{s}\right)^{n_t-1}}{(n_t-1)!}\, d\gamma \\ &= \frac{\left(\frac{\gamma_0}{s}\right)^{n_t}}{(n_t-1)!} \frac{1}{\gamma_0} \int_1^{\infty} e^{-(\gamma_0/s)x} x^{n_t-2} \ln(x)\, dx \end{aligned}$$

This equals (see Appendix-2)

$$\begin{aligned} &= \frac{1}{s(n_t-1)} \left[\int_{\gamma_0/s}^{\infty} \frac{e^{-t}}{t} dt + \sum_{k=0}^{n_t-3} \frac{e^{-\gamma_0/s}}{k+1} \sum_{i=0}^{k} \frac{(\gamma_0/s)^i}{i!} \right] \\ &\le \frac{1}{s(n_t-1)} \left[\int_{\gamma_0/s}^{\infty} \frac{e^{-t}}{t} dt + \sum_{k=0}^{n_t-3} \frac{e^{-\gamma_0/s}}{k+1} e^{\gamma_0/s} \right] \\ &= \frac{1}{s(n_t-1)} \left[\int_{\gamma_0/s}^{\infty} \frac{e^{-t}}{t} dt + \delta \right] \end{aligned}$$

where $\delta = \sum_{k=0}^{n_t-3} \frac{1}{k+1}$(for $n_t = 2$, $\delta = 0$). Therefore,

$$(n_t-1)s - \delta \le \int_{\gamma_0/s}^{\infty} \frac{e^{-t}}{t} dt \tag{10}$$

Thus, using Theorem 2 of [2],

$$(n_t-1)s - \delta \le -\ln(1 - e^{-\gamma_0/s})$$

and hence

$$e^{-(n_t-1)s+\delta} \ge 1 - e^{-\gamma_0/s}$$

Using Appendix-1, we get

$$\frac{\gamma_0}{s} \le k e^{-(n_t-1)s+\delta}. \tag{11}$$

Therefore, using Appendix-2

$$
\begin{aligned}
P_b &= \frac{1}{2}F_{\|\mathbf{H}\|^2}(\gamma_0/s) + \int_{\gamma_0/s}^{\infty} Q\left(\sqrt{2s\alpha P\left(s\alpha\right)}\right) f_{\|\mathbf{H}\|^2}(\alpha)d\alpha \qquad (12) \\
&\leq \frac{1}{2}F_{\|\mathbf{H}\|^2}(\gamma_0/s) + \int_{\gamma_0/s}^{\infty} \frac{1}{2}e^{-s\alpha P(s\alpha)} f_{\|\mathbf{H}\|^2}(\alpha)d\alpha \\
&= 0.5\left[1 - \sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!} + \int_{\alpha=\gamma_0/s}^{\infty} e^{-s\alpha\frac{1}{s\alpha}\ln\left(\frac{\alpha}{\gamma_0/s}\right)}\frac{e^{-\alpha}\alpha^{n_t-1}}{(n_t-1)!}d\alpha\right] \\
&= 0.5\left[1 - \sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!} + \left(\frac{\gamma_0}{s}\right)^{n_t}\int_{\alpha=1}^{\infty}\frac{e^{-(\gamma_0/s)x}x^{n_t-2}}{(n_t-1)!}dx\right] \\
&= 0.5\left[1 - \sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!} + \frac{\frac{\gamma_0}{s}e^{-\gamma_0/s}}{n_t-1}\sum_{i=0}^{n_t-2}\frac{(\gamma_0/s)^i}{i!}\right] \\
&\leq 0.5\left[1 - e^{-\gamma_0/s} + \frac{\gamma_0/s}{n_t-1}\right] \qquad (13)
\end{aligned}
$$

and hence,

$$
P_b \leq 0.5\frac{n_t}{n_t-1}ke^{-(n_t-1)s+\delta}. \qquad (14)
$$

Thus, we see the exponential decay of P_s with SNR. As n_t increases, the rate of BER decay increases.

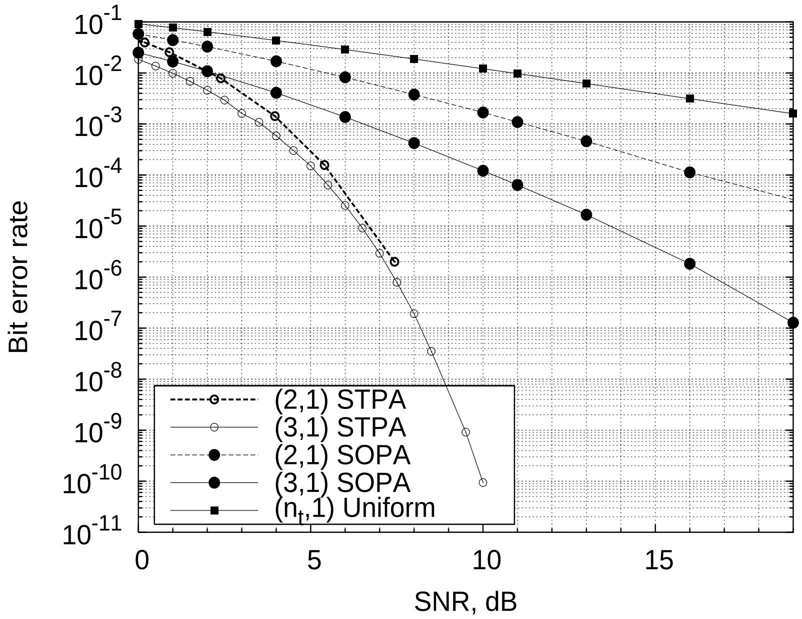

Figure 2: BER versus average SNR for MISO systems with perfect CSIT.

We compare the BER performance of STPA scheme with the SOPA scheme.

For SOPA scheme, the exact BER is (see [8] pp. 42–43)

$$\begin{aligned}
P_b &= \frac{1}{2^{n_t}}\left(1-\sqrt{\frac{s}{1+s}}\right)^{n_t}\sum_{k=0}^{n_t-1}\left[\frac{1}{2^k}\binom{n-1+k}{k}\cdot\left(1+\sqrt{\frac{s}{1+s}}\right)^k\right] \\
&= \frac{1}{2^{n_t}s^{n_t}} + \mathrm{O}\left(\frac{1}{s^{2n_t}}\right) \\
&\to \frac{1}{2^{n_t}s^{n_t}} \text{ as } s \to \infty\cdot
\end{aligned}$$

We plot P_b from Eqn. 12 in Fig. 2 for $n_t = 2$ and 3. For comparison, we have also included the curve for the optimal space-only beamforming power control of [8]. One can see significant gains in BER for our policy as compared to the space-only policy. The bound (Eqn. 13) is too loose and hence not provided.

It is worth mentioning here that the BER performance due to *Uniform Power Allocation* in MISO case is the same as that of SISO case evaluated by replacing s with $n_t \cdot s$, i.e.,

$$\begin{aligned}
P_s &= \frac{1}{2}\left(1-\sqrt{\frac{n_t s}{1+n_t s}}\right) \\
&\to \frac{1}{2n_t s} \text{ as } s \to \infty\cdot
\end{aligned}$$

2.3 MIMO

For this system, the maximum SNR for an average transmit power constraint is achieved by beamforming in the direction of the eigenvector corresponding to the largest eigenvalue, λ of the Wishart distributed matrix, $\mathbf{H}^\dagger\mathbf{H}$. Thus, the SNR of the channel γ is $s\lambda = s\|\mathbf{H}W\|^2$. It is to be noted that the nonzero eigenvalues (and the distribution) of $\mathbf{H}^\dagger\mathbf{H}$ and $\mathbf{H}\mathbf{H}^\dagger$ are same. Hence, without loss of generality, we consider the case $n_t \geq n_r$ (i.e., the BER of the system with n_t transmit antennas and n_r receive antennas will be the same as that of the system with n_r transmit antennas and n_t receive antennas).

The density function of the largest eigenvalue of the Wishart distributed matrix is given by ([3])

$$f_\lambda(\lambda) = \sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}\lambda^m e^{-i\lambda}}{m!}, \tag{15}$$

where $d_{i,m}$'s are constants which depend upon n_t and n_r. Therefore, the cutoff γ_0 is found by solving the equation,

$$\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}(\gamma_0/s)^m}{m!}\int_1^\infty x^{m-1}e^{-i(\gamma_0/s)x}\ln(x)\,dx = s.$$

Solving this equation (for the case $n_t > n_r$), we get (see Appendix-3)

$$\frac{\gamma_0}{s} \lesssim k\exp(-(s-c)/K) \tag{16}$$

where$k$ was defined earlier, c and K are defined in Appendix 3. The average BER satisfies

$$\begin{aligned} P_b &\leq \frac{1}{2}\int_0^{\gamma_0/s} f_\lambda(\lambda)d\lambda + \int_{\gamma_0/s}^{\infty} \frac{1}{2}e^{-s\lambda P(s\lambda)} f_\lambda(\lambda)d\lambda \\ &\lesssim 0.5Kke^{-(s-c)/K}. \end{aligned} \tag{17}$$

The parameter $1/K$ for various values of (n_t,n_r) is numerically computed and shown in Table 1.

(n_t,n_r)	$1/K$	K
(2,2)	2.5887	0.3863
(3,2)	4.0000	0.2500
(4,2)	5.3333	0.1875
(6,2)	7.8769	0.1270
(8,2)	10.3261	0.0968
(4,3)	7.2605	0.1377
(5,3)	8.7465	0.1143
(8,4)	15.2426	0.0656

Table 1: $1/K$ for various (n_t,n_r).

The approximation given by Eqn. 17 is very tight and is plotted in Fig. 3 for (n_t, n_r) = (4,2). It also shows the exponential decay of P_s with SNR and (using Table 1) that the rate of decay increases with (n_t, n_r).

The exact BER expression for SOPA is given in [3] (Eqn. 29). The uniform power allocation case here is also equivalent to a SISO system evaluated by replacing s with $n_t n_r \cdot s$, i.e.,

$$\begin{aligned} P_b &= \frac{1}{2}\left(1 - \sqrt{\frac{n_t n_r s}{1 + n_t n_r s}}\right) \\ &\to \frac{1}{2n_t n_r s} \text{ as } s \to \infty \end{aligned}$$

Fig. 3 shows the BER performance of a MIMO system with $n_t = 4$ and $n_r = 2$. We plot the cases of STPA, SOPA and the Uniform power allocation scheme. It is clear from Fig. 3 that the fall in BER with SNR is exponential for STPA case whereas for SOPA and uniform power allocation cases it is linear. Also, the power saving due to STPA over the other schemes is substantial. For example, STPA achieves a BER of 10^{-10} at 1 dB whereas SOPA requires 6 dB to achieve the same BER. It is also evident from Fig. 3 that this gain margin increases with BER.

3 System With Imperfect CSIT

The system model considered here is the same as in Sec. 2 except that the channel state information available at the transmitter is noisy. We assume that $\mathbf{H}_T$ is the transmitter's estimate of the channel. $\mathbf{H}_T$ could be a delayed version or a finite precision representation or a noisy version of $\mathbf{H}$. We assume that $\mathbf{H}$

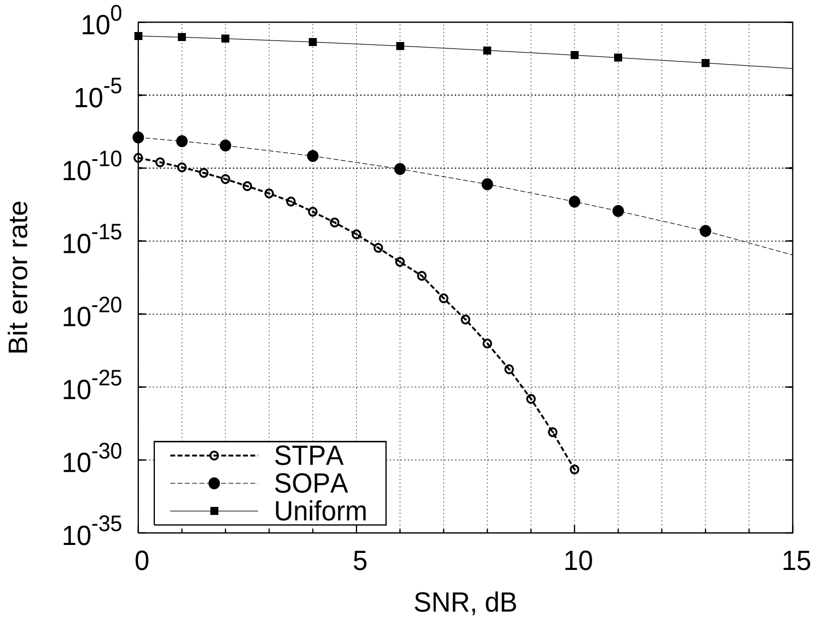

Figure 3: BER versus average SNR for (4,2) MIMO systems with perfect CSIT.

and $\mathbf{H}_T$ are jointly complex Gaussian. Let ρ_T be the correlation between $\mathbf{H}$ and $\mathbf{H}_T$.

Since, the transmitter assumes that the channel it sees is $\mathbf{H}_T$, the optimum transmit power will be the same as given in Eqn. 5. We do beamforming and power allocation as given in Eqn. 5, using $\mathbf{H}_T$ instead of $\mathbf{H}$, i.e., compute $P(\gamma_T)$, where $\gamma_T = s\lambda_T$ and λ_T is the largest eigenvalue of $\mathbf{H}_T^\dagger \mathbf{H}_T$. Coherent signaling is assumed. The output of the matched filter is sampled at symbol duration and the received complex signal vector is given by

$$Y_k \quad = \quad \sqrt{P(\gamma_T)}\; \mathbf{H}_k \; W_{Tk} \; x_k + N_k, \quad k = 0, 1, 2, \ldots,$$

where x_k is the transmitted symbol, and γ_T is the SNR estimate at the transmitter ($\gamma_T = s\lambda_T = s\|\mathbf{H}_T W_T\|^2$), $P(\gamma_T)$ is the instantaneous transmit power, W_{Tk} is the input weight vector (eigenvector of $\mathbf{H}_T^\dagger \mathbf{H}_T$ corresponding to the largest eigenvalue λ_T). Also, we define μ to be $\|\mathbf{H} W_T\|^2$. As before, we drop the time index k.

The model we follow to describe the imperfect CSIT is

$$\mathbf{H}_T \quad = \quad \rho_T \; \mathbf{H} + \sqrt{1 - \rho_T^2}\mathbf{E} \tag{18}$$

where $\mathbf{E}$ is the estimation error independent of $\mathbf{H}$ and has distribution $\sim \mathcal{N}(0, I_{n_t \cdot n_r})$. Equivalently,

$$\mathbf{H} \quad = \quad \rho_T \; \mathbf{H}_T + \sqrt{1 - \rho_T^2}\mathbf{E}' \tag{19}$$

where $\mathbf{E}'$ this time is the estimation error independent of $\mathbf{H_T}$ and has distribution $\sim \mathcal{N}(0, I_{n_t \cdot n_r})$($\mathbf{E}$ and $\mathbf{E}'$ are used interchangebly in the rest of the paper).

We assume that the receiver has perfect CSI and also the transmitter's knowledge of the channel $\mathbf{H}_T$. When $\mathbf{H_T}$ is delayed version of $\mathbf{H}$, then (18) will hold if $\{\mathbf{H_k}\}$ is an autoregressive process. If $\mathbf{H_T}$ is quantized $\mathbf{H}$, then $\mathbf{E}$ can be the quantization noise. This assumption is also made in [7]. The case when the receiver does not know $\mathbf{H_T}$ will be covered in the next section.

The BER is given by

$$P_{b|\mu,\lambda_T} = Q\left(\sqrt{2\ s\ \mu\ P(s\lambda_T)}\right).$$

The average BER is given by,

$$\begin{aligned} P_b &= \int_{\lambda_T=0}^{\gamma_0/s} 0.5 \cdot f_{\lambda_T}(\lambda_T)\ d\lambda_T + \\ &\int_{\mu=0}^{\infty}\int_{\lambda_T=\gamma_0/s}^{\infty} \mathrm{Q}\left(\sqrt{2s\mu P\left(s\lambda_T\right)}\right) \cdot f_{\mu,\lambda_T}(\mu,\lambda_T)\ d\mu\ d\lambda_T \end{aligned} \tag{20}$$

It is to be noted that the density functions, $f_{\lambda_T}(\lambda_T)$ are the same as that of the perfect CSIT case. As with perfect CSIT, Eqn. 20 can be actually computed. But, we obtain bounds on it which provide a better insight on the behaviour of P_b as the SNR changes. Since λ_T has the same distribution as λ corresponding to the perfect CSIT case, the bounds on the cutoff values, γ_0 computed for the perfect CSIT case hold good for the imperfect CSIT case. But, for the imperfect CSIT case, we also need to derive the conditional pdf's $f_{\mu|\lambda_T}(\mu|\lambda_T)$ for the SISO, MISO, and MIMO cases which we obtain in the following.

3.1 SISO

For $n_t = n_r = 1$, the (scalar) channel $\mathbf{H} \sim \mathcal{N}_C(0,1)$. Since $\mathbf{H} = \rho_T\mathbf{H_T} + \sqrt{1-\rho_T^2}\mathbf{E}$, $|\mathbf{H}|^2 = X^2 + Y^2$, where X and Y are the real and imaginary parts of $\mathbf{H}$. Clearly, X and Y are Gaussian distributed (given $\mathbf{H_T}$), with means , $m_x = \rho_T\ \mathrm{Re}(\mathbf{H_T})$ and $m_y = \rho_T\ \mathrm{Im}(\mathbf{H_T})$. Therefore , $m_x^2 + m_y^2 = \rho_T^2|\mathbf{H_T}|^2$. The variance of X and Y is $\frac{1}{2}\sigma_e^2 = \frac{1}{2}(1-\rho_T^2)$. Thus, the density function $f_{|\mathbf{H}|^2||\mathbf{H}_T|^2}(\alpha\big|\beta)$ follows a *noncentral chi-square* distribution

$$\begin{aligned} f_{|\mathbf{H}|^2\big||\mathbf{H}_T|^2}(\alpha\big|\beta) &= \frac{1}{\sigma_e^2}e^{-\left(\beta\rho_T^2+\alpha\right)/\sigma_e^2} I_0\left(\frac{2\sqrt{\alpha\beta}\rho_T}{\sigma_e^2}\right) \\ &= \frac{1}{\sigma_e^2}e^{-\left(\beta\rho_T^2+\alpha\right)/\sigma_e^2} \sum_{k=0}^{\infty}\frac{\left(\frac{-\alpha\beta\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)} \end{aligned}$$

where Γ is the Gamma function. The average BER is therefore given by

$$\begin{aligned} P_b &= 0.5 \cdot F_{|\mathbf{H}_T|^2}(\gamma_0/s) + \\ &\int_{\alpha=0}^{\infty}\int_{\beta=\gamma_0/s}^{\infty} \mathrm{Q}\left(\sqrt{2s\alpha P\left(s\beta\right)}\right) \cdot f_{|\mathbf{H}_T|^2}(\beta) f_{|\mathbf{H}|^2\big||\mathbf{H}_T|^2}(\alpha\big|\beta)\ d\alpha\ d\beta \qquad (21) \\ &\leq 0.5 \cdot F_{|\mathbf{H}_T|^2}(\gamma_0/s) + \\ &\int_{\alpha=0}^{\infty}\int_{\beta=\gamma_0/s}^{\infty} 0.5 \cdot e^{-s\alpha P(s\beta)} \cdot f_{|\mathbf{H}_T|^2}(\beta) f_{|\mathbf{H}|^2\big||\mathbf{H}_T|^2}(\alpha\big|\beta)\ d\alpha\ d\beta \end{aligned}$$

$$
\begin{aligned}
&= 0.5\left[1-e^{-\gamma_0/s}\right]+ \\
&\quad 0.5\int_{\alpha=0}^{\infty}\int_{\beta=\gamma_0/s}^{\infty} e^{-\frac{\alpha}{\beta}\ln\left(\frac{s\beta}{\gamma_0}\right)}\cdot e^{-\beta}\frac{1}{\sigma_e^2}e^{-\left(\beta\rho_T^2+\alpha\right)/\sigma_e^2}\sum_{k=0}^{\infty}\frac{\left(\frac{-\alpha\beta\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)}\,d\alpha\,d\beta \\
&= 0.5\left[1-e^{-\gamma_0/s}\right]+0.5\left(\frac{\gamma_0}{s}\right)^2\int_{x=1}^{\infty}\frac{e^{-(1+\frac{\rho_T^2}{\sigma_e^2})\frac{\gamma_0}{s}x}}{\frac{\gamma_0}{s}+\sigma_e^2\frac{\ln(x)}{x}}\exp\left(\frac{-\left(\frac{\gamma_0}{s}\right)^2\frac{\rho_T^2}{\sigma_e^2}x}{\frac{\gamma_0}{s}+\sigma_e^2\frac{\ln(x)}{x}}\right)\,dx \\
&\leq 0.5\left[1-e^{-\gamma_0/s}\right]+0.5\left(\frac{\gamma_0}{s}\right)^2\int_{x=1}^{\infty}\frac{e^{-(1+\frac{\rho_T^2}{\sigma_e^2})\frac{\gamma_0}{s}x}}{\frac{\gamma_0}{s}+\sigma_e^2\frac{\ln(x)}{x}}\exp\left(\frac{-\left(\frac{\gamma_0}{s}\right)^2\frac{\rho_T^2}{\sigma_e^2}x}{\frac{\gamma_0}{s}+\sigma_e^2\frac{1}{e}}\right)\,dx \\
&= 0.5\left[1-e^{-\gamma_0/s}+\left(\frac{\gamma_0}{s}\right)^2\int_{x=1}^{\infty}\frac{e^{-\alpha' x}}{\frac{\gamma_0}{s}+\sigma_e^2\frac{\ln(x)}{x}}\,dx\right] \\
&\leq 0.5\left[1-e^{-\gamma_0/s}+\left(\frac{\gamma_0}{s}\right)^2\int_{x=1}^{\infty}e^{-\alpha' x}\frac{1}{\frac{\gamma_0}{s}}\,dx\right] \\
&= 0.5\left[1-e^{-\gamma_0/s}+\left(\frac{\gamma_0}{s}\right)\frac{e^{-\alpha'}}{\alpha'}\right] \\
&\leq 0.5\left[\frac{\gamma_0}{s}+\left(\frac{\gamma_0}{s}\right)\frac{e^{-\alpha'}}{\alpha'}\right] \\
&= 0.5\frac{\gamma_0}{s}\left[1+\frac{e^{-\alpha'}}{\alpha'}\right] \\
&\leq ke^{-\sqrt{2s}}+ke^{-\sqrt{2s}}\frac{e^{-\alpha'}}{\alpha'} \qquad (22)
\end{aligned}
$$

where

$$
\alpha' = \left[1+\frac{\rho_T^2}{\sigma_e^2}\right]\frac{\gamma_0}{s}+\frac{\left(\frac{\gamma_0}{s}\right)^2\frac{\rho_T^2}{\sigma_e^2}}{\frac{\gamma_0}{s}+\frac{\sigma_e^2}{e}}. \qquad (23)
$$

We plot the BER (calculated from Eqn. 21) of the power allocation policy defined in Eqn. 5 for different values of ρ_T of a SISO system in Fig. 4. Rayleigh fading is assumed and the average fading gain is taken to be one. It is evident from Fig. 4 that as the quality of CSI at the transmitter degrades, then the BER performance loses exponential fall at high SNR. For $\rho_T = 0.99$,(which implies an error of the order of 15% in the estimation of $\mathbf{H}$) exponential diversity is lost in high SNR regime. Another interesting observation here is that at SNR $\leq$ 8 dB (fortunately, the region of most practical interest), the BER decay is exponential. This means that the BER performance in low SNR region is inherently limited by the additive noise of the channel and in high SNR region it is limited by the channel estimation error. One interesting point to be deduced

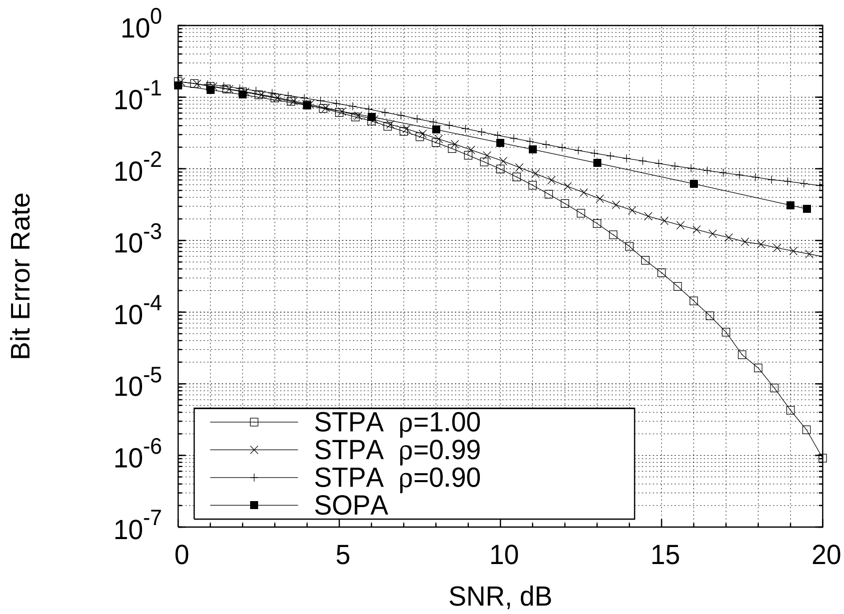

Figure 4: BER versus average SNR for a SISO system with Imperfect CSIT.

from here is that the finite bit precision of the feedback (channel estimate) can be coarser at low SNR but should be finer at high SNR.

The behaviour of BER as a function of SNR for imperfect CSIT explained in the above paragraph is captured partly in the upper bound Eqn. 22. Fig. 5 shows the behaviour of the two terms in Eqn. 22 as a function of SNR. It is

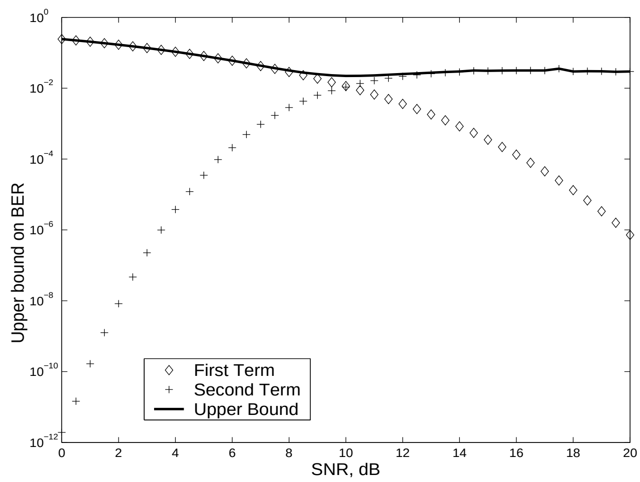

Figure 5: Upper bound versus average SNR of SISO system.

evident from Fig. 5 that upto a certain SNR, the first term dominates and later on the second term. Thus, one can expect an exponential fall of BER in the low SNR region (which is caused by the channel noise, since this term is also present in the perfect CSIT case). As SNR increases, the effect of estimation error is more prominent and the BER fall in SNR becomes polynomial.

3.2 MISO

Here, $\mathbf{H}$ represents the $1 \times n_t$ vector channel and let $\mathbf{H}_T = \rho_T\ \mathbf{H} + \sqrt{1-\rho_T^2}\ \mathbf{E}$ be the CSI available at the transmitter. The transmitter does beamforming by choosing the transmit weight vector $W_T = \mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|$. The instantaneous transmit power is chosen as if $\mathbf{H}_T$ were the true channel. Thus, the instantaneous transmit power is $P(s\|\mathbf{H}_T\|^2)$ (computed using Eqn. 5) and μ is given by $|\mathbf{H}W_T|^2 = |\mathbf{H}\mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|\ |^2$. The SNR is $\gamma = \frac{1}{\sigma^2}\left|\frac{\mathbf{H}\mathbf{H}_T^{\dagger}}{\|\mathbf{H}_T\|}\right|^2 = s\mu$ and the BER for a given μ is

$$P_{b|\mu,\|\mathbf{H}_T\|^2} = \mathrm{Q}\left(\sqrt{2s\mu P(s\|\mathbf{H}_T\|^2)}\right). \tag{24}$$

From Eqn. 11 we obtain the cutoff

$$\frac{\gamma_0}{s} \leq ke^{-(n_t-1)s+\delta}.$$

Let $\mu = X^2 + Y^2$, where X is the real part of $\mathbf{H}\mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|$ and Y is the imaginary part of $\mathbf{H}\mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|$, i.e., $X = \rho_T\|\mathbf{H}_T\| + \mathrm{Re}\Big(\sqrt{1-\rho_T^2}\ \mathbf{E}\mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|\Big)$ and $Y = \mathrm{Im}\Big(\sqrt{1-\rho_T^2}\ \mathbf{E}\mathbf{H}_T^{\dagger}/\|\mathbf{H}_T\|\Big)$. The random variables X and Y are real Gaussian variables (given $\mathbf{H}_T$) with mean m_X and m_Y respectively. Clearly, $m_X^2 + m_Y^2 = \|\mathbf{H}_T\|^2\rho_T^2$ and the variance of X and Y is $\frac{1}{2}\sigma_e^2 = \frac{1}{2}(1-\rho_T^2)$. The conditional pdf $f_{\mu\big|\|\mathbf{H}_T\|^2}(\lambda|\alpha)$ is *noncentral chi-square* distribution,

$$f_{\mu\big|\|\hat{h}\|^2}(\lambda\big|\alpha) = \frac{1}{\sigma_e^2}e^{-\left(\alpha\rho_T^2+\lambda\right)/\sigma_e^2}\sum_{k=0}^{\infty}\frac{\left(\frac{-\lambda\alpha\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)}.$$

The average BER is given by

$$\begin{aligned} P_b &= 0.5F_{\|\mathbf{H}_T\|^2}(\gamma_0/s) + \\ &\quad \int_{\alpha=\gamma_0/s}^{\infty}\int_{\lambda=0}^{\infty}\mathrm{Q}\left(\sqrt{2s\lambda P\left(s\alpha\right)}\right)f_{\mu\big|\|\mathbf{H}_T\|^2}(\lambda|\alpha)d\lambda f_{\|\mathbf{H}_T\|^2}(\alpha)d\alpha \\ &\leq 0.5F_{\|\mathbf{H}_T\|^2}(\gamma_0/s) + \int_{\alpha=\gamma_0/s}^{\infty}\int_{\lambda=0}^{\infty}\frac{1}{2}e^{-s\lambda P(s\alpha)}f_{\mu\big|\|\mathbf{H}_T\|^2}(\lambda|\alpha)d\lambda f_{\|\mathbf{H}_T\|^2}(\alpha)d\alpha \\ &= 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right] + \\ &\quad 0.5\int_{\alpha=\gamma_0/s}^{\infty}\left[\int_{\lambda=0}^{\infty}e^{-s\lambda P(s\alpha)}f_{\mu\big|\|\mathbf{H}_T\|^2}(\lambda|\alpha)d\lambda\right]f_{\|\mathbf{H}_T\|^2}(\alpha)d\alpha \end{aligned} \tag{25}$$

$$
\begin{aligned}
= \;& 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+ \\
& 0.5\int_{\alpha=\gamma_0/s}^{\infty}\frac{e^{-\alpha}\alpha^{n_t-1}}{(n_t-1)!} \\
& \left[\int_{\lambda=0}^{\infty}e^{-\frac{\lambda}{\alpha}\ln\left(\frac{\alpha}{\gamma_0/s}\right)}\frac{1}{\sigma_e^2}e^{-\left(\alpha\rho_T^2+\lambda\right)/\sigma_e^2}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{\lambda\alpha\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)}d\lambda\right]d\alpha \\
= \;& 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+ \\
& 0.5\frac{1}{\sigma_e^2}\int_{\alpha=\gamma_0/s}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{\alpha\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)}e^{-\left(\alpha\rho_T^2\right)/\sigma_e^2} \\
& \left[\int_{\lambda=0}^{\infty}\lambda^k e^{-\left[\frac{\ln\left(\frac{\alpha}{\gamma_0/s}\right)}{\alpha}+\frac{1}{\sigma_e^2}\right]\lambda}d\lambda\right]\frac{e^{-\alpha}\alpha^{n_t-1}}{(n_t-1)!}d\alpha \\
= \;& 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+ \\
& 0.5\frac{1}{\sigma_e^2}\int_{\alpha=\gamma_0/s}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{\alpha\rho_T^2}{\sigma_e^4}\right)^k}{k!\Gamma(k+1)}e^{-\alpha\rho_T^2\sigma_e^2} \\
& \left[\frac{k!}{\left(\frac{\ln\left(\frac{\alpha}{\gamma_0/s}\right)}{\alpha}+\frac{1}{\sigma_e^2}\right)^{k+1}}\right]\frac{e^{-\alpha}\alpha^{n_t-1}}{(n_t-1)!}d\alpha \\
= \;& 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+ \\
& 0.5\frac{1}{\sigma_e^2}\int_{x=1}^{\infty}\sum_{k=0}^{\infty}\frac{(-1)^k\left(\frac{(\gamma_0/s)\rho_T^2x}{\sigma_e^4}\right)^k}{k!}e^{-\left(1+\rho_T^2/\sigma_e^2\right)\gamma_0/s\ x} \\
& \left[\frac{1}{\left(\frac{\ln(x)}{\gamma_0/s\ x}+\frac{1}{\sigma_e^2}\right)}\right]^{k+1}\frac{(\gamma_0/s\ x)^{n_t-1}}{(n_t-1)!}\frac{\gamma_0}{s}dx \\
= \;& 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+0.5\frac{\left(\frac{\gamma_0}{s}\right)^{n_t+1}}{(n_t-1)!}\int_{x=1}^{\infty}\left(\frac{e^{-(1+\frac{\rho_T^2}{\sigma_e^2})\frac{\gamma_0}{s}x}}{\frac{\sigma_e^2\ln(x)}{x}+\gamma_0/s}\right) \\
& \exp\left(\frac{-\left(\frac{\gamma_0}{s}\right)^2\frac{\rho_T^2}{\sigma_e^2}x}{\frac{\gamma_0}{s}+\sigma_e^2\frac{\ln(x)}{x}}\right)x^{n_t-1}dx
\end{aligned}
$$

$$
\begin{aligned}
&= 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+0.5\frac{\left(\frac{\gamma_0}{s}\right)^{n_t+1}}{(n_t-1)!}\int_{x=1}^{\infty}\left(\frac{e^{-\alpha' x}}{\frac{\sigma_e^2\ln(x)}{x}+\gamma_0/s}\right)x^{n_t-1}dx\\
&\leq 0.5\left[1-\sum_{i=0}^{n_t-1}\frac{e^{-\gamma_0/s}(\gamma_0/s)^i}{i!}\right]+0.5\frac{\left(\frac{\gamma_0}{s}\right)^{n_t+1}}{(n_t-1)!}\int_{x=1}^{\infty}e^{-\alpha' x}\frac{1}{\gamma_0/s}x^{n_t-1}dx\\
&= 0.5\left[1-e^{-\gamma_0/s}\sum_{i=0}^{n_t-1}\frac{(\gamma_0/s)^i}{i!}\right]+0.5\left(\frac{\frac{\gamma_0}{s}}{\alpha'}\right)^{n_t}e^{-\alpha'}\sum_{j=0}^{n_t-1}\frac{{\alpha'}^j}{j!}\\
&\leq 0.5\left[1-e^{-\gamma_0/s}\sum_{i=0}^{n_t-1}\frac{(\gamma_0/s)^i}{i!}\right]+0.5\left(\frac{\frac{\gamma_0}{s}}{\alpha'}\right)^{n_t}e^{-\alpha'}\sum_{j=0}^{n_t-1}\frac{{\alpha'}^j}{j!}\\
&\leq 0.5\left[1-e^{-\gamma_0/s}\right]+0.5\left(\frac{\frac{\gamma_0}{s}}{\alpha'}\right)^{n_t}e^{-\alpha'}e^{\alpha'}\\
&= 0.5\left[1-e^{-\gamma_0/s}\right]+0.5\left(\frac{\frac{\gamma_0}{s}}{\alpha'}\right)^{n_t}\\
&\leq 0.5\left[\frac{\gamma_0}{s}\right]+0.5\left(\frac{\frac{\gamma_0}{s}}{\alpha'}\right)^{n_t}\\
&\leq 0.5ke^{-(n_t-1)s+\delta}\left[1+k^{n_t-1}\frac{e^{-(n_t-1)^2s+(n_t-1)\delta}}{{\alpha'}^{n_t}}\right]
\end{aligned}
$$

where α' is given in Eqn. 23.

The BER performance of (4,1) system is shown in Fig. 6. Here also, we observe the exponential fall of BER at low SNR. But, at high SNR the exponential decay is lost. We have also studied the BER performance of (2,1) and

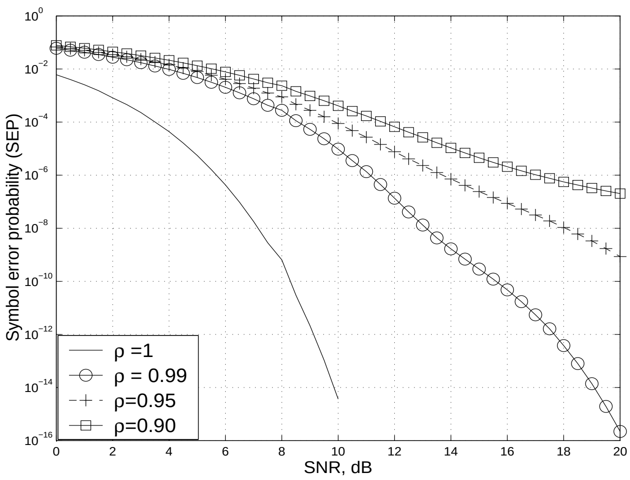

Figure 6: BER versus average SNR for different values of ρ of MISO (4,1) system.

(3,1) systems and found that the behaviour of these systems is similar to that of the (4,1) system. This behaviour can be again explained from Eqn. 26 in the same way as is done in Fig. 5.

3.3 MIMO

As in the perfect CSIT case, the transmitter does beamforming by forming a transmit weight vector, W_T which is the eigenvector corresponding to the largest eigenvalue of $\mathbf{H}_T^\dagger \mathbf{H}_T$. Thus, the signal model is

$$Y = \sqrt{P(s\|\mathbf{H}_T W_T\|^2)}\ \mathbf{H} W_T x + N$$

and the power allocation policy is given by $P(s\|\mathbf{H}_T W_T\|^2)$ (See Eqn. 5). We define μ as $\|\mathbf{H}W_T\|^2$. The average BER is

$$P_b = \int_{\mu=0}^{\infty}\int_{\lambda'=0}^{\infty} \mathrm{Q}\left(\sqrt{2s\mu \mathrm{P}(s\lambda')}\right) f_{\mu,\lambda'}(\mu,\lambda^{'})\ d\mu\ d\lambda^{'}. \tag{26}$$

As the joint density function $f_{\mu,\lambda'}(\mu,\lambda^{'})$ is not analytically tractable, we simulate the system for $n_t = 4$ and $n_r = 2$ for different values of ρ_T. The BER performance results are provided in Fig. 7. The exponential fall in low SNR

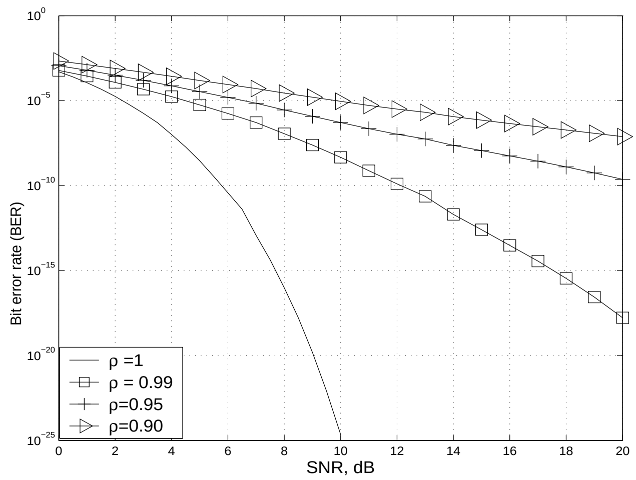

Figure 7: BER versus average SNR for different values of ρ of MIMO (4,2) system.

regime is observed in this case also. Similarly, one can observe at high SNR, the imperfect CSIT causes the BER to fall polynomially with SNR. Comparing Figs. 6 and 7, we can see that as n_r increases from 1 to 2, there is a substantial increase in diversity gain. For $\rho_T = 1$, there is a gain in SNR of 3 dB for achieving a BER of 10^{-5} by (4,2) system over (4,1) system. The diversity gain diminishes as ρ_T becomes small. For $\rho_T = 0.90$, we can see that the BER performance of (4,1) and (4,2) systems is almost the same.

4 Beamforming System With Imperfect CSIR

The system model considered here is the same as in Sec. 3. except that the channel estimate at the receiver, $\mathbf{H}_R$ is not perfect. Let $x \in \{+1, -1\}$ be the symbol to be transmitted (i.e., BPSK modulation is assumed). The transmitter does *beamforming* and transmits the vector $X = W_T x$, where W_T is the beamforming vector. The output of the matched filter is sampled at symbol duration and the received complex signal is given by

$$Y_k \quad = \quad \sqrt{\mathrm{P}(\gamma_T)}\ \mathbf{H}_k W_{T\,k} x + N_k, \quad k = 0, 1, 2, \ldots, \tag{27}$$

where γ_T is the SNR estimate at the transmitter. As explained before (in Sec. 2), we drop the time index k. For SOPA, $\mathrm{P}(\gamma_T) = 1$, whereas for STPA, $\mathrm{P}(\gamma_T)$ is given in Eqn. 5. There are two possible scenarios that may arise viz. the receiver may or may not know the CSIT(H_T and W_T) which are necessary for deciding the bit. The penaltys for not knowing the CSIT is different in MIMO and MISO cases as we shall show. We consider MISO and MIMO systems with imperfect CSIT and imperfect CSIR in the following (SISO case is just a special case of MISO).

4.1 MISO

We consider a system with $n_t \geq 1$ transmit antennas and $n_r = 1$ receive antenna. The optimal beamforming vector is given by $W_T = \frac{\mathbf{H}_T^\dagger}{\|\mathbf{H}_T\|}$ and the SNR estimate at the transmitter, γ_T is $s\|\mathbf{H}_T\|^2$.

4.1.1 CSIT not known at the receiver

We first consider the case when the receiver does not know CSIT. The decoding will be done as per equation (2) (with $\mathbf{H_R}$ and W_R used instead of $\mathbf{H}$ and W). It is to be noted that only the sign of the real part of $\frac{(\mathbf{H_R} W_R)^\dagger Y}{\sqrt{P(\gamma_R)}||\mathbf{H_R} W_R||^2}$ is of importance and not its magnitude. From (2) we can say that the imaginary part of the above quantity does not influence the decision, since the magnitude of the imaginary part of $\hat{x}$ does not depend on x.
$\frac{(\mathbf{H_R} W_R)^\dagger}{\sqrt{P(\gamma_R)}||\mathbf{H_R} W_R||^2}$ will always be a positive number and so will have no role in the decoding process. Hence the channel state information at the receiver is not important for decoding nor does it influence the performance in single receiver antenna systems. When the CSIT is perfect , the analysis of Section 2 holds, but when the CSIT is imperfect, the analysis of Section 3 does not hold, because there it was assumed that the CSIT as well as perfect CSIR was available at the receiver. When the CSIT is imperfect and not available at the receiver, we analyze the performance here.

The sign of $Re(Y)$ is the decoding criteria. Thus the probability of error is given by

$$P_b \quad = \quad P[Re(Y) < 0/x = 1]P(x = 1) + P[Re(Y) > 0/x = -1]P(x = -1) \cdot$$

By symmetry,

$$\begin{aligned} P_b &= P[Re(Y) < 0/x = 1] \\ &= P[Re(\sqrt{P(\gamma_T)}\mathbf{H}W_T + N) < 0] \end{aligned}$$

As per the model assumed for imperfect CSIT in Section 3,

$$\mathbf{H} = \rho_T \mathbf{H_T} + \sqrt{1 - {\rho_T}^2}\mathbf{E}.$$

Hence,

$$\mathbf{H}W_T = \rho_T ||\mathbf{H}_T|| + \sqrt{1 - {\rho_T}^2}\mathbf{E}W_T.$$

Since the components of $\mathbf{E}$ are independent complex Gaussian random variables,

$$\sqrt{1 - {\rho_T}^2}\mathbf{E}W_T = N'$$

and N' is circularly symmetric complex Gaussian given $\mathbf{H}_T$ (and hence W_T) ,with variance $= 1 - \rho_T^2$, since $||W_T||^2 = 1$. We assume that the channel noise N and N' are independant. Then

$$Y = \rho_T ||\mathbf{H}_T|| \sqrt{P(\gamma_T)} + N''$$

where N'' is complex Gaussian with variance $1/s + (1 - \rho_T^2)P(\gamma_T)$. Therefore the performance of this system can be interpreted as the output of an AWGN channel with SNR $= \frac{\rho_T^2 s ||\mathbf{H}_T||^2 P(\gamma_T)}{1 + sP(\gamma_T)(1-\rho_T^2)}$. Thus the average BER is given by,

$$E_{\gamma_T}\left[Q\left(\sqrt{\frac{2\rho_T^2\gamma_T P(\gamma_T)}{1 + sP(\gamma_T)(1 - \rho_T^2)}}\right)\right] \tag{28}$$

since $\gamma_T = s||\mathbf{H}_T||^2$ by definition.

The distribution of γ_T is given by the $Erlang_{n_t}(1/s)$ distribution. We have numerically computed the above expression and plotted in Fig 8. We observe that at smaller values of SNR , the exponential diversity is maintained but at large values of SNR, the BER is more or less constant. This can be predicted just by having a closer look at (28).As $s \to 0$, (for $ln\gamma_0 < 0$, which happens for most SNR (SNR $>$ 2dB for (2,1)case)) $sP(\gamma_T)(1-\rho_T^2) \to 0$ and hence for smaller s, the term in the square root in (28) becomes approximately $2\rho_T^2\gamma_T P(\gamma_T)$ and hence by the results of Section 2, we get exponential diversity. As $s \to \infty$, $sP(\gamma_T)(1-\rho_T^2) \to \infty$ and hence the term in the square root tends to $\frac{2\rho_T^2||\mathbf{H_T}||^2}{(1-\rho_T^2)}$. Thus for large s , the BER becomes approximatley constant with changing SNR. In the intermediate values of channel SNR, we get diversity order changing from exponential to linear to sublinear to constant. The important parameter here is $sP(\gamma_T)(1 - \rho_T^2)$ the magnitude of which will decide the diversity order. An appropriate system can be designed such that at the working SNR levels, we can guarentee exponetial diversity , but the price will be a higher value of ρ_T , that is a better feedback and estimation. Also to be noted is that if we need to operate in higher regions of SNR, we should have even higher values of ρ_T, while if we desire to operate at lower SNR levels, we can work with lower values of ρ_T.

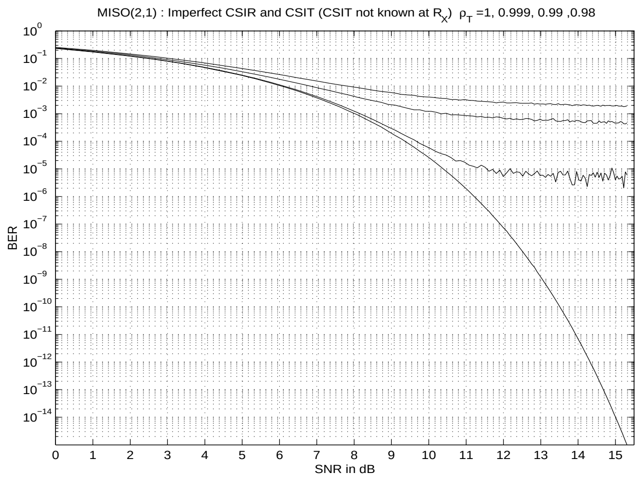

Figure 8: average BER versus SNR for different values of ρ_T of MISO (2,1) system, where the receiver does not know the CSIT.

4.1.2 CSIT known at the receiver

Here we consider the case when the CSIT ($\mathbf{H_T}$ and hence W_T) are known at the receiver. Suppose we decide to use the $\mathbf{H_T}$ for decoding in (2), we shall get exactly the same BER as in section 4.1.1, because for BPSK, it is only the sign of the real part of $\frac{(\mathbf{H_R}W_R)^{\dagger}Y}{\sqrt{P(\gamma_R)}||\mathbf{H_R}W_R||^2}$ that matters and by using $\mathbf{H_T}$ and W_T as $\mathbf{H_R}$ and W_R, we do not change the decision in anyway. But suppose we have a scenario where the receiver has a better channel estimate than the transmitter and also knows the CSIT, there arises a question as to which $\mathbf{H}$ to use for decoding. We shall address this question here.It is clear that we have to use W_T and not W_R for decoding ,in which case we get the same performance as in the previous section. Then we will use $\mathbf{H_R}$ and W_T and compare the BER with that of the previous section.The model assumed for $\mathbf{H_R}$ will be similar to the model for $\mathbf{H_T}$.

Define

$$\theta = (\mathbf{H_R}W_T)^{\dagger}(\mathbf{H}W_T\sqrt{P(\gamma_T)} + N).$$

The sign of the real part of θ will decide the decoding, and thus the probability of error is given by,

$$P_e = P[Re(\theta) < 0/x = 1]P(x = 1) + P[Re(\theta) > 0/x = -1]P(x = -1)$$

By symmetry,

$$P_e = P[Re(\theta) < 0/x = 1]\cdot$$

Since the model for $\mathbf{H_R}$ is similar to that of $\mathbf{H_T}$, from (19)

$$\mathbf{H_R} = \rho_R\mathbf{H} + \sqrt{1-\rho_R^2}\mathbf{E},$$

and therefore,

$$\begin{aligned}
[\theta/x=1] &= (\rho_R \mathbf{H} W_T + \sqrt{(1-\rho_R^2)}\mathbf{E} W_T)^\dagger (\mathbf{H} W_T \sqrt{P(\gamma_T)} + N) \\
&= \rho_R ||\mathbf{H} W_T||^2 \sqrt{P(\gamma_T)} + \sqrt{1-\rho_R^2}\sqrt{P(\gamma_T)} (\mathbf{E} W_T)^\dagger (\mathbf{H} W_T) \\
&\quad + (\rho_R \mathbf{H} W_T)^\dagger N + \sqrt{(1-\rho_R^2)} (\mathbf{E} W_T)^\dagger N.
\end{aligned}$$

Given ($\mathbf{H}$ and $\mathbf{H_T}$), the last term in the above expression is super-Gaussian and also has variance considerably lesser (because ρ_R is close to 1) than the other Gaussian noise terms. Thus, it can be neglected with respected to the other 2 Gaussian noise terms. So ,by arguments similar to Section 4.1.1, we can interpret the above situation as an AWGN channel and

$$[\theta/x=1] = \rho_R ||\mathbf{H} W_T||^2 \sqrt{P(\gamma_T)} + N'$$

where N'has a variance

$$\sigma^2 = \rho_R^2 \frac{||\mathbf{H} W_T||^2}{s} + (1-\rho_R^2) P(\gamma_T) ||\mathbf{H} W_T||^2.$$

Thus, the BER is given by ,

$$BER = E_{\mathbf{H},\mathbf{H_T}} \left[Q\left(\sqrt{\frac{2sP(\gamma_T)||\mathbf{H} W_T||^2}{1 + sP(\gamma_T)\frac{1-\rho_R^2}{\rho_R^2}}} \right) \right]. \tag{29}$$

Here too we see that the diversity order changes from exponential to sub-linear, for reasons similar to the ones mentioned in the previous section. See Fig 9 .By examining the denominators of (28) and (29), we can observe that when $\rho_R > \rho_T$ we can have a better performance if we decode using $\mathbf{H_R}$ as against using $\mathbf{H_T}$. If $\rho_R < \rho_T$, the CSIR should not be used for decoding.

Thus we may conclude that for the single antenna system, CSIR gives an advantage only if we know the CSIT at the receiver and the CSIR is better than the CSIT. If the CSIR is exact, then knowing CSIT at the receiver provoides polynomial diversity at high SNR as against constant (0 diversity) when CSIT is not known at receiver.

4.2 MIMO

Let $n_t > 1$ and $n_r > 1$. The optimal beamforming vector, W_T is given by the eigenvector corresponding to the largest eigenvalue (λ_T) of $\mathbf{H}_T^\dagger \mathbf{H}_T$ and the SNR estimate at the transmitter, γ_T is $s\lambda_T$.

The detection will be done by the rule (2).

$$\arg\min_x \left| x - \frac{(\mathbf{H}_R W_R)^\dagger Y}{\sqrt{P(\gamma_R)} \, || \, \mathbf{H}_R W_R \, ||^2} \right|^2 = \arg\min_x \; |x - \hat{x}|^2 \tag{30}$$

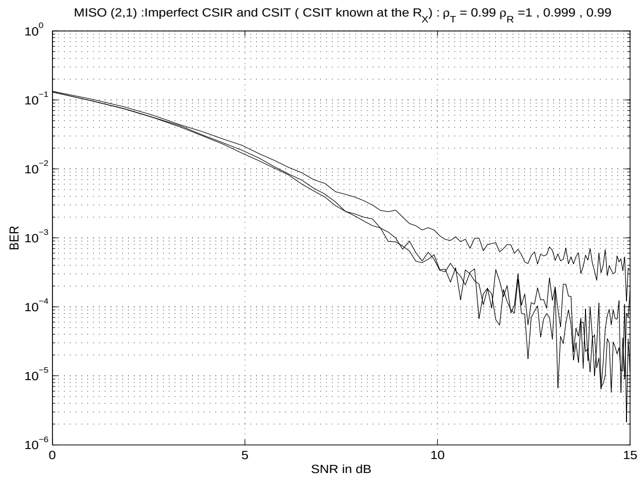

Figure 9: average BER versus SNR for different values of ρ_R when $\rho_R = 0.99$ for a MISO(2,1) system when the receiver knows CSIT.

where W_R is the eigenvector of $\mathbf{H}_R^\dagger \mathbf{H}_R$ corresponding to its largest eigenvalue, λ_R, $\gamma_R = s\lambda_R$ is the SNR estimate at the receiver, and

$$\begin{aligned} \hat{x} &\triangleq \frac{(\mathbf{H}_R W_R)^\dagger Y}{\sqrt{P(\gamma_R)}\|\mathbf{H}_R W_R\|^2} \\ &= r^\dagger Y \end{aligned} \tag{31}$$

where r is the receive weight vector used for detection.

Since we consider the BPSK modulation scheme (x is real), it is enough to consider only the real part of Eqn. 30(the imaginary part offers no performance improvement). The detection rule is

$$|1 - r^\dagger Y|^2 \underset{-1}{\overset{+1}{\lessgtr}} |1 + r^\dagger Y|^2$$

We observe that by virtue of BPSK modulation, only the sign and not the magnitude of $Re(r^\dagger Y)$ is important. We analyze the BER performance for the cases when the CSIT is known and when the CSIT is not known at the receiver. The model assumed for CSIT and CSIR will be the same as in the previous sections.

4.2.1 CSIT not known at the receiver

Since the receiver does not know CSIT, the decoding will have to be done using $\mathbf{H_R}$ and W_R which is the eigenvalue of $\mathbf{H_R}^\dagger \mathbf{H_R}$. Unlike the MISO case, now the W_T is very crucial. We simulate the system for $\rho_T = 0.99$ and $\rho_R = 0.9, 0.95, 0.99$ and 1.It is very clear from the Fig. 10 that even with the CSIR being perfect, but not the same as CSIT, a severe penalty has to be paid. The reason for this is that we do not know the beamformer exactly at the receiver. Most of the errors are due to this and not due to the channel noise.Thus when the transmit weight vector is not available at the reciver, we have to have a very

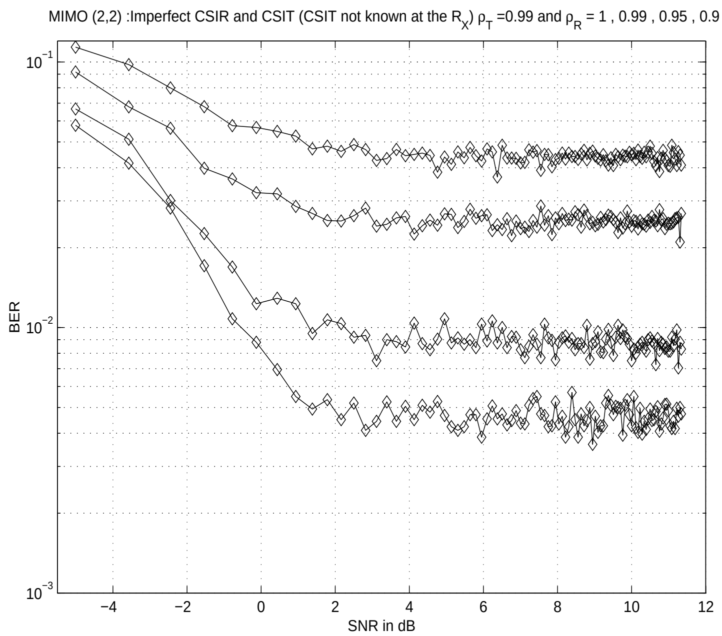

Figure 10: avearge BER versus SNR for a MIMO(2,2) system with imperfect CSIT(CSIT not known at the R_X),imperfect CSIR .

good estimate of the channel at the transmitter($\rho_T > 0.999$) as well as at the reciver ($\rho_R > 0.999$)to get a reasonable $BER \approx 10^-4$.

4.2.2 CSIT known at the receiver

Now the reciver knows the CSIT, and hence the transmit weight vector W_T. There arises a question as to which $\mathbf{H}$ ($\mathbf{H_R}$ or $\mathbf{H_T}$) should be used for decoding. We shall show that using the better $\mathbf{H}$(higher ρ) invariant of whether it is the transmitter's estimate or the receiver's estimate will give a better average BER. As said earlier, it is only the sign of $r^\dagger Y$ that will decide the bit and hence also the BER. Define $\theta = (\mathbf{H_R} W_T)^\dagger Y$. The probabilty of error (because of the symmetry), is given by

$$P_e = P[Re(\theta) < 0/x = 1].$$

Now,

$$\begin{aligned}
[\theta/x = 1] &= (\mathbf{H_R} W_T)^\dagger (\mathbf{H} W_T \sqrt{P(\gamma_T)} + N) \\
&= [(\rho_R \mathbf{H} + \sqrt{1 - \rho_R^2}\mathbf{E}) W_T]^\dagger (\mathbf{H} W_T \sqrt{P(\gamma_T)} + N) \\
&= \rho_R ||\mathbf{H} W_T||^2 \sqrt{P(\gamma_T)} + \rho_R (\mathbf{H} W_T)^\dagger N + \\
&\quad \sqrt{1 - \rho_R^2}\sqrt{P(\gamma_T)} (\mathbf{E} W_T)^\dagger (\mathbf{H} W_T) + \sqrt{1 - \rho_R^2} (\mathbf{E} W_T)^\dagger N.
\end{aligned}$$

The final term is super-gaussian and can be neglected with respect to the other two Gaussian noise random variables since it has significantly lesser variance(since ρ_R is close to 1)at high SNR cases, the case of interest to us. As in section 4.1.2, the above system can be interpreted as an AWGN channel and,

$$\theta \;\; = \;\; \rho_R ||\mathbf{H}W_T||^2 \sqrt{P(\gamma_T)} + N'$$

where N' is circularly symmetric complex Gaussian with variance

$$\sigma^2 = \rho_R^2 \frac{||\mathbf{H}W_T||^2}{s} + (1-\rho_R^2) P(\gamma_T) ||\mathbf{H}W_T||^2.$$

Therefore, the average BER is given by

$$E_{\mathbf{H},\mathbf{H_T}} \left[Q \left(\sqrt{\frac{2s||\mathbf{H}W_T||^2 P(\gamma_T)}{1 + sP(\gamma_T)\frac{1-\rho_R^2}{\rho_R^2}}} \right) \right]$$

It is to be noted that if we had used $\mathbf{H_T}$ instead of $\mathbf{H_R}$, we would have arrived at the same result(with ρ_T in place of ρ_T). Thus the choice of using $\mathbf{H_R}$ or $\mathbf{H_T}$ for decoding is to be decided by which of them is better. Similar to the arguments in section 4.1.1, we can explain the diversity order going from exponential to constant.Fig 11 has the simulation results.

From the results of the above two sections, we observe that knowledge of CSIT at the receiver is very important (especially for MIMO systems). The quality of CSIR becomes important only if it is better than the quality of CSIT available. In the scenario where CSIT is not available at the receiver, one should use CSIR for MIMO systems although we have observed that for MISO, it is not useful at all.

5 OSTBCs with Imperfect CSIT

In this section, we define the system model of a space-time coded system and analyze its BER performance with our power allocation policy defined in Eqn. 5. We follow the approach given in [7] and we borrow the results given in [7] for the sake of completeness. Let $\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_K\}$ denote the set of codewords, where K is the number of codewords. We assume that the codewords are of length T, i.e., each codeword is described by a $n_t \times T$ matrix. Actually, the codeword $\mathbf{C}_k$ is generated by a Space-Time encoder, $\overline{\mathbf{C}}_k$ followed by a precoder matrix, $\mathbf{W}_T$, i.e., $\mathbf{C}_k = \mathbf{W}_T \overline{\mathbf{C}}_k$ ($\mathbf{W}_T$ is chosen such that it optimizes certain performance metric). The channel between the j^{th} transmit antenna and the i^{th} receive antenna, h_{ij} is modeled as a zero mean complex Gaussian random variable. Thus, the $n_r \times n_t$ matrix, $\mathbf{H}$ represents the MIMO channel. We assume a *block fading* channel where T successive realizations of the channel corresponding to one codeword are same and the channel varies independently from one codeword to another. Let $\mathbf{C} \in \mathcal{C}$ be the transmitted codeword. The received complex signal vectors corresponding to one codeword may then be arranged in an $n_r \times T$ matrix $\mathbf{Y}$, given by

$$\mathbf{Y} \;\; = \;\; \mathbf{H}\mathbf{W}_T\overline{\mathbf{C}} \; + \; \mathbf{N},$$

where $\mathbf{Y} \in \mathbb{C}^{n_r \times T}$, $\overline{\mathbf{C}} \in \mathbb{C}^{n_t \times T}$ and $\mathbf{N} \in \mathbb{C}^{n_r \times T}$ represents the AWGN over T symbol durations at the n_r receive antennas.

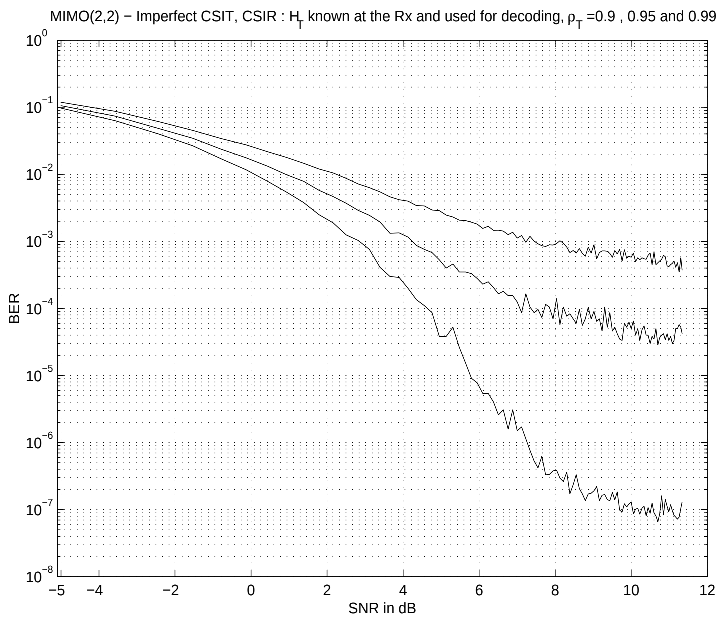

Figure 11: BER versus average SNR for a MIMO(2,2) system with imperfect CSIT(CSIT known at the R_X) and is used for decoding. CSIR quality is worse than CSIT.

The ML decoding of $\mathbf{C}$ given $\mathbf{Y}$ corresponds to

$$\hat{\mathbf{C}} = \arg\min_{\mathbf{C}\in\mathcal{C}} \|\mathbf{Y} - \mathbf{H}\mathbf{C}\|_F^2$$

where $\|.\|_F$ represents the Frobenius norm.

Let h =vec$(\mathbf{H}^\dagger)$[3]. we assume that $\mathbf{H}_T$ is the knowledge of the channel available at the transmitter. Also, let $\hat{h} = \text{vec}(\mathbf{H}_T) = [\hat{h}_1, \hat{h}_2, \ldots, \hat{h}_{n_r n_t}]^T$. Suppose, $\mathbf{C}_k$ is the transmitted codeword. Then, $P(\mathbf{C}_k \to \mathbf{C}_l \mid \hat{h})$, the probability of wrongly decoding C_k as C_l given the side information $\hat{h}$ is called the pairwise error probability (PEP) between $\mathbf{C}_k$ and $\mathbf{C}_l$. The overall design goal is to find $\mathbf{W}_T$ which minimizes PEP with respect to the set of all codewords. This is equivalent to finding the optimum $\mathbf{W}_T$ which minimizes PEP for the worst-case pair $\mathbf{C}_k$, $\mathbf{C}_l$.

We start by conditioning on the true channel realization h and utilize a well-known upper bound on the Gaussian tail function to arrive at

$$P(\mathbf{C}_k \to \mathbf{C}_l \mid h, \hat{h}) \leq \frac{1}{2} e^{-d^2(\mathbf{C}_k,\mathbf{C}_l)/4\sigma^2} \tag{32}$$

where $d(\mathbf{C}_k, \mathbf{C}_l)$ is the Euclidean distance between the codewords $\mathbf{C}_k$ and $\mathbf{C}_l$,

$$\begin{aligned} d^2(\mathbf{C}_k, \mathbf{C}_l) &= \|\mathbf{H}(\mathbf{C}_k - \mathbf{C}_l)\|_F^2 \\ &= \text{tr}(\mathbf{H}(\mathbf{C}_k - \mathbf{C}_l)(\mathbf{C}_k - \mathbf{C}_l)^\dagger \mathbf{H}^\dagger) \end{aligned}$$

[3] vec(.) is the vectorization operator which stacks the columns of its argument into a vector.

$$
\begin{aligned}
&= \operatorname{tr}(\mathbf{H}\ A(\mathbf{C}_k,\mathbf{C}_l)\ \mathbf{H}^\dagger) \\
&= \big(\operatorname{vec}(A(\mathbf{C}_k,\mathbf{C}_l)\ \mathbf{H}^\dagger)\big)^\dagger\ \operatorname{vec}(\mathbf{H}^\dagger) \\
&= h^\dagger\big(I_{n_r}\otimes A(\mathbf{C}_k,\mathbf{C}_l)\big)h
\end{aligned}
$$

where $\otimes$ represnts the Kronecker product and $A(\mathbf{C}_k,\mathbf{C}_l) = (\mathbf{C}_k-\mathbf{C}_l)(\mathbf{C}_k-\mathbf{C}_l)^\dagger$.

Since the true channel, h and the estimate $\hat{h}$ are jointly complex Gaussian, the pdf of h conditioned on $\hat{h}$ is given by

$$
p_{h|\hat{h}}(h|\hat{h}) = \frac{e^{-(h-m_{h|\hat{h}})^\dagger R_{hh|\hat{h}}^{-1}(h-m_{h|\hat{h}})}}{\pi^{n_t n_r}\det(R_{hh|\hat{h}})}
$$

where $m_{h|\hat{h}}$ and $R_{hh|\hat{h}}$ represent the conditional mean and covariance respectively. Then from Eqn. 32

$$
P(\mathbf{C}_k\to\mathbf{C}_l\mid\hat{h}) \le \int \frac{1}{2}e^{-d^2(\mathbf{C}_k,\mathbf{C}_l)/4\sigma^2}\ p_{h|\hat{h}}(h|\hat{h})\ dh.
$$

Let $\Psi(\mathbf{C}_k,\mathbf{C}_l) = (I_{n_r}\otimes A(\mathbf{C}_k,\mathbf{C}_l))/4\sigma^2 + R_{hh|\hat{h}}^{-1}$. The exponent of the above integral is

$$
\begin{aligned}
& \frac{-h^\dagger\big(I_{n_r}\otimes A(\mathbf{C}_k,\mathbf{C}_l)\big)h}{4\sigma^2} - (h-m_{h|\hat{h}})^\dagger R_{hh|\hat{h}}^{-1}(h-m_{h|\hat{h}}) \\
= \; & -h^\dagger\big(\Psi(\mathbf{C}_k,\mathbf{C}_l) - R_{hh|\hat{h}}^{-1}\big)h - (h-m_{h|\hat{h}})^\dagger R_{hh|\hat{h}}^{-1}(h-m_{h|\hat{h}}) \\
= \; & -h^\dagger\Psi(\mathbf{C}_k,\mathbf{C}_l)h + h^\dagger R_{hh|\hat{h}}^{-1}h - h^\dagger R_{hh|\hat{h}}^{-1}h + m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}h + h^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} - \\
& m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -h^\dagger\Psi(\mathbf{C}_k,\mathbf{C}_l)h + m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}h + h^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} - m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -h^\dagger(\Psi(\mathbf{C}_k,\mathbf{C}_l)h - R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}h - m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -h^\dagger\Psi(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}h - m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} + \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}})^\dagger\Psi(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + \\
& m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}h - m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}})^\dagger\Psi(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + \\
& m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}\Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} - m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} \\
= \; & -(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}})^\dagger\Psi(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + \\
& m_{h|\hat{h}}^\dagger R_{hh|\hat{h}}^{-1}(\Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}} - R_{hh|\hat{h}}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) \\
= \; & -(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}})^\dagger\Psi(h - \Psi^{-1}R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}) + \\
& m_{h|\hat{h}}^{*}R_{hh|\hat{h}}^{-1}(\Psi^{-1} - R_{hh|\hat{h}})R_{hh|\hat{h}}^{-1}m_{h|\hat{h}}.
\end{aligned}
$$

Thus, the PEP between $\mathbf{C}_k$ and $\mathbf{C}_l$ is bounded by

$$
\begin{aligned}
& P(\mathbf{C}_k \to \mathbf{C}_l \mid \hat{h}) \\
\leq \; & \frac{1}{2} \int \frac{e^{-(h-\Psi^{-1}R^{-1}_{hh|\hat{h}} m_{h|\hat{h}})^* \Psi (h-\Psi^{-1}R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}) + m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} (\Psi^{-1} - R_{hh|\hat{h}}) R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\pi^{n_t n_r} \det(R_{hh|\hat{h}})} \, dh \\
= \; & \frac{1}{2} \int \frac{e^{-(h-\Psi^{-1}R^{-1}_{hh|\hat{h}} m_{h|\hat{h}})^* \Psi (h-\Psi^{-1}R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}) + m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} (\Psi^{-1} - R_{hh|\hat{h}}) R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\pi^{n_t n_r} \det(R_{hh|\hat{h}}) \det(\Psi^{-1}) \det(\Psi)} \, dh \\
= \; & \frac{1}{2} \frac{e^{m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} (\Psi^{-1} - R_{hh|\hat{h}}) R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\det(R_{hh|\hat{h}}) \det(\Psi)}.
\end{aligned}
$$

From the above equation, we see that the objective function to minimize the PEP is given by,

$$
\begin{aligned}
V(\mathbf{C}_k, \mathbf{C}_l) &= \frac{e^{m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} (\Psi^{-1} - R_{hh|\hat{h}}) R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\det(R_{hh|\hat{h}}) \det(\Psi)} \\
&= \left\{ \frac{e^{m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} \Psi^{-1} R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\det(\Psi)} \right\} \frac{e^{-m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\det(R_{hh|\hat{h}})} \\
&= \left\{ e^{l(\mathbf{C}_k, \mathbf{C}_l)} \right\} \frac{e^{-m^*_{h|\hat{h}} R^{-1}_{hh|\hat{h}} m_{h|\hat{h}}}}{\det(R_{hh|\hat{h}})}
\end{aligned}
$$

Hence, we need to minimize $l(\mathbf{C}_k, \mathbf{C}_l)$ for the worst case codeword pairs $(\mathbf{C}_k, \mathbf{C}_l)$. This optimization problem is solved by Jongren [7]. Since, [7] solves the above optimization problem with trace$(\mathbf{W}\mathbf{W}^\dagger) = 1$ as a constraint ($\mathbf{W}$ is the precoder matrix), the solution of this problem will provide the precoder matrix $\mathbf{W}$, which will allocate the power optimally in space. The second term in the above equation is a constant and does not depend on the power allocation policy. We combine our power allocation allocation policy given by Eqn. 5 with that of Jongren's policy to get an optimal power allocation in space and time for Orthogonal Space-Time Block Codes (OSTBCs).

We now look into the Alamouti code and the optimal precoder matrix W_T that minimizes the PEP for the worst case pair ($\mathbf{C}_k$, $\mathbf{C}_l$). [7] gives closed form expression for W_T for the Alamouti code,

$$
\overline{\mathbf{C}} = \begin{bmatrix} s_1 & s_2 \\ -s_2^* & s_1^* \end{bmatrix}
$$

and $\mathbf{C} = \mathbf{W}_T \overline{\mathbf{C}}$. For the Alamouti code, the optimum $\mathbf{W}_T$ (for SOPA) is given by

$$
\mathbf{W}_{T,\text{ SOPA}} = \frac{1}{\sqrt{|\hat{h}_1|^2 + |\hat{h}_2|^2}} \begin{bmatrix} -\hat{h}_2^* & \hat{h}_1 \\ \hat{h}_1^* & \hat{h}_2 \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}
$$

The following algorithm calculates λ_1 and λ_2. We will use the following notation. Let $\alpha = 1 - \rho_T^2$. The OSTBCs (e.g., Alamouti code) have the property that $A(\overline{\mathbf{C}}_K, \overline{\mathbf{C}}_l) = \mu_{kl} I_{n_t}$, $\forall k \neq l$. Let μ_{min} be defined by $\min_{k,l} \mu_{kl}$ and μ_{min} = 4 for Alamouti code. Denote η by $\mu_{\text{min}}/4\sigma^2 = 1/\sigma^2$. Let $\kappa = \alpha(2+\alpha\eta)$. Let

$$\mu = \frac{\eta\left(3\kappa + |\rho|^2\|\hat{h}\|^2 + \sqrt{6\kappa|\rho|^2\|\hat{h}\|^2 + |\rho|^4\|\hat{h}\|^4 + \kappa^2}\right)}{2(2+\alpha\eta)^2}$$

and $\lambda = \frac{1}{\mu} - \frac{1}{\alpha\eta}$. If $\lambda > 0$ then set $\lambda_1 = \lambda$ and set $\lambda_2 = 1$ - λ. If $\lambda \leq 0$ then set $\lambda_1 = 0$ and set $\lambda_2 = 1$.

The optimum $\mathbf{W}_T$ for Space-Time power allocation is given by

$$\mathbf{W}_{T,\ \text{STPA}} = \frac{\sqrt{P\left(s(|\hat{h}_1|^2 + |\hat{h}_2|^2)\right)}}{\sqrt{|\hat{h}_1|^2 + |\hat{h}_2|^2}} \begin{bmatrix} -\hat{h}_2^* & \hat{h}_1 \\ \hat{h}_1^* & \hat{h}_2 \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{bmatrix}.$$

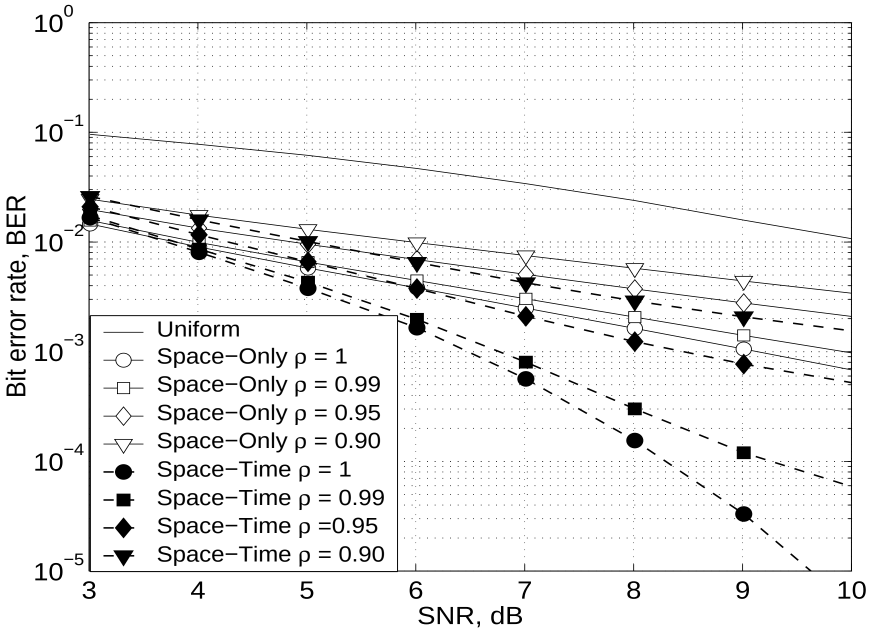

Figure 12: BER versus average SNR for an Alamouti and convolutionally coded $(G(D) = \{1+D^2, 1+D+D^2\})$ (2,1) system with Imperfect CSIT.

Fig. 8 shows the BER performance of a coded system with an outer rate-1/2 *convolutional code* and an inner *Alamouti code* for $n_t = 2$ and $n_r = 1$. We observe that for the perfect CSIT case, our power allocation policy provides *exponential order diversity* gain which is substantially more than the conventional *space-only (SOPA)* and *uniform* power allocation schemes. Also, when the quality of CSIT degrades ($\rho \neq 1$), the exponential diversity is lost at high SNR for our policy. The interesting observation here is that we still achieve exponential diversity at low SNR (up to 9 dB in the systems we studied). In Fig. 8, we note that at $\rho = 1$ and $P_e = 10^{-3}$ the gain in STPA compared to SOPA is 2.5 dB and at $\rho = 0.99$ the gain is 3.2 dB. As, ρ can be as large as 0.99 for practical systems (the channel estimation in GSM is done by a 26-bit midamble sequence; a ML estimate of H from this sequence can provide ρ as large as 0.99 for an SNR of 10 dB) our power allocation policy provides significant improvement in BER performance over the existing policies.

The fading in our system is assumed to be i.i.d.. Therefore, interleaving will not improve the performance.

6 Conclusions

We have analyzed the power allocation problem in space and time for Rayleigh fading channels. Similar Analysis can also be done for other fading distributions. We have shown that the optimal power allocation in space and time substantially reduces the BER. We observed that in the beamforming case, when perfect CSIT is available, our power allocation policy provides *exponential order diversity* gain for both coded and uncoded systems. The diversity gain that we achieve is substantially more than the conventional Space-Only and uniform power allocation schemes. Also, when the quality of CSIT degrades, the exponential diversity is lost at high SNR. The interesting observation here is we still achieve exponential diversity at low SNR.Furthermore, we have found that it is important to have CSIT available at the receiver also, more so for $n_t > 1$.

We also study the BER performance of S-T codes (Alamouti code) with our Space-Time power allocation scheme. We have observed that our power allocation policy improves the BER performance of S-T codes also. Infact, the performance of S-T codes for imperfect CSIT case is better than the power allocation policy given by [7] and the beamforming as observed in [9] and [12] for space-only power control.

Acknowledgment

This work has benifitted from discussions with Raghava H Nanjundaswamy, Dept of Electrical Engineering , IIT Madras, Chennai.

APPENDIX-1: An Inequality

Let $f(x)$ = -ln$(1-x)$ defined in the interval $x \in [0,1)$. Then, it is clear that $f(x)$ is a convex increasing function of x. Thus, for all $x \in [0, x_0], (x_0 \ll 1), f(x) \leq kx$, where k depends on x_0 and is given by $k = f(x_0)/x_0$. We use this bound extensively in our derivations.

APPENDIX-2: Calculations for MISO, Perfect CSIT

$$
\begin{aligned}
\text{Let } I_m &= \int_1^\infty x^m e^{-\alpha x} \ln(x)\, dx \\
&= x^m \ln(x) \frac{e^{-\alpha x}}{-\alpha} |_1^\infty + \frac{1}{\alpha} \int_1^\infty e^{-\alpha x} \left(\ln(x) m x^{m-1} + x^{m-1} \right) dx \\
&= \frac{m}{\alpha} I_{m-1} + \frac{1}{\alpha} U_{m-1} \\
&= \frac{m!}{\alpha^m} I_0 + \sum_{k=0}^{m-1} \frac{m!}{(k+1)!} \frac{1}{\alpha^{m-k}} U_k
\end{aligned}
$$

where

$$
\begin{aligned}
U_k &= \int_1^\infty x^k e^{-\alpha x} dx \\
&= \sum_{i=0}^{k} \frac{e^{-\alpha}}{\alpha} \frac{k!}{(k-i)!} \frac{1}{\alpha^i}
\end{aligned}
$$

Therefore

$$
\begin{aligned}
I_m &= \frac{m!}{\alpha^m} I_0 + \sum_{k=0}^{m-1} \frac{m!}{(k+1)!} \frac{1}{\alpha^{m-k}} \sum_{i=0}^{k} \frac{e^{-\alpha}}{\alpha} \frac{k!}{(k-i)!} \frac{1}{\alpha^i} \\
&= \frac{m!}{\alpha^m} I_0 + \frac{m! e^{-\alpha}}{\alpha^{m+1}} \sum_{k=0}^{m-1} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha^{k-i}}{(k-i)!} \\
&= \frac{m!}{\alpha^m} I_0 + \frac{m! e^{-\alpha}}{\alpha^{m+1}} \sum_{k=0}^{m-1} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha^i}{i!} \\
&= \frac{m!}{\alpha^m} \left[I_0 + \frac{e^{-\alpha}}{\alpha} \sum_{k=0}^{m-1} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha^i}{i!} \right]
\end{aligned}
$$

where

$$
\begin{aligned}
I_0 &= \int_1^\infty e^{-\alpha x} \ln(x) dx \\
&= \ln(x) \frac{e^{-\alpha x}}{-\alpha} |_1^\infty + \frac{1}{\alpha} \int_1^\infty e^{-\alpha x} \frac{1}{x} dx \\
&= \frac{1}{\alpha} \int_\alpha^\infty e^{-\alpha x} \frac{1}{\alpha x} d(\alpha x) \\
&= \frac{1}{\alpha} \int_\alpha^\infty \frac{e^{-t}}{t} dt
\end{aligned}
$$

Thus,

$$I_m = \frac{m!}{\alpha^{m+1}}\left[\int_\alpha^\infty \frac{e^{-t}}{t}dt + e^{-\alpha}\sum_{k=0}^{m-1}\frac{1}{k+1}\sum_{i=0}^{k}\frac{\alpha^i}{i!}\right]$$

APPENDIX-3: Calculations for MIMO, Perfect CSIT

In the following, we will consider $n_t > n_r$ and $n_t = n_r$ cases seperately.

Case1: $n_t > n_r$

From (7),

$$P = 0.5\left(\int_0^{\gamma_0} f_\gamma(\gamma)d\gamma + \int_{\gamma_0}^\infty e^{-ln(\gamma/\gamma_0)}f_\gamma(\gamma)d\gamma\right).$$

Using (15),

$$\begin{aligned}\int_0^{\gamma_0} f_\gamma(\gamma)d\gamma &= \int_0^{\gamma_0}\frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}(\gamma/s)^m e^{-i\gamma/s}}{m!}d\gamma \\ &= \frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}}{m!s^m}\int_0^{\gamma_0}\gamma^m e^{-i\gamma/s}d\gamma.\end{aligned}$$

Let

$$\begin{aligned}I'_m &= \int_0^\beta x^m e^{-\alpha x}dx \\ &= \frac{\beta^m e^{-\alpha\beta}}{-\alpha} + \frac{m}{\alpha}I'_{m-1} \\ &= \frac{m!}{\alpha^m}\left(I'_o - \frac{e^{-\alpha\beta}}{\alpha}\left[\sum_{k=1}^{m}\frac{(\alpha\beta)^k}{k!}\right]\right). \qquad (33)\end{aligned}$$

Since

$$\begin{aligned}I'_o &= \frac{1-e^{-\alpha\beta}}{\alpha}, \\ I'_m &= \frac{m!}{\alpha^{m+1}}\left(1 - e^{-\alpha\beta}\left[\sum_{k=0}^{m}\frac{(\alpha\beta)^k}{k!}\right]\right).\end{aligned}$$

Thus, we get

$$\int_0^{\gamma_0} f_\gamma(\gamma)d\gamma = \frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}s\left[1 - e^{-i\gamma_0/s}\left[\sum_{k=0}^{m}\frac{(i\gamma_0/s)^k}{k!}\right]\right]. \qquad (34)$$

Now consider the integral,

$$\begin{aligned}\int_{\gamma_0}^\infty \frac{1}{\gamma}f_\gamma(\gamma)d\gamma &= \frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\int_{\gamma_0}^\infty\frac{1}{\gamma}\frac{i^{m+1}(\gamma/s)^m e^{-i\gamma/s}}{m!}d\gamma \\ &= \frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}}{s^m m!}\int_{\gamma_0}^\infty\gamma^{m-1}e^{-i\gamma/s}d\gamma.\end{aligned}$$

Let,

$$
\begin{aligned}
I''_m &= \int_{\beta}^{\infty} x^m e^{-\alpha x} dx \\
&= \frac{\beta^m e^{-\alpha\beta}}{\alpha} + \frac{m}{\alpha} I''_{m-1} \\
&= \frac{m!}{\alpha^m} \left(I''_o + \frac{e^{-\alpha\beta}}{\alpha} \left[\sum_{k=1}^{m} \frac{(\alpha\beta)^k}{k!} \right] \right). \qquad (35)
\end{aligned}
$$

Since

$$
\begin{aligned}
I''_o &= \frac{e^{-\alpha\beta}}{\alpha} \\
I''_m &= \frac{m!}{\alpha^{m+1}} \left(e^{-\alpha\beta} \left[\sum_{k=0}^{m} \frac{(\alpha\beta)^k}{k!} \right] \right)
\end{aligned}
$$

Thus, we get

$$
\int_{\gamma_0}^{\infty} \frac{\gamma_0}{\gamma} f_\gamma(\gamma) d\gamma = \frac{\gamma_0}{s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left[e^{-i\gamma_0/s} \left[\sum_{k=0}^{m} \frac{(i\gamma_0/s)^k}{k!} \right] \right]. \qquad (36)
$$

On computation , we find that $\gamma_0/s \ll 1$ ($\gamma_0/s < 0.01$ for all $SNR > 3dB$). Then using (7), (34) , (36),

$$
\begin{aligned}
P_b &= \frac{1}{2s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \left\{ s \left(1 - e^{-i\gamma_0/s} \left[\sum_{k=0}^{m} \frac{(i\gamma_0/s)^k}{k!} \right] \right) + \frac{i\gamma_0}{m} \left[e^{-i\gamma_0/s} \left[\sum_{k=0}^{m} \frac{(i\gamma_0/s)^k}{k!} \right] \right] \right\} \\
&\approx \frac{1}{2s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \left\{ s \left(1 - e^{-i\gamma_0/s}[1] \right) + \frac{i\gamma_0}{m} \left[e^{-i\gamma_0/s} [e^{i\gamma_0/s}] \right] \right\} \\
&\approx \frac{1}{2s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \left\{ i\gamma_0 + \frac{i\gamma_0}{m} \right\} \\
&\approx \frac{\gamma_0}{2s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} i + \frac{\gamma_0}{s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m}
\end{aligned}
$$

Since the first term is zero,

$$
\begin{aligned}
P_b &\approx \frac{\gamma_0}{2s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \\
&= \frac{\gamma_0}{2s} K \qquad (37)
\end{aligned}
$$

where,

$$
K = \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m}. \qquad (38)
$$

Now let us calculate γ_0/s.

$$\begin{aligned}
1 &= E[P(\gamma)] \\
&= \int_{\gamma_0}^{\infty} \frac{1}{\gamma} ln\left(\frac{\gamma}{\gamma_0}\right) f_\gamma(\gamma) d\gamma \\
&= \int_{0}^{\gamma_0} \frac{1}{\gamma} ln\left(\frac{\gamma}{\gamma_0}\right) \frac{1}{s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i^{m+1}(\gamma/s)^m e^{-i\gamma/s}}{m!} d\gamma
\end{aligned}$$

which can be shown to be equal to

$$\frac{1}{s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{\alpha_i^m i}{m!} \int_1^{\infty} x^{m-1} e^{-\alpha_i x} ln(x) dx$$

where $\alpha_i = i\gamma_0/s$.As in(33) and (35), we can rewrite it as

$$1 = \frac{1}{s} \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left(\int_{\alpha}^{\infty} \frac{e^{-t}}{t} dt + e^{-\alpha_i} \sum_{k=0}^{m-2} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha_i^i}{i!} \right). \quad (39)$$

Using the fact that $\gamma_0/s \ll 1$,

$$\begin{aligned}
E1(\frac{\gamma_0}{s}) = \int_{\gamma_0/s}^{\infty} \frac{e^{-t}}{t} dt &= \int_{i\gamma_0/s}^{\infty} \frac{e^{-t}}{t} + \int_{\gamma_0/s}^{i\gamma_0/s} \frac{e^{-t}}{t} \\
&\approx \int_{i\gamma_0/s}^{\infty} \frac{e^{-t}}{t} + \int_{\gamma_0/s}^{i\gamma_0/s} \frac{1}{t} \\
&= \int_{i\gamma_0/s}^{\infty} \frac{e^{-t}}{t} + ln(i) \\
&= E_1(\frac{i\gamma_0}{s}) + ln(i) \quad (40)
\end{aligned}$$

where$E_1(t)$ is the exponential integral([2]).Thus from (39)

$$\begin{aligned}
s &= \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left(\int_{\alpha_i}^{\infty} \frac{e^{-t}}{t} dt + e^{-\alpha_i} \sum_{k=0}^{m-2} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha_i^i}{i!} \right) \\
&\approx \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left(E_1(\frac{\gamma_0}{s}) - ln(i) + e^{-\alpha_i} \sum_{k=0}^{m-2} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha_i^i}{i!} \right) \\
&= KE_1(\frac{\gamma_0}{s}) + c
\end{aligned}$$

and hence

$$\frac{(s-c)}{K} \approx E_1(\frac{\gamma_0}{s}) \quad (41)$$

where K was defined earlier and

$$
\begin{aligned}
c &= \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left(e^{-\alpha_i} \sum_{k=0}^{m-2} \frac{1}{k+1} \sum_{i=0}^{k} \frac{\alpha_i^i}{i!} - ln(i) \right) \\
&\approx \sum_{i=1}^{n_r} \sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i}{m} \left(\sum_{k=0}^{m-2} \frac{1}{k+1} - ln(i) \right). \qquad (42)
\end{aligned}
$$

From Theorm2 of [2] $E_1(\frac{\gamma_0}{s})$ is upper bounded by $-ln(1-e^{-\gamma_0/s})$. Thus,

$$
\begin{aligned}
-ln(1-e^{-\gamma_0/s}) &\gtrsim \frac{s-c}{K} \\
1-e^{-\gamma_0/s} &\lesssim e^{-\frac{s-c}{K}} \\
\frac{\gamma_0}{s} &\lesssim k\ exp(-\frac{s-c}{K}) \qquad (43)
\end{aligned}
$$

where k was defined in section 3.1. Hence from bound (37), P_b is

$$
P_b \lesssim \frac{Kk}{2} exp\left(-\frac{s-c}{K}\right). \qquad (44)
$$

The k for values of $SNR \geq 0dB$ is very close to 1 and can be left out in most of the cases. Also the bound (44) is very tight for $SNR \geq 3dB$.

Case2: $n_t = n_r$

The analysis is similar to Case1 until (34).
In the following we use the following property of the distribution function which we prove in Appendix-5,

$$
\sum_{i=1}^{n_t} i d_{i,0} = 0\cdot \qquad (45)
$$

when $n_t = n_r$. Now consider the integral,

$$
\begin{aligned}
\int_{\gamma_0}^{\infty} \frac{1}{\gamma} f_\gamma(\gamma) d\gamma &= \frac{1}{s} \sum_{i=1}^{n_r} \sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m} \int_{\gamma_0}^{\infty} \frac{1}{\gamma} \frac{i^{m+1}(\gamma/s)^m e^{-i\gamma/s}}{m!} d\gamma \\
&\quad + \frac{1}{s} \sum_{i=1}^{n_r} d_{i,0} i \int_{\gamma_0}^{\infty} \gamma_0 \frac{e^{-i\gamma/s}}{\gamma} d\gamma \\
&= \frac{1}{s} \sum_{i=1}^{n_r} \sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m} \frac{i^{m+1}}{s^m m!} \int_{\gamma_0}^{\infty} \gamma^{m-1} e^{-i\gamma/s} d\gamma + \\
&\quad \frac{\gamma_0}{s} \sum_{i=1}^{n_r} d_{i,0} i E_1(i\gamma_0/s).
\end{aligned}
$$

By (40),

$$\begin{aligned}
\frac{\gamma_0}{s}\sum_{i=1}^{n_r} d_{i,0} i E_1(i\gamma_0/s) &\approx \frac{\gamma_0}{s}\sum_{i=1}^{n_r} d_{i,0} i\left(E_1(\gamma_0/s) - ln(i)\right)\\
&= -\frac{\gamma_0}{s}\sum_{i=1}^{n_r} d_{i,0} i ln(i)
\end{aligned}$$

since the first term is zero because of (45). By (35)

$$\begin{aligned}
\int_{\gamma_0}^{\infty}\frac{\gamma_0}{\gamma}f_\gamma(\gamma)d\gamma &\approx \frac{\gamma_0}{s}\sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left[e^{-i\gamma_0/s}\left[\sum_{k=0}^{m}\frac{(i\gamma_0/s)^k}{k!}\right]\right] - \\
&\quad \frac{\gamma_0}{s}\sum_{i=1}^{n_r} d_{i,0} i ln(i).\\
&\approx K\frac{\gamma_0}{s}
\end{aligned}$$

$$\begin{aligned}
\text{where,}\qquad K &= \sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left[e^{-i\gamma_0/s}\left[\sum_{k=0}^{m}\frac{(i\gamma_0/s)^k}{k!}\right]\right] - \\
&\quad \sum_{i=1}^{n_r} d_{i,0} i ln(i).
\end{aligned} \tag{46}$$

Let us now evaluate γ_0/s. By the average power constraint,

$$\begin{aligned}
1 &= E[P(\gamma)]\\
&= \int_{\gamma_0}^{\infty}\frac{1}{\gamma}ln\left(\frac{\gamma}{\gamma_0}\right)\frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=n_t-n_r}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i^{m+1}(\gamma/s)^m e^{-i\gamma/s}}{m!}d\gamma.
\end{aligned}$$

As in case1,

$$\begin{aligned}
1 &= \frac{1}{s}\sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left(\int_{\alpha}^{\infty}\frac{e^{-t}}{t}dt + e^{-\alpha}\sum_{k=0}^{m-2}\frac{1}{k+1}\sum_{i=0}^{k}\frac{\alpha^i}{i!}\right) + \\
&\quad \frac{1}{s}\sum_{i=1}^{n_r} i d_{i,0}\int_{1}^{\infty}\frac{e^{(\frac{-i\gamma_0 x}{s})}ln(x)}{x}dx.
\end{aligned}$$

Now,

$$\begin{aligned}
\sum_{i=1}^{n_r} d_{i,0} i\int_{1}^{\infty}\frac{e^{(\frac{-i\gamma_0 x}{s})}ln(x)}{x}dx &\le \sum_{i=1}^{n_r} d_{i,0} i\int_{1}^{\infty}\frac{e^{(\frac{-i\gamma_0 x}{s})}ln(2x-1)}{x}dx\\
&= \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left[-E_1(\frac{\gamma_0}{s})\right]^2 \qquad \text{(by [5])}\\
&\approx \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left[E_1(\frac{\gamma_0}{s}) - ln(\frac{i}{2})\right]^2
\end{aligned}$$

$$
\begin{aligned}
= \quad & \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}[E_1(\frac{\gamma_0}{s})]^2 + \\
& \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left\{(ln(i/2))^2 - 2E_1(\frac{\gamma_0}{s})ln(i/2)\right\}.
\end{aligned}
$$

The first term is 0 by (45) and hence the right hand side in the above equation equals

$$
\begin{aligned}
& \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left\{(ln(i/2))^2 - 2E_1(\frac{\gamma_0}{s})ln(i/2)\right\} \\
= \quad & \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left\{(ln(i/2))^2 - 2E_1(\frac{\gamma_0}{s})ln(i) + 2ln(2)E_1(\frac{\gamma_0}{s})\right\} \\
= \quad & \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left\{(ln(i/2))^2 - 2E_1(\frac{\gamma_0}{s})ln(i)\right\}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
s \quad \lesssim \quad & \sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left(\int_{\alpha_i}^{\infty}\frac{e^{-t}}{t}dt + e^{-\alpha_i}\sum_{k=0}^{m-2}\frac{1}{k+1}\sum_{i=0}^{k}\frac{\alpha_i^i}{i!}\right) \\
& + \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}\left\{(ln(i/2))^2 - 2E_1(\frac{\gamma_0}{s})ln(i)\right\}.
\end{aligned}
$$

On simplification, we get

$$
E_1(\frac{\gamma_0}{s}) \gtrsim \frac{s-c}{K}
$$

where K was defined earlier and

$$
\begin{aligned}
c \quad = \quad & \sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left(e^{-\alpha}\sum_{k=0}^{m-2}\frac{1}{k+1}\sum_{i=0}^{k}\frac{\alpha^i}{i!} - ln(i)\right) + \\
& \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}(ln(i/2))^2 \\
\approx \quad & \sum_{i=1}^{n_r}\sum_{m=1}^{(n_t+n_r)i-2i^2} d_{i,m}\frac{i}{m}\left(\sum_{k=0}^{m-2}\frac{1}{k+1} - ln(i)\right) + \sum_{i=1}^{n_r} d_{i,0}\frac{i}{2}(ln(i/2))^2.
\end{aligned}
$$

Since the expressional form of the bound of $E_1(\frac{\gamma_0}{s})$ is the same as in case 1, the same expression will follow and so we obtain,

$$
P_b \lesssim \frac{Kk}{2}exp\left(-\frac{s-c}{K}\right). \tag{47}
$$

This upper bound is tight for $SNR \geq 3dB$.

APPENDIX-4: Proof of $\sum_{i=1}^{i=n_r} id_{i,0} = 0$

From (22) of [3], $c_{i,m}$ = coeffecient of $e^{-iu}u^m$ in $\frac{d}{du}\det(\mathbf{S}(\mathbf{u}))$. Now, from (20b) of [3]

$$(\mathbf{S}(\mathbf{u}))_{\mathbf{k},\mathbf{l}} = \mathbf{\Gamma}(\mathbf{n_r} - \mathbf{n_t} + \mathbf{k} + \mathbf{l} - \mathbf{1}, \mathbf{u})$$

where Γ is the incomplete Gamma function. Since $n_t = n_r$,

$$(\mathbf{S}(\mathbf{u}))_{\mathbf{k},\mathbf{l}} = \mathbf{\Gamma}(\mathbf{k} + \mathbf{l} - \mathbf{1}, \mathbf{u}) \tag{48}$$

and by (21) of [3]

$$\Gamma(k+1, u) = k! \left[1 - e^{-u} \sum_{m=0}^{k} \frac{u^m}{m!}\right]. \tag{49}$$

Let us have a closer look at $\det\mathbf{S}(\mathbf{u})$.It is a square matrix of dimensions n_t x n_t. Suppose we expand it like a determinant, the number of terms will be $n_t!$. Each of the $n_t!$ terms will be the product of n_t incomplete Gamma functions. If we were to differentiate these $n_t!$ terms one by one, by the product rule we will have $n_t!n_t$ terms in the expression for the differential of the determinant. We have to look at the coefficients of $e^{-iu}u^0$ and sum them up to get $c_{i,0}$. Let us look at the differential of one gamma function

$$\begin{aligned} \frac{d}{du}\Gamma(k+1, u) &= k!e^{-u} \sum_{m=0}^{k} \frac{u^m}{m!} - k!e^{-u} \sum_{m=1}^{k} \frac{u^{m-1}}{(m-1)!} \\ &= e^{-u}u^k. \end{aligned}$$

In the $n_t!n_t$ terms of the differential of the determinant we have to look for the terms that will contribute to the $e^{-iu}u^0$. By observing that each of the $n_t!n_t$ terms is a product of $n_t - 1$ Gamma functions and one differential, we can say that if a term were to contribute to $c_{i,0}$, it's differential term should not have any u terms. In other words the k (in $\Gamma(k+1, u)$) has to be zero in the Gamma function that was differentiated. Hence out of the $n_t!n_t$ terms , the only ones that contribute are the ones in which $\Gamma(1, u)$ is the differentiated term. We observe that in the matrix ($\mathbf{S}$), only 1 term satisfies this requirement which is $(\mathbf{S})_{\mathbf{1},\mathbf{1}}$. Rest will have a $k \geq 1$. Now the total number of contributing terms will $(n_t - 1)!$. We claim

$$det\mathbf{S}(\mathbf{u}) = \begin{bmatrix} \Gamma(1,u) & \Gamma(2,u) & \cdots & \Gamma(n_t,u) \\ \Gamma(2,u) & \Gamma(3,u) & \cdots & \\ . & & & \\ . & & & \\ . & & & \\ \Gamma(n_t,u) & & \cdots & \Gamma(n_t+n_t-1,u) \end{bmatrix}.$$

This is because if we fix $k = 0$ term in the determinant expansion, there will be a matrix of $n_t - 1$ x $n_t - 1$, which contributes $(n_t - 1)!$ factorial terms and when differentiating, the only term that contributes is the term in which

$\Gamma(1, u)$ is differentiated, i.e., one term from every one of the $(n_t - 1)!$ terms in the determinant expansion (the rest of the $n_t - 1$ terms do not contribute because the differentiated Gamma function will have a $k \geq 1$).

Let us concentrate on one of the $(n_t - 1)!$ terms. This term will be equal to $e^{-u} \prod_{l=1}^{n_t} \Gamma(k_l + 1, u)$ which equals

$$e^{-u} \prod_{l=1}^{n_t - 1} k_l! \left[1 - e^{-u} \sum_{m=0}^{k_l} \frac{u^m}{m!} \right]$$

where k_ls are the coefficients of the chosen Γ functions in the smaller $(n_t - 1)!$ matrix. On expanding this expression we find it is sufficient to consider

$$e^{-u} \prod_{l=1}^{n_t - 1} k_l! \left[1 - e^{-u} \right]$$

since the rest of the terms have a u term in them and hence will not contribute to $c_{i,0}$. Further, this expression can be written as

$$e^{-u}[1 - e^{-u}]^{n_t - 1} r$$

where r is some constant. We observe that on expanding this term , we get only $c_{i,0}$ and there are no u terms. Thus we have succeded in extracting the $c_{i,0}$ terms from one of the $(n_t - 1)!$ terms. The rest of the terms have similar expansions, therefore the summation of all contributing terms will be

$$\sum_{i=1}^{n_t} c_{i,0} e^{-iu} = e^{-u}[1 - e^{-u}]^{n_t - 1} \left(\sum_{l=1}^{(nt-1)!} r_l \right) \tag{50}$$

where r_l are the respective constants of the $(n_t - 1)!$ terms. Now, the sum of all $c_{i,0}$s is obtained by substituting $u = 0$ in (50), and thus

$$\sum_{i=1}^{n_t} c_{i,0} \quad = \quad 0.$$

From (24) of [3], we now obtain

$$\sum_{i=1}^{n_t} i d_{i,0} \quad = \quad 0.$$

References

[1] S. S. Ahuja and V. Sharma, "Optimal power control for convolutional and turbo codes over fading channels," *Proc. of IEEE Conf. Globecom '02,*

[2] H. Alzer, "On some inequalities for the incomplete gamma function," *Mathematics of Computation,* AMS, Vol. 66, pp. 771–778, http://www.ams.org/mcom/1997-66-218/.

[3] P. A. Dighe, R. K. Mallik and S. S. Jamuar, "Analysis of Transmit–Receive Diversity in Rayleigh Fading," *IEEE Trans. on Communications*, Vol. 51, pp. 694–703, April 2003.

[4] A. J. Goldsmith and P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. on Inform. Theory*, Vol. 43, No. 6, pp. 1986-1992, Nov. 1997.

[5] Gradshteyn and Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, 2000.

[6] J. F. Hayes, “Adaptive feedback communication,” *IEEE Trans. on Commun. Technol.,* Vol. Com-16, pp. 29-34, Feb. 1968.

[7] G. Jongren, M. Skoglund, and B. Ottersten, “Combining Beamforming and Orthogonal Space-Time Block Coding, ” *IEEE Trans. on Inform. Theory,”* Vol. 48, pp. 611–627, March 2002

[8] E. G. Larrson and P. Stoica, *Space-Time Block Coding for Wireless Communications*, Cambridge University Press, 2003.

[9] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, 2003.

[10] J. G. Proakis, “Digital Communications,” McGraw-Hill Edition, 2001.

[11] A. Rangarajan, V. Sharma and S. K. Singh, “Information-theoretic and Communication-theoretic Optimal Power Allocation for Fading Channels,” *Proc. of the IEEE Intl. Symposium on Information Theory (ISIT)*, Yokohama, Japan, p. 246, 2003.

[12] A. Rangarajan, and V. Sharma, “Achieving Exponential Diversity Order in Rayleigh Fading Channels,” *Proc. National Conf. on Commun. (NCC),* Bangalore, Jan. 2004.

[13] Simon, M. K. and Alouini, M. S., “Digital Communications over Fading Channels: a unified approach to performance analysis,” John Wiley and Sons, New York, NY.