

Return Address:

Dr. Raymond KNOPP
Mobile Communications Laboratory
Swiss Federal Institute of Technology (Lausanne)
INR Ecublens
1015 Lausanne, SWITZERLAND

Power Control and Beamforming for Systems with Multiple Transmit and Receive Antennas

Raymond Knopp, Giuseppe Caire¹

Technical Subject Area: Communication Theory, Radio Communication, Smart Antennas and Interference
Cancellation

Tel: (+41) 21 693 5657

Fax: (+41) 21 693 4312

Email: raymond.knopp@epfl.ch

¹G. Caire is with *Institut Eurécom*, B.P. 193, 06904 Sophia Antipolis, France

Abstract

This paper investigates the performance of narrowband, slowly fading, and delay-limited multiple-antenna systems where *channel state information* (CSI) is available at the transmission end. This situation can arise in *time-division duplex* (TDD) based two-way systems where channel state estimation can be performed using the signal received from the opposite link. Power control methods which attempt to keep the transmission rate constant at the expense of randomizing the transmit power are considered. It is shown that significant savings in average transmit power (sometimes on the order of tens of dB) can be expected compared to systems which keep the total transmit power constant. Several practical channel coding examples using are illustrated and their bit and frame error rate performance are discussed.

1 Introduction

The aim of this paper is to investigate the performance improvements which can be gained by employing *channel state information* (CSI or side information) in the form of *power control* at the transmission end of a multi-antenna radio communication system. The first question that one may ask is how realistic is it to assume that quasi-perfect CSI can be made available at the transmission end. The answer depends strongly on the system architecture. Recent proposals for next generation global wireless communications systems include both time division (TDD) and frequency division (FDD) duplex multiplexing strategies (e.g. [1]). If we employ the same antenna array for transmission and reception in a TDD system then channel reciprocity allows us to use our channel estimates obtained during reception for transmission. We will show that the performance improvements over modern antenna processing schemes are dramatic when CSI is exploited. This provides a strong argument for employing TDD, when possible, in next generation wireless systems.

Despite the reciprocity of the propagation channel, in systems with multiple transmit antennas, there are non-trivial implementation issues which make the problem harder, namely the effects of transmission and reception electronics. More precisely, the channel responses are cascades of the physical channel and the responses of the transmission and reception electronics. Since the transmitter and receiver electronics do not have the same response, they must be estimated using self-calibrating circuits. This is less of an issue in single-antenna systems, since estimates of the cascaded channels phase responses are not required. With multiple-antennas accurate phase estimates are required to use techniques such as *beamforming*. Some important practical issues are discussed in [2]. In this work we assume that the cascaded channel amplitude and phase responses can be estimated quasi-perfectly so that the signal at the receiver can be controlled.

This paper approaches the problem more from the point of view of communication theory rather than signal processing. Recent papers on antenna processing using a similar approach include [3, 4, 5, 6, 7]. We will make use of some key results from these studies in order to analyse the benefits of employing CSI at the transmitter.

In this work we focus on delay-sensitive systems and/or non-ergodic channels. For such cases (see e.g. [8, 9, 10, 11, 12, 13]) the instantaneous average mutual information between the transmitted and received signals depends on the channel state and is therefore random. In practice, a situation such as this will arise when the number of antennas is small and little or no time/frequency variation occurs in the duration/band of the transmitted signal. Mathematically speaking, the number of significant degrees of freedom characterizing the randomness of the channel process is small enough that ergodic arguments cannot be invoked in analyzing system performance. To counter the randomness of the channel, generalized power control can be employed to maintain a constant average mutual information between transmitter and receiver, at the expense of rendering the transmit power random. The direct consequence of this will be that for certain channel realizations the expended transmit power will be higher than for others. Our figure of merit, therefore, is the average transmit power to maintain a constant average mutual information (or maximum information rate for which reliable communication is possible) between the transmitter and receiver.

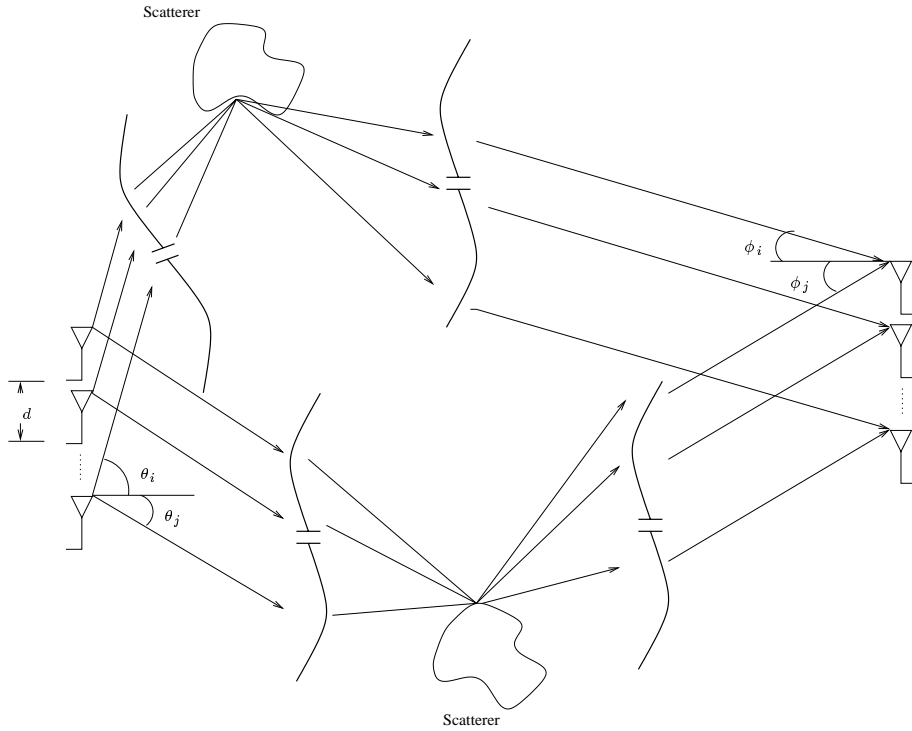


Figure 1: Transmit and Receive Antenna System

The paper is organized as follows: Section 2 presents the system model under consideration, namely a slowly-fading multipath channel with multiple transmit and receive antennas. We consider two multi-antenna architectures, linear antenna arrays and dual (horizontal/vertical) polarized antennas. A discrete-time model based on the *singular value decomposition* is reviewed. Section 3 examines the important special case where only a single antenna is present on one end. We show that power control can yield reductions in transmit power greater than 10dB compared to schemes such as *space-time coding*[3]. The effect of element spacing is also discussed. In section 4 we consider the more general case where multiple-antennas are present on both ends. Again, impressive power savings can be expected with optimal and sub-optimal power control methods. Finally, in Section 5 we address the issue of outdated channel estimates and their detrimental effect on the performance of the proposed system.

2 Multiantenna Transmission/Reception

Consider multi-antenna diversity signaling scheme shown in Figure 1 which uses M antennas in the transmitter and N antennas in the receiver. The same antenna is used for transmission and reception at the same carrier frequency so that channel reciprocity holds. We assume that the M transmitted signals, $s_i(t), i = 1, \dots, M$ are narrowband QAM signals of the form

$$\tilde{s}_i(t) = \sum_k \sqrt{P_i} u_{i,k} g(t - kT), \quad i = 1, 2, \dots, M \quad (1)$$

where $u_{i,k}$ is the k^{th} complex symbol on the i^{th} transmit antenna and $g(t)$ is a signaling pulse such that its Fourier transform, $G(f)$ is zero for frequencies $f > W/2$, where W is the bandwidth of the transmitted signal. We assume that $g(t)*g^H(-t)$ satisfies Nyquist's criterion for zero intersymbol interference and that the average energy of the $u_{i,k}$ is $\mathbb{E} |u_{i,k}|^2 = 1$. The \mathcal{P}_i are the instantaneous powers allocated to each transmit antenna element. The real transmitted signals are

$$s_i(t) = \text{Re}(\tilde{s}_i(t)e^{j2\pi f_c t}), \quad i = 1, 2, \dots, M \quad (2)$$

where f_c is the carrier frequency.

The $s_i(t)$ are transmitted over a static L -path multipath channel with NM impulse responses

$$h_{ij}(t) = \sum_{l=1}^L a_l \delta(t - d_l(i, j)) \quad (3)$$

where a_l and $d_l(i, j)$ are the gain and delay of the l^{th} path between transmitter i and receiver j . The channel remains static during the transmission of long codewords but can change from codeword to codeword. We assume that $W(d_l(i, j) - d_1(i, j)) \ll 1$ so that we may approximate the complex baseband equivalent signal seen by the j^{th} receiver by

$$\tilde{r}_j(t) = \sum_{i=1}^M \sum_k \sqrt{\mathcal{P}_i} A_{i,j} u_{i,k} g(t - kT) + \tilde{z}_j(t), \quad j = 1, 2, \dots, N \quad (4)$$

where A_{ij} is the complex gain of the i^{th} transmitter and j^{th} receiver given by

$$A_{ij} = \sum_{l=1}^L a_l e^{j2\pi f_c d_l(i,j)}. \quad (5)$$

The $\tilde{z}_j(t)$ are circularly symmetric additive white (in the band of the signal) Gaussian noise with power spectral density N_0 .

As is common in the literature we take the A_{ij} to be complex Gaussian random variables with variance σ_A^2 and mean μ_{ij} . Note that not all the A_{ij} have the same mean since they differ in phase but that their variances are the same. For convenience, the average attenuation of the channel $\sum_l \mathbb{E} a_l^2$ is assumed to be normalized to one so that

$$\sigma_A^2 + |\mu_{ij}|^2 = 1 \quad (6)$$

A non-zero μ_{ij} indicates the presence of a direct path between the transmitter and receiver and the power ratio $K = |\mu_{ij}|^2 / \sigma_A^2$ is commonly referred to as the *Ricean Factor*. Assuming that the receiver employs maximum-likelihood detection using filters $g^H(-t)$ sampled at instants $t = kT$ we have the discrete-time channel model

$$r_{k,j} = \sum_{i=1}^M \sqrt{\mathcal{P}_i} A_{i,j} u_{k,i} + z_{k,j}, \quad j = 1, \dots, M \quad (7)$$

or in vector form

$$\mathbf{r}_k = \mathbf{A} \text{diag}(\mathcal{P}_i) \mathbf{u}_k + \mathbf{z}_k \quad (8)$$

where \mathbf{A} is an $N \times M$ matrix of complex channel gains.

2.1 Receiver and Transmitter Channel State Information

The A_{ij} are assumed to be known perfectly to the receiver. This can be achieved by inserting training sequences (possibly a different one for each transmit antenna!) which allow for quasi-perfect estimation of the A_j and at the same time do not significantly reduce information rates. As mentioned in the introduction, the results of this work are intended to provide a strong argument for employing time-division duplex with the same antennas used for transmission and reception. We will therefore also assume that the A_{ij} may be known to the transmitter. The channel state information will be estimated from the signal received during previous time-slots.

There will be two sources of error in the channel gain estimates. The first will be due to additive noise and affects both the transmitter and receiver CSI in an identical fashion. It can be made insignificant by using fairly long training sequences. In addition, assuming correct decisions in the received signal the transmission CSI can be improved using decision feedback estimation. The second source of error affects only the transmitter CSI and is due to imperfect estimates due to channel variation. Since the estimate is based on a signal which was received at an earlier point in time, the channel may have changed across time slots. This will be addressed in Section 5.

2.2 Linear Arrays and Dual Polarized Antennas

If we consider linear antennas with element spacing d in both the transmitter and receiver as shown in Figure 1, the wavefront associated with each scatterer will have *two* associated angles, θ (transmission end) and ϕ (reception end). Using the standard 2-dimensional land-mobile model [14] we assume the scatterers are uniformly distributed around both the transmitter and receiver. Moreover, we may take the θ_i and ϕ_i to be independent and uniformly distributed on $[0, 2\pi)$. We may also approximate the path difference between trajectory i, j (transmitter i to receiver j) and i', j' (transmitter i' to receiver j') for each scatterer as

$$d_{i,j,i',j'} = (i - i')d \cos \theta + (j - j')d \cos \phi \quad (9)$$

Generalizing the standard result [14] we have that the A_{ij} have cross-correlation

$$\begin{aligned} K_{i,j,i',j'} &= \mathbb{E} (A_{ij} - \overline{A_{ij}})(A_{i'j'}^H - \overline{A_{i'j'}^H}) \\ &= \mathbb{E} e^{j2\pi \frac{f_c d}{c}(i-i') \sin \theta} e^{j2\pi \frac{f_c d}{c}(j-j') \sin \phi} \\ &= J_0 \left(2\pi \frac{f_c d}{c} |i - i'| \right) J_0 \left(2\pi \frac{f_c d}{c} |j - j'| \right) \end{aligned} \quad (10)$$

where $c = 3 \times 10^8$ m/s is the speed of light.

For dual-polarized antennas in either or both the transmitter and receiver, the matrix \mathbf{A} will contain correlated components with varying average powers. In the basestation antenna, because of the dominant horizontal component in the propagation path, the vertical polarization will be strongest, typically by a factor 10 [15]. This may be reduced if a dual-polarized antenna is used in the handset. Experiments have shown further that

the two components can have a normalized correlation coefficient around .2 [16]. For these types of systems we will consider only single-element antennas with both polarizations in either the transmitter or receiver but not both. The gain in the horizontal component will be denoted λ_h and in the vertical component λ_v , and the correlation between the two, ρ_{vh} .

2.3 Parallel Channel Decomposition

As in [5, 4] which generalize the continuous-time frequency-selective channel described in [17, Chap. 8] to a discrete-time multi-antenna system. We decompose \mathbf{A} using its *singular value decomposition* $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ where \mathbf{U} and \mathbf{V} are $N \times N$ and $M \times M$ unitary matrices and

$$\mathbf{\Sigma} = \begin{cases} \begin{bmatrix} \sqrt{\mathbf{\Lambda}} & \mathbf{0} \end{bmatrix} & M \geq N \\ \begin{bmatrix} \sqrt{\mathbf{\Lambda}} \\ \mathbf{0} \end{bmatrix} & M < N \end{cases}$$

and $\mathbf{\Lambda}$ is a $\min(N, M)$ -dimensional diagonal matrix containing the non-identically zero eigenvalues of $\mathbf{A}\mathbf{A}^H$ or $\mathbf{A}^H\mathbf{A}$. Since \mathbf{A} is known to the transmitter, $\mathbf{U}, \mathbf{\Sigma}$ and \mathbf{V} can, at least in principle, be computed before transmission or reception. We may therefore transform the detection problem, without loss of generality, as

$$\mathbf{r}'_k = \mathbf{U}^H \mathbf{r}_k = \mathbf{\Sigma} \text{diag}(\mathcal{P}'_i(\mathbf{\Lambda})) \mathbf{u}'_k + \mathbf{z}'_k \quad (11)$$

where $\text{diag}(\mathcal{P}'_i) \mathbf{u}'_k = \mathbf{V}^H \text{diag}(\mathcal{P}_i) \mathbf{u}_k$ and \mathbf{u}'_k has unit variance components. The \mathcal{P}'_i are the power of the transmitted signal components in the transform domain. Note that either the transformed input or transformed output may be reduced in dimension if \mathbf{A} is not square. We have the equivalent parallel channel representation

$$r'_{i,k} = \sqrt{\lambda_i} \mathcal{P}'_i(\mathbf{\Lambda}) u'_{i,k} + z'_{i,k}, \quad i = 1, \dots, \min(M, N) \quad (12)$$

which is conceptually the same as frequency-based diversity system (see e.g. [18, 10, 11, 19], except in the transform domain. The main difference between the two systems is that the statistics of $\mathbf{\Lambda}$ are more complex than those of the channel strengths at different frequencies, since they are eigenvalues of a (typically) Gaussian matrix. The matrix $\mathbf{A}\mathbf{A}^H$ is known as a *Wishart matrix* [20] when the A_{ij} are Gaussian, and is the matrix generalization of chi-square random variables. Except in the special case of i.i.d. Gaussian distributed components in \mathbf{A} [21, 5], the distribution of the λ_i are unknown.

Here we allow the instantaneous transmit power

$$P = \sum_{i=0}^{\min(M,N)} |u_{i,k}|^2 = \sum_{i=0}^{\min(M,N)} \mathcal{P}'_i \quad (13)$$

to vary with the channel subject to the power constraint

$$\sum_{i=0}^{\min(M,N)} \mathbb{E} \mathcal{P}'_i \leq \bar{P} \quad (14)$$

where \bar{P} is the average transmit power. We will express or compare the performance of systems in terms of the average transmit power to received noise power ratio \bar{P}/WN_0 where W is the total bandwidth used by the signaling scheme and N_0 is the two-sided noise spectral density. Additionally we also characterize systems in terms of the average bit signal energy to noise ratio E_b/N_0 , which may cause some confusion. To see this first suppose we keep both the average transmit power, P , and transmit bandwidth W constant and use a coded-modulation scheme of rate R bits/symbol transmitting with M antennas. The overall information rate is MR which increases linearly with the number of antennas, despite the constant bandwidth. Suppose now we compare systems with constant bit rate R_b and constant transmit power $\bar{P} = R_b E_b$, where E_b is the energy per bit. The energy per transmit symbol with, say, uniform power allocation is $MR E_b/M = R E_b$ which is a constant independent of M . This implies that by increasing M we have both an increase in performance (constant symbol energy) and a reduction in bandwidth. Note that this difference is not apparent with multi-antenna receivers.

3 $N \times 1$ and $1 \times M$ Antenna Diversity

Let us first consider the simplest case which would apply to mobile systems which employ only a single antenna in the mobile terminal. As an example, let us consider the downlink channel with M basestation antennas and power control. In the parallel channel decomposition there is a single eigenvalue $\lambda_1 = \sum_{i=1}^M |A_{1,i}|^2$ and the transformed input is one-dimensional. This means that the optimal coding scheme is repetition coding on each antenna with gain $A_{1,i}^H/\sqrt{\lambda_1}$. This amounts to performing *maximal-ratio combining* in the transmitter which can also be found by choosing the antennas gains/phases to maximize the signal-to-noise ratio at the receiver [22]. It can also be interpreted as *beamforming*. We will see that impressive gains can be obtained by using this technique instead of space-time coding. Similar techniques can be used to reduce interference in multiuser systems ([23, 24])

In this simple case without multipath, we see that $\lambda_1 = M$ so that the beamforming antenna essentially provides a gain of M *at the receiver* without increasing the transmit power. We do not violate the law of conservation of energy since we just focus the energy at the receiver. There will be other points on the sphere of radius equal to the distance between the receiver and transmitter which will contain nulls.

Assuming we transmit long codewords and the channel is static during the codeword, the instantaneous average mutual information (channel capacity) as a function of λ_1 and $\mathcal{P}(\lambda_1)$ is given by

$$C(\mathcal{P}(\lambda_1), \lambda_1) = W \log_2 \left(1 + \frac{\mathcal{P}(\lambda_1)\lambda_1}{WN_0} \right) \text{ bits/s} \quad (15)$$

assuming an average symbol energy of one and minimum bandwidth pulse shapes $g(t)$. It is achieved if the $w'_{1,k}$ are zero-mean, independent, circular symmetric complex Gaussian random variables [17, Chap. 7]. The meaning of this quantity is that reliable communication is impossible if $R > C(\mathcal{P}(\lambda_1), \lambda_1)$, or that R must be adjusted as a function of $\mathcal{P}(\lambda_1)\lambda_1$. If we wish to maintain a constant channel capacity (or practically transmit

at a constant rate) then the power controller must satisfy

$$\mathcal{P}(\lambda_1) = \frac{WN_0(2^{W^{-1}R} - 1)}{\lambda_1} \quad (16)$$

and the average transmitted power (taken over all realizations of λ_1) is

$$\bar{P} = WN_0(2^{W^{-1}R} - 1) \int_0^\infty \frac{f_{\Lambda_1}(u)}{u} du \quad (17)$$

In single-antenna systems with narrowband signals, this quantity is infinite [25]. We will see that with multiple transmit and/or receive antennas it can be finite. In fact, we will see that by using power control we can actually achieve a significant transmitter power savings (sometimes by a factor 10 or more) with respect to using a constant transmit power. This reduction comes at the expense of an increased *peak-to-average* power ratio. Current and next generation wireless systems can have transmit power amplifiers with gains which can vary over as much as 90dB. With power control we will simply operate well below maximum most of the time, with rare spurts of more powerful transmissions. This power savings translates directly either to reduced battery size or increased battery life for mobile terminals, which is of course of great importance.

3.1 P.D.F. of λ_1 Rayleigh/Ricean Fading and Linear/Polarized Arrays

The single channel coefficient $\lambda_1 = \sum_{i=1}^M |A_{1,i}|^2$ in Rayleigh/Ricean fading is a trivial quadratic form of correlated complex Gaussian random variables. The covariance matrix \mathbf{K} has components given by (10) for a linear antenna array and by

$$\mathbf{K} = \begin{pmatrix} \epsilon_v & \rho_{vh} \\ \rho_{vh} & \epsilon_h \end{pmatrix} \quad (18)$$

where $\epsilon_v = E\lambda_v$ and $\epsilon_h = E\lambda_h$ for a dual-polarized antenna.

From [26] we have that the moment generating function of λ_1 is given by

$$G_{\Lambda_1}(s) = \frac{e^{s\mathbf{m}^H(\mathbf{I}-s\mathbf{K})^{-1}\mathbf{m}}}{\det(\mathbf{I}-s\mathbf{K})} \quad (19)$$

where \mathbf{m} is the mean of the each antenna element gain (i.e. the complex gain of the direct path). For the linear array this will be given by

$$\mathbf{m} = |\mu_{ij}| e^{j\phi} \left(1 \quad e^{j\frac{f_c d \cos(\psi)}{c}} \quad e^{j\frac{2f_c d \cos(\psi)}{c}} \quad \dots \quad e^{j\frac{(M-1)f_c d \cos(\psi)}{c}} \right)^T \quad (20)$$

where ψ is either the angle of departure (transmitter array) or arrival (receiver array) of the direct path with respect to the array. Note that the performance is dependent on ψ which is random, even though it can be estimated. The phase ϕ is the random phase of the first element.

For non-zero covariance between elements in a no line-of-sight situation, the eigenvalues $\{\eta_i\}$ of \mathbf{K} will be distinct and we may easily invert $G_{\lambda_1}(s)$ yielding the p.d.f.

$$f_{\Lambda_1}(u) = \sum_{i=1}^M B_i e^{-u/\eta_i} \quad (21)$$

where $B_i = \eta_i^{N-1} \left(\prod_{j=1, \dots, n, j \neq i} (\eta_i - \eta_j) \right)^{-1}$. With a non-zero line-of-sight component, the initial value theorem applied to $G_{\Lambda_1}(s)$ can be invoked to approximate the p.d.f. for small values of λ_1 which dominate the average transmit power [22] as

$$f_{\Lambda_1}(u) \approx \frac{u^{M-1} e^{-u}}{\Gamma(M)} \left(\prod_{i=1}^M \eta_i \right)^{-1} e^{-.5 \mathbf{m}^H \mathbf{K}^{-1} \mathbf{m}} \quad (22)$$

so that it behaves as a central chi-square random variable around the origin with the geometric mean of the η as parameter. When the $\{A_{1,i}\}$ are uncorrelated and unit mean-square we have the p.d.f.

$$f_{\Lambda_1}(u) = (1+K) \left((1+1/K) \frac{u}{M} \right)^{M-\frac{1}{2}} \exp(-(1+K)u + MK) I_{2M-1} \left(\sqrt{uMK(K+1)} \right), u \geq 0 \quad (23)$$

Using these p.d.f. we may easily evaluate (17). For the case of uncorrelated zero-mean components we see immediately that

$$\frac{\bar{P}}{WN_0} = \frac{1}{M-1} \quad (24)$$

For $M = 2, 3$ and correlated zero-mean components we have that

$$\frac{\bar{P}}{WN_0} = \begin{cases} \frac{\ln \eta_1 / \eta_2}{\eta_1 - \eta_2}, & \eta_1 > \eta_2, M = 2 \\ \frac{\eta_1 \eta_2 \ln \eta_1 / \eta_2 - \eta_1 \eta_3 \ln \eta_1 / \eta_3 + \eta_2 \eta_3 \ln \eta_2 / \eta_3}{(\eta_1 - \eta_2)(\eta_1 - \eta_3)(\eta_2 - \eta_3)} & \eta_1 > \eta_2 > \eta_3, M = 3 \end{cases} \quad (25)$$

For non-zero mean uncorrelated components, we may use [27, 6.643] yielding

$$\frac{\bar{P}}{WN_0} = \frac{(1+K)^M}{2M-1} \left[\frac{1+1/K}{M} \right]^{M-1/2} \frac{2e^{9MK/8}}{\sqrt{MK(K+1)}} M_{1-M, M-1/2} \left(\frac{MK}{4} \right) \quad (26)$$

where $M_{\mu, \nu}(x)$ is the *Whittaker function*. It is simpler to use numerical means to compute the average transmit power. For the general Ricean case we may use the approximation in (22) to arrive at

$$\frac{\bar{P}}{WN_0} \approx \frac{e^{-.5 \mathbf{m}^H \mathbf{K}^{-1} \mathbf{m}}}{M-1} \left(\prod_{i=1}^M \eta_i \right)^{-1/M} \quad (27)$$

which for uncorrelated gains simplifies to

$$\frac{\bar{P}}{WN_0} \approx \frac{e^{-K}}{(M-1)(1+K)} \quad (28)$$

We remark that $1 \times M$ and $N \times 1$ systems are equivalent when power control is used. This is not the case without power control [5]. As we mentioned in the previous section, this is because power control allows us to perform maximal-ratio combining/beamforming in the transmitter. In the special (worst) case of Rayleigh fading, the loss in transmit SNR with respect to a non-fading channel with multiple antennas is $1 - 1/M$, which very quickly becomes negligible. Thus with multiple antennas and power control, we can eliminate the effect of multipath fading and still benefit from large gains due to beamforming. Moreover, the presence of a direct path will also not significantly influence the average transmit power.

In figure 2 we plot the necessary transmit SNR to achieve 2 bits/s/Hz for a linear array as a function of the number of antennas and the element spacing in Rayleigh fading. We see that for spacings greater than a quarter wavelength we obtain close to asymptotic performance. For a 2GHz carrier frequency this is on the order of 3.75cm which may be reasonable even for some mobile terminals with micro-strip antennas. Figure 3 shows the performance of a two-element (e.g. dual-polarized) antenna as a function of the normalized cross-correlation ρ_{vh} and the gain ratio $\frac{\bar{\lambda}_v}{\bar{\lambda}_h}$ assuming $\bar{\lambda}_v = 1$.

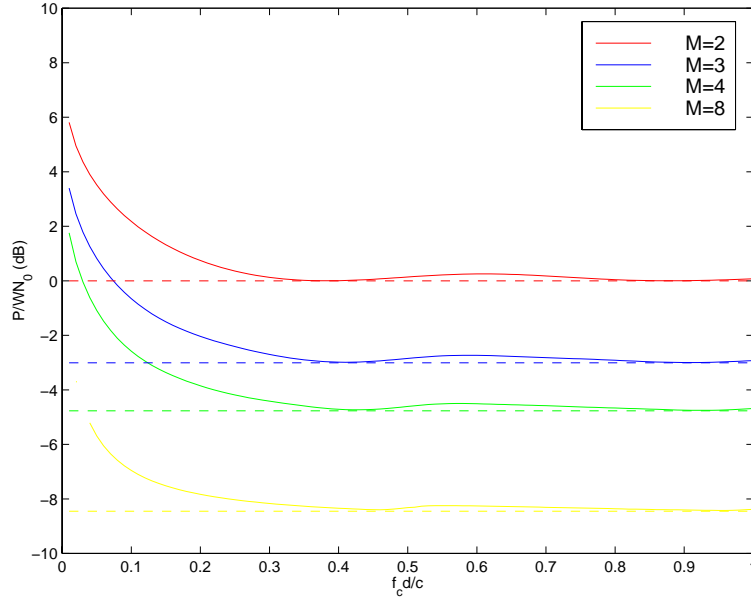


Figure 2: SNR reduction as a function of antenna spacing d

3.1.1 Comparison with Multiantenna Channel Coding Schemes

In the case where equal power \bar{P}/M is assigned to each transmit antenna, the information outage probability is [5, 6]

$$P_{\text{out}}(R) = \text{Prob} \left(\log \det \left(\mathbf{I} + \frac{\bar{P}}{MWN_0} \mathbf{A}^H \mathbf{A} \right) < W^{-1}R \right) \quad (29)$$

$$= \text{Prob} \left(\sum_{i=1}^{\min(M,N)} \log_2 \left(1 + \frac{\bar{P}}{MWN_0} \lambda_i \right) < W^{-1}R \right) \quad (30)$$

This outage rate indicates the practical *frame error rate* (FER) performance of multiantenna coding schemes which do not exploit channel state information (e.g. space-time coding). For a $1 \times M$ system this simplifies to

$$P_{\text{out}}(R) = F_{\Lambda_1} \left[\left(2^{W^{-1}R} - 1 \right) \frac{MWN_0}{\bar{P}} \right] \quad (31)$$

where $F_{\Lambda_1}(\cdot)$ is the cumulative distribution function (c.d.f.) of Λ_1 . In Ricean fading this quantity is given by

$$P_{\text{out}}(R) = 1 - Q_M \left(\sqrt{K}, \sqrt{(K+1)\beta} \right) \quad (32)$$

where $\beta = \frac{MWN_0}{\bar{P}} (2^{W^{-1}R} - 1)$ and $Q_M(\cdot)$ is the *Marcum Q function of order M*. This is plotted in figure 4 for Rayleigh fading. The effect of beamforming is evident since we obtain gains on the order of 12dB compared to

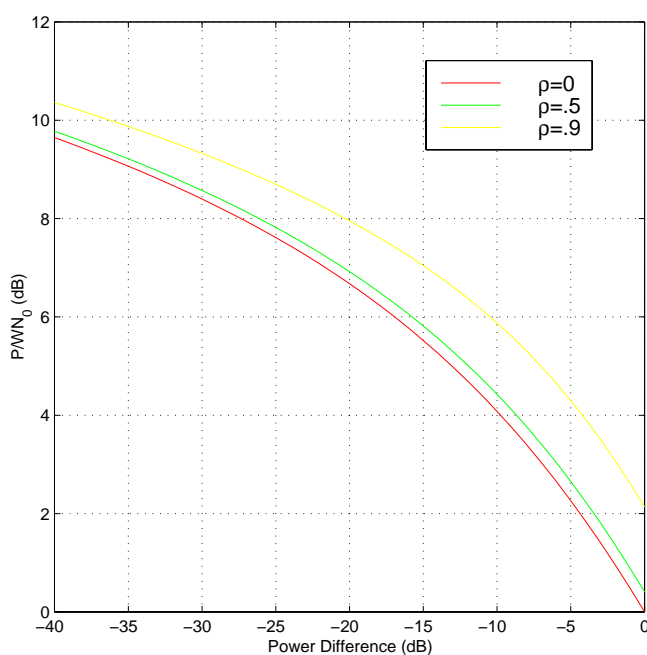


Figure 3: SNR reduction in a dual-polarized antenna systems as a function of normalized cross-correlation ρ_h and element gain ratio ϵ_v/ϵ_h

space-time coding for frame error-rates around 10^{-2} . Much more impressive gains can be had at lower FER. We may conclude, therefore, that under the assumption that fairly simple codes *designed for the AWGN channel* can bring us to within a few dB from channel capacity, we may obtain huge reductions in transmit power with respect to an *optimal* multi-antenna coding scheme which does not exploit CSI at the transmitter.

This is demonstrated in Figure 5 where we show the simulated *bit-error rate (BER)* and FER of a dual transmit antenna ($M = 2$) QPSK 4-state space-time code (taken from [3]) with fixed transmit power and uncoded QPSK with power control for the same *average* transmitted $\mathcal{E}_b/N_0 = \bar{P}/RN_0$, where $R = 2$ is now the number of bits per symbol. This was carried out under the assumption of independent Rayleigh fading for each antenna element for which the BER and FER for uncoded QPSK with power control are given by

$$\text{BER} = Q\left(\sqrt{\frac{2\mathcal{E}_b}{(M-1)N_0}}\right)$$

$$\text{FER} = 1 - (1 - \text{BER})^L$$

where L is the number of information bits per frame. For the results of figure 5 we chose $L = 130$. Note that even with uncoded modulation, the effect of power control is a dramatic reduction in average transmit power. Moreover, since power control keeps the *received* SNR constant, any standard coding scheme for AWGN channels can be employed while maintaining the same amount of coding gain.

Although the necessary transmitted signal-to-noise ratio may be small (e.g. 0 dB), at the receiver we obtain an effective gain of $M - 1$. This is important since frequency and timing estimation algorithms cannot operate at very low *received* signal-to-noise ratios.

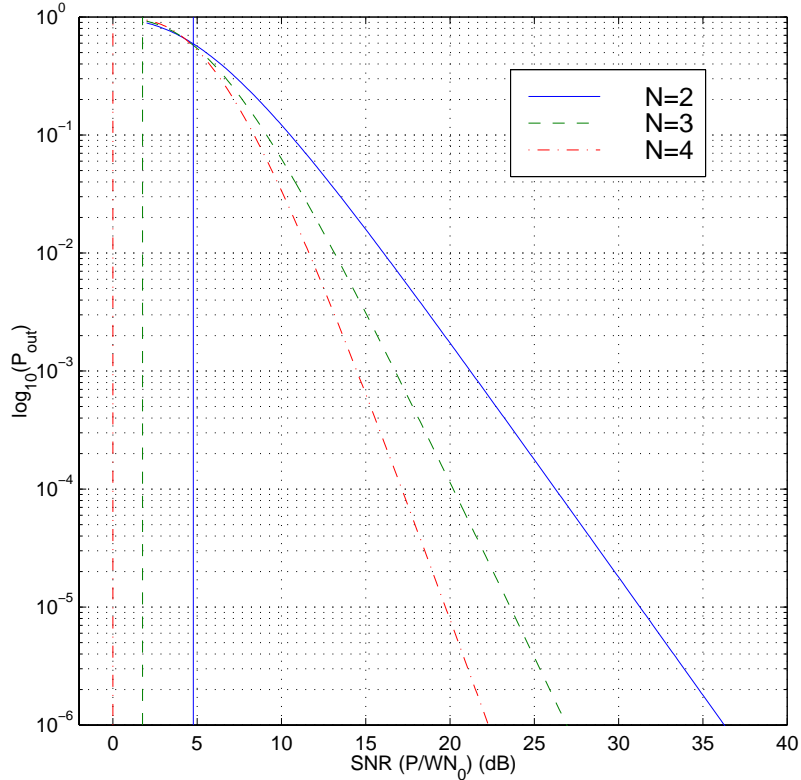


Figure 4: Space-time Codes Comparison

4 Generalized Beamforming

Let us now consider the more general beamforming case where $M, N > 1$. For the set of $\min(M, N)$ parallel Gaussian channels in (12) we have that the instantaneous channel capacity is given by [17, Chap 7]

$$C(\mathcal{P}'_i(\mathbf{\Lambda}), \mathbf{\Lambda}) = \sum_{i=1}^{\min(M,N)} R_i = \sum_{i=1}^{\min(M,N)} \log_2 \left(1 + \frac{\mathcal{P}'_i(\mathbf{\Lambda}) \lambda_i}{W N_0} \right) \quad (33)$$

and is achieved by using input symbols u'_{ik} which are independent Gaussian random variables. Moreover, we are interested in choosing the $\mathcal{P}'_i(\mathbf{\Lambda})$ such that the total power is minimum for each realization of the $\{\lambda_i\}$ subject to the fixed rate constraint $\sum_{i=1}^{\min(M,N)} R_i = R$. Note that this is not exactly the standard water-filling optimization [17, Chap. 7]. Here we assure that we always transmit reliably at a fixed rate while minimizing the long-term average power. This optimization is a special case of the general multiuser framework considered in [9]. In addition it is also an important special case of the single-user parallel channel information outage probability minimization [13]. Specifically we have

$$\min \sum_{i=1}^{\min(M,N)} \mathcal{P}'_i(\mathbf{\Lambda}) = \sum_{i=1}^{\min(M,N)} \frac{2^{R_i} - 1}{\lambda_i} \quad \text{subject to} \quad \sum_{i=1}^{\min(M,N)} R_i = R \quad (34)$$

which is a standard concave optimization problem (see e.g. [17, Chap 4]). The Kuhn-Tucker conditions yield the system of inequalities

$$\frac{\partial \mathcal{P}'_i(\mathbf{\Lambda})}{\partial R_i} = \frac{2^{R_i}}{\lambda_i} \geq B(\mathbf{\Lambda}) \quad (35)$$

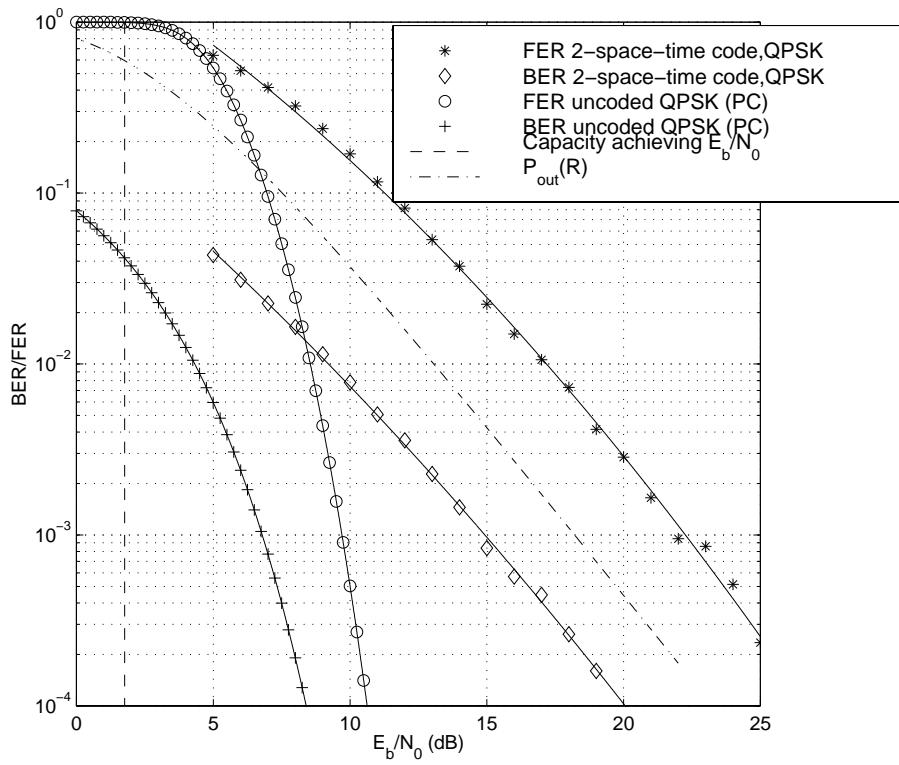


Figure 5: BER/FER Comparison of Power Controlled QPSK and 2-Space-time QPSK Codes (4-state)

with equality if and only if $R_i > 0$. Note that the optimization constant $B(\mathbf{\Lambda})$ is dependent on the realization of $\mathbf{\Lambda}$. This says that the effective rate on each sub-channel should be adjusted according to the *strengths of all channels* and the threshold $B(\mathbf{\Lambda})$ is chosen to satisfy the fixed rate constraint. This is very reminiscent of an OFDM system where responses at the different carrier frequencies are now replaced by singular values of a transition matrix. With wideband signals, this idea can be extended further by defining a set of singular values for each frequency carrier [4].

In many instances it would be unlikely (although not impossible) to employ more than 2 antennas in the mobile terminal, both for size and receiver complexity reasons. In wideband wireless local loop systems, with fixed terminals, it could, however, be very beneficial to examine many antennas both in the transmitter and receiver [4]. Here we restrict ourselves to the 2 channel case so that $\min(M, N) = 2$. The corresponding power controllers are

$$\begin{cases} \mathcal{P}'_1(\mathbf{\Lambda}) = W N_0 \frac{2^R - 1}{\lambda_1}, \mathcal{P}'_2(\mathbf{\Lambda}) = 0 & \lambda_1 \geq 2^R \lambda_2 \\ \mathcal{P}'_1(\mathbf{\Lambda}) = 0, \mathcal{P}'_2(\mathbf{\Lambda}) = W N_0 \frac{2^R - 1}{\lambda_1} & \lambda_2 \geq 2^R \lambda_1 \\ \mathcal{P}'_1(\mathbf{\Lambda}) = W N_0 \frac{2^{R/2}}{\sqrt{\lambda_1 \lambda_2}} - \frac{1}{\lambda_1}, \mathcal{P}'_2(\mathbf{\Lambda}) = \frac{2^{R/2}}{\sqrt{\lambda_1 \lambda_2}} - \frac{1}{\lambda_2} & \text{otherwise} \end{cases}$$

In the optimal power control scheme, *Scheme 1*, we see that if one channel (eigenvector) is much stronger (depending on the rate constraint) than the other, only the stronger one is used. Otherwise, both are used. Note that in the latter case, the received SNR on each channel is not constant, but depends on the relative strengths of the two channels. Unlike the single channel case, this implies that standard AWGN codes need not be effective

for this type of system, and special codes must be designed.

A simple sub-optimal modification of the optimal scheme would be to use a rate R code on the stronger channel when one channel is K times stronger than the other and a rate $R/2$ code on each channel otherwise. K is a parameter to be determined. When $K = 1$ we simply select the best channel (selection diversity). In this power control scheme, we could keep the received SNR constant and use AWGN codes with predictable performance. Let us denote this scheme by *Scheme 2*. The power controllers are

$$\begin{cases} \mathcal{P}'_1(\mathbf{\Lambda}) = WN_0 \frac{2^R - 1}{\lambda_1}, \mathcal{P}'_2(\mathbf{\Lambda}) = 0 & \lambda_1 \geq K\lambda_2 \\ \mathcal{P}'_1(\mathbf{\Lambda}) = 0, \mathcal{P}'_2(\mathbf{\Lambda}) = WN_0 \frac{2^R - 1}{\lambda_2}, & \lambda_2 \geq K\lambda_1 \\ \mathcal{P}'_1(\mathbf{\Lambda}) = WN_0 \frac{2^{R/2} - 1}{\lambda_1}, \mathcal{P}'_2(\mathbf{\Lambda}) = WN_0 \frac{2^{R/2} - 1}{\lambda_2} & \text{otherwise} \end{cases}$$

Another even simpler sub-optimal signaling scheme, *Scheme 3*, which does not require an eigenvalue decomposition is as follows. On the end with $U = \max(M, N)$ antennas, we perform either beamforming (transmission) or maximal ratio combining (reception) to obtain the largest gain. On the end with $L = \min(M, N)$ antennas we select the antenna yielding the strongest received signal-to-noise ratio. In this case the single channel gain is $\alpha = \max_i \|\mathbf{a}_i\|^2$, $i = 1, \dots, L$ and \mathbf{a}_i is the i^{th} U -dimensional row or column of \mathbf{A} .

In order to determine the average transmit power needed to communicate reliably at rate R for the three signaling schemes, we must determine $f_{\lambda_1, \lambda_2}(u, v)$. For i.i.d. Gaussian components in \mathbf{A} the distribution of the ordered eigenvalues is derived in closed-form [21, 5] and similarly that of the unordered eigenvalues in [5]. The ordered p.d.f. is given by

$$f_{\lambda_{\max}, \lambda_{\min}}(u, v) = K_2(D) e^{-u-v} u^D v^D (u-v)^2 \quad (36)$$

where $K_2(D)$ is a normalizing constant given by (using [27, 3.381, 6.455])

$$\begin{aligned} K_2(D) = & \left[\frac{\Gamma(D+4)}{(D+1)2^{2D+4}} {}_2F_1(1, 2D+4; D+2; 1/2) \right. \\ & - \frac{\Gamma(D+4)}{(D+2)2^{2D+3}} {}_2F_1(1, 2D+4; D+3; 1/2) \\ & \left. + \frac{\Gamma(D+4)}{(D+3)2^{2D+4}} {}_2F_1(1, 2D+4; D+4; 1/2) \right]^{-1}, \quad (37) \end{aligned}$$

$D = |M - N|$ is the absolute difference between the number of transmit and receive antennas, and ${}_2F_1(a, b; c; d)$ is an hyper-geometric number. We may write the average powers for Schemes 1 and 2 as

$$\begin{aligned} \overline{P}_1 &= WN_0 \int_0^\infty \left[\int_{2^R v}^\infty \frac{(2^R - 1) f_{\lambda_{\max}, \lambda_{\min}}(u, v)}{u} du + \int_v^{2^R v} f_{\lambda_{\max}, \lambda_{\min}}(u, v) \left(\frac{2^{1+R/2}}{\sqrt{uv}} - \frac{1}{u} - \frac{1}{v} \right) du \right] dv \\ \overline{P}_2 &= WN_0 \int_0^\infty \left[\int_{Kv}^\infty \frac{(2^R - 1) f_{\lambda_{\max}, \lambda_{\min}}(u, v)}{u} du + (2^{R/2} - 1) \int_v^{Kv} f_{\lambda_{\max}, \lambda_{\min}}(u, v) \left(\frac{1}{u} + \frac{1}{v} \right) du \right] dv \end{aligned} \quad (38)$$

These can be computed analytically, however the expressions involve sums of hyper-geometric numbers, which are difficult to handle. They are more easily computed using Monte Carlo averaging. It is straightforward to show that $K = 1$ minimizes the average transmit power in Scheme 2, which yields a very simple one-dimensional transmission technique, since the transmit signal lies solely in the dimension of the eigenvector corresponding to the largest eigenvalue of $\mathbf{A}\mathbf{A}^H$.

In Scheme 3 for Rayleigh fading each of the $\|\mathbf{a}_i\|^2$ are i.i.d. central chi-square random variables with U degrees of freedom so the the average transmit power is

$$\overline{P}_3 = LWN_0(2^R - 1) \int_0^\infty \frac{1}{\Gamma(U)^L} x^{U-2} e^{-x} [\gamma(U, x)]^{L-1} dx \quad (39)$$

where $\gamma(n, x) = \int_0^x t^{n-1} e^{-t} dt$ is the incomplete Gamma function. For $L = 2$ we have the closed-form expression [27, 6.455]

$$\overline{P}_3 = \frac{2WN_0(2^R - 1)\Gamma(2U - 1)}{\Gamma(U + 1)2^{2U-1}} {}_2F_1(1, 2U - 1; U + 1; 1/2) \quad (40)$$

which for $U = 2, 3, 4$ yields $P_{\text{subopt}2} = WN_0(2^R - 1)\{1/2, 5/16, 5/64\}$

4.1 Numerical Results

In figures 6,7 we show the necessary transmit power to achieve 2 and 4 bits/s/Hz for the three schemes. We also show the outage probabilities for optimal space-time coding schemes. We first note that the gains due to beamforming are smaller than in the $1 \times N$ antenna diversity case, especially when there are more antennas at the receiver than at the transmitter. This is because with many receive antennas the transmit antenna coefficients must somehow be chosen to achieve an acceptable signal level on each receive antenna.

Secondly we note that, at least for $R = 2$ bits/s/Hz, the first suboptimal transmission technique where we transmit only in the direction of the eigenvector with the larger eigenvalue (B_2) suffers very little in terms of average transmit power. At $R = 4$ bits/s/Hz this is no longer the case. The simple sub-optimal scheme (B_3) which does not use an eigenvalue decomposition suffers a small additional power penalty (1-2dB) which is most likely due to the selection component. This penalty will most likely become more severe when $\min(M, N) > 2$.

Finally another interesting observation is that the required average power for a 1×4 system (see figure 4) is slightly less than that of a 2×2 system. From an implementation point of view this is interesting since it shows that it may be unnecessary to distribute antennas evenly between transmitter and receiver. We should point out that this behaviour is different from the case of the average capacity [5]. This observation must be studied in greater depth.

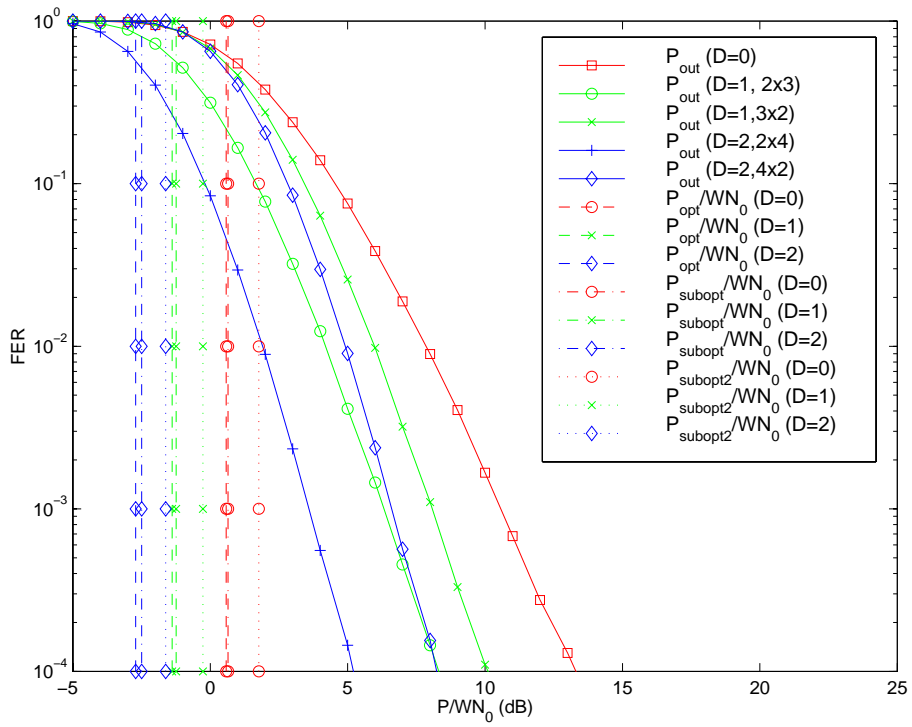


Figure 6: Generalized Beamforming, Ideal Performance ($R = 2$ bits/s/Hz)

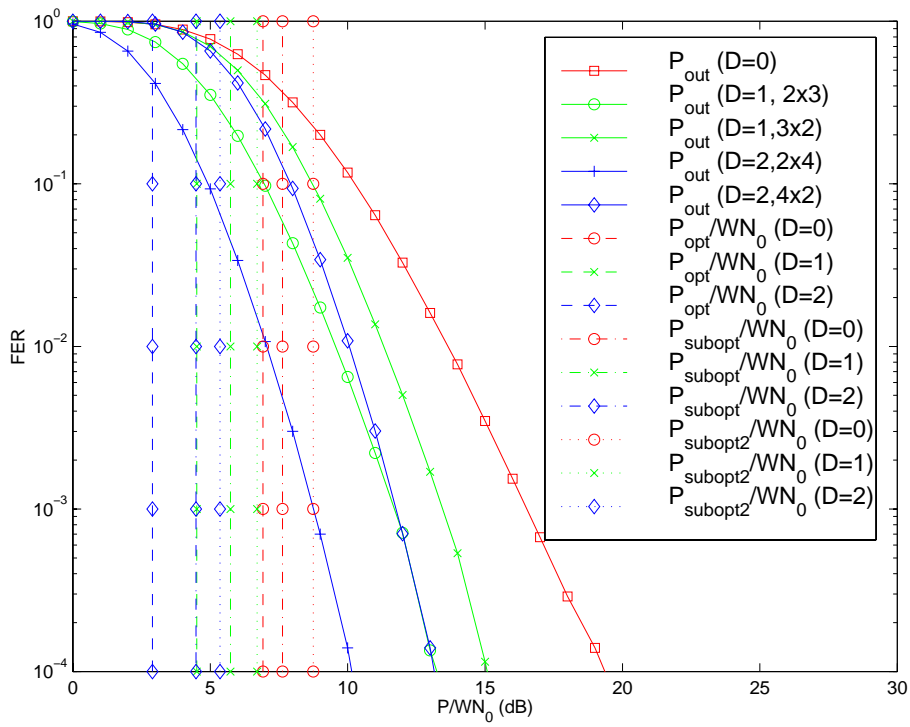


Figure 7: Generalized Beamforming, Ideal Performance ($R = 4$ bits/s/Hz)

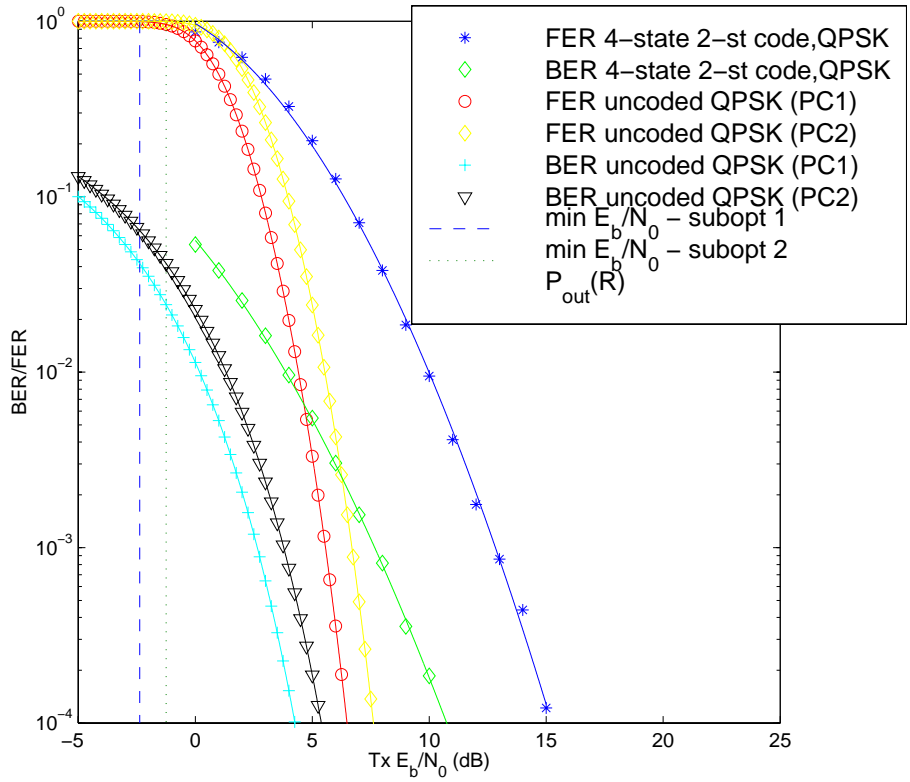


Figure 8: BER/FER Comparison ($R = 2$ bits/s/Hz)

5 The effect of imperfect antenna gain estimates

A potential practical problem with power control schemes is the effect of imperfect channel state information at the transmitter. This would be the case, for instance, for systems with quickly moving terminals. Let us now address this issue for the single channel case ($M \times 1, 1 \times N$). In the same vein, the recent work [28] analyses a multiple-transmit antenna system with imperfect CSI at the transmission end for systems without delay constraints (i.e. ergodic channels).

We will assume that the channel estimates in block k , $\tilde{A}_{1,i}(k)$, $i = 1, \dots, \{M, N\}$, are linear predictions based on information received in D_p prior code blocks and are thus correlated with the current channel gains $A_{1,i}(k)$, $i = 1, \dots, \{M, N\}$. We write the estimates in code block k as

$$\tilde{A}_{1,i}(k) = \sum_{j=1}^{D_p} b_j A_{1,i}(k-j) \quad (41)$$

where b_j are the estimation coefficients. We may choose, for example, a linear minimum mean-square error (MMSE) predictor where $\mathbf{b} = [b_1 \dots b_{D_p}] = \mathbf{r}_A^H \mathbf{R}_A^{-1}$ with

$$\mathbf{r}_A^H = \mathbb{E} A_{1,i}(k) [A_{1,i}^H(k-1) \dots A_{1,i}^H(k-D_p)] \quad (42)$$

$$R_A^{(k,l)} = \mathbb{E} A_{1,i}(k) A_{1,i}^H(l), \quad k, l = 1, \dots, D_p. \quad (43)$$

Let us denote the correlation between the predicted channel gains and the actual gains by ρ . For the linear MMSE predictor $\rho = \mathbf{r}_A^H \mathbf{R}_A^{-1} \mathbf{r}_A$. Its value will depend on the rate of change of the channel (Doppler frequency) and the prediction order D_p . For the classic land-mobile uniform scatterer model we may write

$$\mathbb{E} A_{1,i}(k) A_{1,i}^H(l) = J_0(2\pi f_D |k-l|T) \quad (44)$$

where $f_D = v f_c / c$ is the Doppler frequency, T is the code block duration, v is the relative speed of the receiver/transmitter pair, f_c is the carrier frequency and c is the speed of light.

Our underlying assumption is that the receiver uses the actual channel gain $A_{1,i}(k)$ and the transmitter uses $\tilde{A}_{1,i}(k)$. In the case of receiver diversity with power control we use $A_{i,1}$ to combine the antenna outputs and $\tilde{A}_{i,1}$ to adjust the transmit power. The resulting equivalent channel will be

$$r' = \sqrt{\frac{\sum_{i=1}^M |A_{i,1}|^2}{\sum_{i=1}^M |\tilde{A}_{i,1}|^2}} u' + z' \quad (45)$$

For the case of transmit diversity, the expression will be slightly different since we use $\tilde{A}_{i,1}$ both for combining and power control. Since the signals are not properly combined at the transmitter, there will be a residual phase offset at the receiver (i.e. the signal will not be purely real). Assuming the receiver compensates for this phase offset, the equivalent received signal will be

$$r' = \frac{\left| \sum_{i=1}^N A_{1,i} \tilde{A}_{1,i}^H \right|}{\sum_{i=1}^N |\tilde{A}_{1,i}|^2} u' + z' \quad (46)$$

We expect that the receiver diversity case suffers less with imperfect estimates since the estimates only effect the power control and not the combining.

Defining $\eta_R = \sqrt{\frac{\sum_{i=1}^M |A_{i,1}|^2}{\sum_{i=1}^M |\hat{A}_{i,1}|^2}}$ and $\eta_T = \frac{|\sum_{i=1}^N A_{1,i} \hat{A}_{1,i}^H|}{\sum_{i=1}^N |\hat{A}_{1,i}|^2}$ we have that the BER with QPSK signals in Rayleigh fading with uncorrelated antenna gains are

$$P_{e,R} = E \left[Q \left(\sqrt{2(M-1)\bar{P}\eta_R} \right) \right] \quad (47)$$

$$P_{e,T} = E \left[Q \left(\sqrt{2(N-1)\bar{P}\eta_T} \right) \right] \quad (48)$$

It seems that p.d.f.'s of η_R and η_T cannot be expressed in closed form. We therefore resort to Monte Carlo estimation of the above error probabilities. The results for $\rho = .95, .99, .999$ are shown in Fig. 9, where we observe that a correlation coefficient between the channel gain estimate and the actual gain should be greater than .99 to avoid significant degradation. In table 1 we show the resulting correlation coefficients for different prediction orders and mobile speeds using Jakes autocorrelation model. We have assumed a 2GHz carrier frequency and $T = 1\text{ms}$ which are characteristic of 3rd generation systems. We may conclude, therefore, that the effect of imperfect estimates is not significant even for high mobile speeds (both for TX and RX diversity) when adequate prediction orders are used.

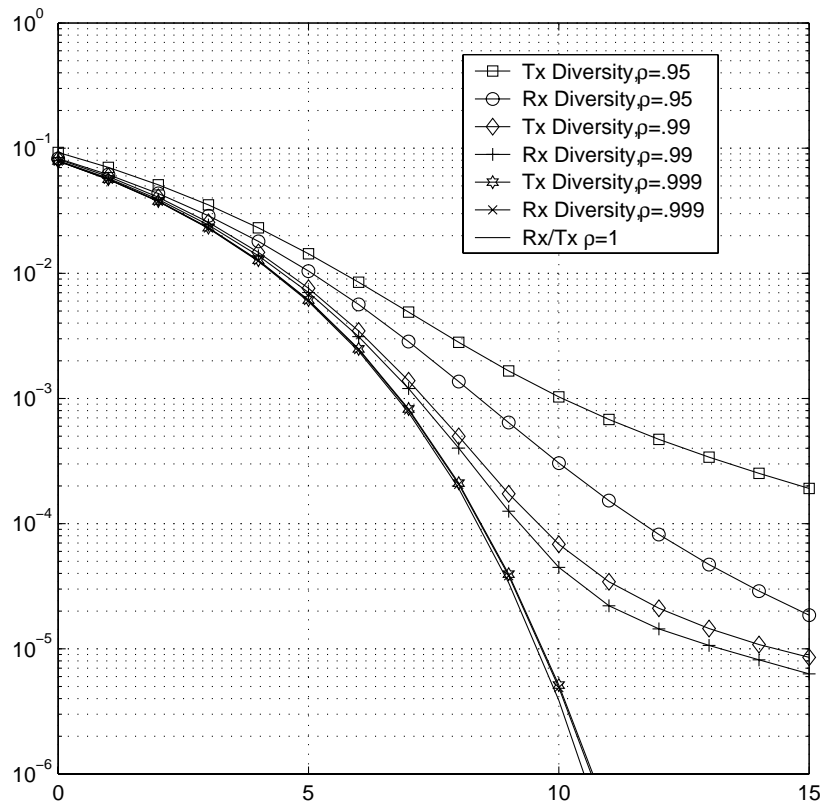


Figure 9: Effect of imperfect channel estimates, $M = 2$ or $N = 2$

$v(m/s)$	$D_P = 2$	$D_P = 3$	$D_P = 4$	$D_P = 5$	$D_P = 6$	$D_P = 7$
1	1.000	1.000	1.000	1.000	1.000	1.000
10	.9963	.9998	1.000	1.000	1.000	1.000
30	.7855	.9255	.9743	.9911	.9969	.9989

Table 1: Correlation Coefficients (ρ) vs. Prediction Order (D_P) and Mobile Velocity (v)

6 Conclusions

In this paper, we examined dynamic power control schemes based on quasi-perfect channel gain estimates for narrowband, delay-limited multi-antenna systems with slow Rayleigh/Ricean fading. The principal results is that significant reductions in average transmit power can be expected compared to systems where the transmit power is fixed. Specifically when a single antenna is present on one end, the power reduction can be greater than 10dB even with a small array (<4 elements). Moreover, the performance is independent of whether the array is transmitting or receiving, which is not the case in fixed power systems [5]. This is because CSI at the transmitter allows us to employ *beamforming* which is analogous to *maximal-ratio combining* at the transmitter, in conjunction with power-control.

We have analyzed the effect of element gain correlation and asymmetries, both of which arise in dual-polarization antennas. In addition, for a linear antenna array we have shown that element spacings greater than one quarter of a wavelength are sufficient to achieve close to the minimum average transmit power.

We then considered generalized beamforming systems where multiple-antennas are present both in the transmitter and receiver. Here the performance is again independent of the direction of communication, which is not the case in systems with fixed transmit power. We show that when the minimum number of elements is two, considerable power savings can be expected, however less than the case where a single-antenna is present on one end. The optimal beamforming scheme requires a complete eigen-decomposition of the channel gain matrix which may be difficult in practice. We therefore illustrate two simpler sub-optimal schemes which do not suffer significant power penalties for moderate information rates. One requires only the eigenvector corresponding to the largest eigenvalue of the channel gain matrix and the second requires no information regarding the eigen-decomposition.

Finally the effect of outdated channel estimates was shown to be small when simple linear prediction is used to estimate the channel gains based on measurements made in previous code blocks. This would make the proposed schemes very interesting for 3rd generation TDD system variants.

References

- [1] "IEEE Personal Communications Magazine," April 1998.
- [2] P. Mogensen, F. Frederiksen, H. Dam, K. Olesen, and S. Larsen, "A Hardware Testbed for Evaluation of Adaptive Antennas in GSM/UMTS," vol. IEEE PIMRC'96, pp. 540–544, Oct. 1996.
- [3] V. Tarrokh, N. Seshadri, and A. Calderbank, "Space-Time Coding for High Data Rate Wireless Communication: Performance Analysis and Code Construction," *IEEE Transactions on Information Theory*, vol. 44, pp. 744–765, March 1998.
- [4] G. Raleigh and J. Cioffi, "Spatio-Temporal Coding for Wireless Communication," *IEEE Trans. on Communications*, vol. 46, pp. 357–366, March 1998.
- [5] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," tech. rep., AT&T Bell Laboratories, 1995.
- [6] G. Foschini and J. Gans, "On Limits of Wireless Communication in a Fading Environment when Using Multiple-Element Antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.
- [7] A. Narula, M. Trott, and G. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. on Information Theory*, vol. 45, pp. 2418–2433, Nov. 1999.
- [8] L. Ozarow, S. Shamai (Shitz), and A. Wyner, "Information Theoretic Considerations for Cellular Mobile Radio," *IEEE Transactions on Vehicular Technology*, vol. 43, pp. 359–378, May 1994.
- [9] Tse, D.N. and Hanly, S.V., "Multi-Access Fading Channels: Part II: Delay-Limited Capacities," *IEEE Transactions on Information Theory*, vol. 44, p. 2815, November 1998.
- [10] E. Malkamäki and H. Leib, "Coded Diversity on Block-Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 771–781, March 1999.
- [11] R. Knopp and P. Humblet, "On Coding for Block-Fading Channels," *IEEE Trans. on Information Theory (to appear)*, vol. 46, pp. 189–205, January 2000.
- [12] G. Caire, R. Knopp, and P. Humblet, "System Capacity of F-TDMA Cellular Systems," *IEEE Trans. on Comm.*, vol. 46, pp. 1649–1661, December 1998.
- [13] G. Caire, G. Taricco, and E. Biglieri, "Optimum Power Control over Fading Channels," *IEEE Transactions on Information Theory*, vol. 45, pp. 1468–1489, July 1999.
- [14] W. Jakes, ed., *Microwave Mobile Communications*. New York: John Wiley and Sons, 1974.
- [15] J. Wagen and M. Keer, "Comparison of Diffraction Coefficients for Propagation Prediction in Micro-Cell," in *COST231, Grimstad, Norway*, May 1993.

- [16] F. Lotse, J. Berg, U. Forssen, and P. Idahl, "Base Station Polarization Diversity Reception in Macrocellular Systems at 1800MHz," in *IEEE Vehicular Tech. Conf.*, pp. 1643–1646, 1996.
- [17] R. Gallager, *Information Theory and Reliable Communication*. John Wiley and Sons, 1968.
- [18] Y. Leung, S. Wilson, and J. Ketchum, "Multi-frequency Trellis Coding with Low-Delay for Fading Channels," *COM*, vol. 41, pp. 1450–1459, Oct. 1993.
- [19] R. Wesel and J. Cioffi, "Fundamentals of Broadcast OFDM," in *Proc. of the 29th Asilomar Conference on Signals Systems and Computers*, 1995.
- [20] K. Mardia, J. Kent, and J. Bibby, *Multivariate Analysis*. Academic Press, 1979.
- [21] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*. PhD thesis, Dept. of Mathematics, Massachusetts Institute of Technology, 1989.
- [22] P. Humblet, *Digital Communications and Radio Transmission*. Institut Eurecom Class Notes, 1998.
- [23] F. Rashid-Farrokhi, K. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE Journal on Selected Areas in Communications*, 1998.
- [24] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *IEEE VTC 99*, pp. 851–856, May 1999.
- [25] A. Goldsmith and P. Varaiya, "Capacity of Fading Channels with Channel Side Information," *IEEE Transactions on Information Theory*, vol. 43, pp. 1986–1992, Nov. 1997.
- [26] M. Schwartz, W. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw Hill, 1966.
- [27] I. Gradshteyn and I. Ryzhik, *Tables of Integrals, series and products*. New York: Academic Press, 1980.
- [28] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *submitted for publication*, 2000.