Outage Probability at Finite SNR

by

Cemal Akçaba

B.Sc. Massachusetts Institute of Technology, Cambridge (2003)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

Master of Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2004

Author ....
Department of Electrical Engineering and Computer Science
July 12, 2004

Certified by ....
Lizhong Zheng
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

Outage Probability at Finite SNR

by

Cemal Akçaba

SUBMITTED ON JULY 12, 2004
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING
IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

Abstract

In this thesis, we present a technique to reduce the outage probability of a single user multiple input multiple output (MIMO) channel when a sub-optimal transceiver architecture is used. We show that in slow-fading scenarios, the outage probability can be improved by using unitary rotations combined with coding. We provide analysis and simulations supporting our approach.

Thesis Supervisor: Lizhong Zheng
Title: Assistant Professor

Acknowledgments

I would like to thank my family for their never-ending support and love. I am indebted to my advisor, Prof. Lizhong Zheng, for his guidance and mentorship. I cannot thank enough to Nicole Zacharia, Eser Keskiner, Gul Gurkan, Oskar Bruening and Jack Holloway who not only provided their friendship, but also taught me a lot things about life. Last but not least, I would like to offer my thanks to my friends, residents of 35-303, who either academically or not-so-academically contributed to completion of this work; Shashi Borade, Baris Nakiboglu, Murtaza Zafer, Anant Srinivas, Amir Khandani, Alessandro Tarello, Masha Ishutkina and Emmanuel Abbe.

Contents

1 Introduction **11**

1.1 Diversity 11

1.2 Spatial Multiplexing 12

1.3 Diversity-Multiplexing Tradeoff 13

1.4 Contribution of this thesis 14

2 Practical Transceiver Designs **17**

2.1 The Decorrelator 18

2.1.1 Description 18

2.1.2 Performance 21

2.1.3 Diversity Order: Decorrelator with Outer Code 23

2.1.4 Diversity Order: Decorrelator with SIC 25

2.2 MMSE 26

2.2.1 Description 26

2.2.2 Performance 27

3 Beyond Diversity **29**

3.1 Formulation 31

3.2 An Example: Decorrelator Bank 33

3.3 Outage Analysis 35

3.3.1 Decorrelator: Outage with Rotations 36

3.3.2 Simulation Results 37

4 Geometric Analysis **41**

4.1 Decorrelator 44

4.2 MMSE 46

5 Conclusions **49**

A Decorrelator Diversity **51**

A.1 Diversity: Sub-block coding 51

B Optimal Receivers with Rotations **53**

List of Figures

2-1 Projection of h_2 into the subspace orthogonal to h_1. 19

3-1 Depiction of sub-blocks 31

3-2 Decorrelator 4-rotation capacity distribution 38

3-3 MMSE 4-rotation capacity distribution 38

3-4 1% Outage Capacity for MMSE and Decorrelator 39

4-1 Depiction of outage event on rotation space 44

4-2 Decorrelator rotation capacity versus θ 46

4-3 MMSE rotation capacity versus θ 48

4-4 MMSE rotation capacity versus θ (varying SNR) 48

Chapter 1

Introduction

Multiple Input Multiple Output (MIMO) systems when used in a point-to-point communication scenario provide substantial capacity increase compared to single antenna systems. We also know that MIMO systems can increase the reliability of a point-to-point link through appropriate design choices. Zheng & Tse demonstrated the existence of a fundamental tradeoff between these two performance measures of wireless systems [6]. Namely, there is a tradeoff between the data rate and diversity provided by a MIMO channel. Diversity can be thought of as a measure of the reliability of a link, and will be explained in detail later in this chapter. Diversity-data rate plots are, therefore, of great importance for system designers and architects as they demonstrate the effect of increasing the data rate of a system on the reliability of the same system. Often systems do not operate at either extreme of the diversity-rate tradeoff, and the diversity-rate tradeoff is an important tool in deciding how to design systems and where to operate them.

1.1 Diversity

Diversity is a measure of reliability of a link. In multiple antenna systems, there is a path between each transmit and receive antenna. If each of these paths are independently faded, then we can obtain independently faded replicas of data symbols by sending them through different paths. Then we can use these independently faded replicas to improve our reception. Diversity can also be thought of as the number of independent fading coefficients in the channel. In a system with n_t transmit, n_r receive antennas, assuming the channel is Rayleigh faded, the maximum diversity gain is $n_t n_r$, the number of independent fading coefficients in the channel. In other words, the fastest error probability can decay with SNR is $\mathsf{SNR}^{-n_t n_r}$. Diversity is obtained via averaging over multiple path gains (fading coefficients). By averaging, the reliability of the overall link is increased, since the failure of the link depends on more fading coefficients with averaging. Intuitively, depending on average of k independent fading coefficients is more reliable than depending on any one of those k fading coefficients. For the averaging to fail, all of the k independent replicas of the data symbol must be lost.

1.2 Spatial Multiplexing

In addition to being used to increase link reliability (provide diversity), MIMO systems can also be used to increase the data rate of a point to point link. Multiplexing can be thought of as the number of parallel spatial channels that can be created using a point-to-point MIMO channel. If the path gains between transmit and receive antenna pairs are independent, multiple spatial channels are created. By sending independently coded data streams through these parallel channels, the data rate of the point to point link can be increased. In an $n_r \times n_t$channel, where path gains between individual antenna pairs are i.i.d Rayleigh variates, ergodic capacity with channel side information at receiver (CSIR) only is given by:

$$C(\mathsf{SNR}) = \mathcal{E}\left[\sum_{i=1}^{\min\{n_t,n_r\}} \log\left(1 + \frac{\mathsf{SNR}}{n_t}\lambda_i^2\right)\right]$$

where $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_{\min\{n_t,n_r\}}$ are the ordered singular values of H, the channel matrix. At high SNR, we can ignore the addition of 1 and approximate the above expression as [4]:

$$C(\mathsf{SNR}) = \min\{n_t, n_r\} \log \mathsf{SNR} + O(1)$$

Thus the channel is said to have $\min\{n_t, n_r\}$ degrees of freedom. $\min\{n_t, n_r\}$ are the maximum number 'single antenna channels' we can embed in the $n_t \times n_r$ system. As we increase SNR, the capacity scales with $\min\{n_t, n_r\}$. Hence, if $n_r > n_t$ adding one more receive antenna will not have a degree of freedom gain. A degree of freedom can only be obtained if we increase $\min\{n_t, n_r\}$ by one.

1.3 Diversity-Multiplexing Tradeoff

Achieving maximal diversity gain $n_t n_r$ requires that we communicate at a fixed rate R, since we are using all the fading coefficients to combat fading. In a way, we are sending replicas of the same data stream through independent paths. A fixed data rate becomes vanishingly small compared to the fast fading ergodic capacity which grows like $\min\{n_t, n_r\} \log \mathsf{SNR}$ at high SNR. Thus, even though we achieve maximal diversity gain, we are not harnessing any of the degrees of freedom offered by the $n_t \times n_r$ channel. In order to harness at least a fraction of the degrees of freedom offered by the channel, we communicate at a rate, $R \approx r \log \mathsf{SNR}$, which scales with increasing SNR. As defined in [6], we define *spatial multiplexing gain* r and *diversity gain* d as

$$\lim_{\mathsf{SNR}\to\infty} \frac{R(\mathsf{SNR})}{\log \mathsf{SNR}} = r$$

$$\lim_{\mathsf{SNR}\to\infty} \frac{P_e(\mathsf{SNR})}{\log \mathsf{SNR}} = -d$$

Then the diversity gain is the rate, average error probability decays in SNR. That is $P_e(\mathsf{SNR}) = \mathsf{SNR}^{-d}$. According to the above definition, any scheme that achieves full spatial multiplexing gain, has zero diversity gain. It is shown in [6] and [4] that for the i.i.d. Rayleigh fading MIMO channel with n_t transmit, n_r receive antennas, the high SNR outage probability at rate $R = r \log \mathsf{SNR}$ is given by

$$P_{outage} \sim \mathsf{SNR}^{-(n_t-r)(n_r-r)}$$

for integer $r = 0, 1, \cdots, min(n_r, n_t)$ [6]. The interpretation of the above result is simple; when operating at a multiplexing gain of r, the maximal diversity gain we can get is $(n_t - r)(n_r - r)$. That is, it is possible to reach a positive diversity gain while attaining some degrees of freedom provided by the channel. There is a fundamental tradeoff between diversity and multiplexing. Higher multiplexing gain comes at the price of lower diversity and higher diversity necessarily decreases the achievable multiplexing gain.

1.4 Contribution of this thesis

There are systems that can achieve the optimal tradeoff between data rate and diversity [1, 6]. We shall call such systems optimal systems from this point on. However, often these systems are hard to deploy due to the complications in their implementations or other practical reasons. Therefore, suboptimal architectures (architectures that cannot achieve the entire optimal tradeoff curve for a MIMO channel) are often deployed in practice. These systems suffer performance penalties either in diversity or data rate due to their simple implementations. Since the reliability of these systems is low compared to optimal architectures, it is often of great interest to increase their reliability without increasing the complexity of their implementations.

This thesis describes a particularly simple method to improve the reliability of some suboptimal architectures. In particular, we will be focusing on the decorrelator and the MMSE transceiver architectures. The method we describe uses unitary rotations with joint-coding in order to improve the diversity performance of these systems. The idea is still to average over multiple path gains to increase reliability. The rotations merely serve as bringing out some of the fading coefficients that are lost due to nulling the interference from other sub-streams. By rotating the channel each time a symbol is transmitted, and coding over the rotated replicas of the same channel realization, we show that the reliability of the decorrelator and the MMSE systems can be improved for Rayleigh fading channels. The observed improvement in diversity is less than unity, but might be significant in some scenarios.

The organization of this thesis is as follows. Chapter 2 describes the suboptimal transceiver architectures we will analyze, and Chapter 3 introduces the rotation method we propose. Chapter 4 provides deeper analysis of the rotation method. We conclude our findings in Chapter 5.

Chapter 2

Practical Transceiver Designs

Practical transceiver implementations rarely employ information theoretically optimum receivers due to complications in the implementation of such receiver architectures. *Maximum likelihood* (ML) detection can be fairly difficult to implement on a cellular phone or wireless sensor due to lack of processing power, battery, time and sometimes all three of these factors. Therefore, suboptimal receiver architectures such as the decorrelator and the *minimum mean squared estimator*, MMSE, are of great practical importance to many applications .

There are schemes that achieve information theoretically optimum capacity using linear receiver architectures in conjunction with successive cancellations [1, 5] These schemes are known as successive interference cancellation (SIC) schemes and require the receiver to successfully decode data of a sub-stream, and encode the successfully decoded data and subtract it from the received data vector, prior to decoding other sub-streams. The process of decoding and subtracting out interference is a significant overhead, and increases the total decoding time. Therefore, most practical receivers do not employ successive interference cancellation.

In what follows, we briefly explain the decorrelator and the MMSE architectures in an effort to provide some background.

2.1 The Decorrelator

2.1.1 Description

The data transmitted on a $n_t \times n_r$ multi-antenna array is often separated into n_t independent data streams that can be coded independently. The reason for this being the fact that code generation for a multi-antenna area is quite complicated and there are no known good coding mechanisms. The symbols across the streams at time m form the vector $\mathbf{x}[m]$ which is transmitted through a $n_r \times n_t$ channel matrix H. The received signal is also corrupted by noise, in which case we have the familiar time-invariant MIMO channel:

$$\mathbf{y}[m] = \mathsf{H}\mathbf{x}[m] + \mathbf{w}[m], \quad m = 1, 2, \cdots \tag{2.1}$$

where $\mathbf{x}[m] \,\epsilon\, \mathcal{C}^{n_t}$, $\mathbf{y}[m] \,\epsilon\, \mathcal{C}^{n_r}$ and $\mathbf{w}[m] \sim \mathcal{CN}(0, I_{n_r})$ are the transmitted signal, received signal and white Gaussian noise at time m respectively.

Therefore, receivers need to separate the data streams that from different 'users'. The term 'user' here is a bit misleading, and it refers to the virtual user which owns one of the independently coded data streams. All the streams in a point to point communication scenario belong to the same user. However, we will stick to this convention and associate each data stream with a virtual 'user'.

It is not clear whether the receiver can separate the data streams efficiently enough so that the resulting system attains full degrees of freedom of a $n_t \times n_r$ antenna array. In such a system, the k^{th} user's data stream faces interference from all other users, and is also corrupted by noise. In such scenarios, our aim is to maximize the signal to interference ratio (SINR) as opposed to signal to noise (SNR) ratio. The matched filtering (maximal ratio combining) aims to preserve as much energy in the signal as possible at the expense of facing high inter-stream interference. The decorrelator is motivated by the same idea, but works at the other end of the spectrum. The idea is to remove the inter-stream interference by projecting the received signal $\mathbf{y}$ on to a subspace that is orthogonal to one containing all the other subspaces. This is

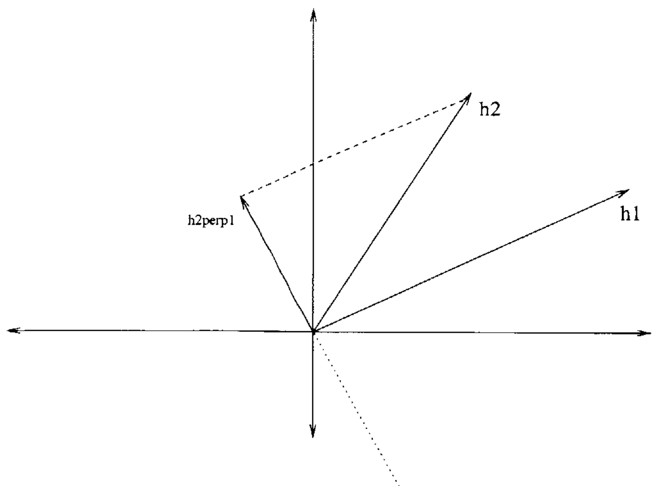

Figure 2-1: Projection of h_2 into the subspace orthogonal to h_1.

the subspace orthogonal to the one spanned by the vectors $h_1, ...h_{k-1}, \cdots h_{k+1}, h_{n_t}$ (h_j's are columns of channel matrix H). The projection operation can be represented by matrix transformation, i.e. projection into a subspace V_k of dimension d_k can be carried out with multiplying by a d_k by n_r matrix Q_k. The rows of Q_k form an orthonormal basis of V_k and hence multiplying any vector v by Q_k will have the effect of projecting v onto V_k. This is shown in figure 2-1.

Obviously, for the interference cancellation to be successful the spatial signature, h_k, of data stream k must not be linear combination of the spatial signatures of the other data streams. If this is the case, the projection will result in a zero vector, and the receiver will fail. Now, if we use more than n_r data streams, this strategy of interference cancellation cannot be successful, since at least one of the processed received streams will be completely zero. Hence we have a natural constraint here, the number of data streams must be equal to or less than n_r. Assuming $n_t \leq n_r$, after the projection the channel for the k^{th} sub-stream looks like

$$y[m] = Q_k h_k x_k[m] + Q_k \mathbf{w}$$

which is the familiar scalar channel with projected spatial signature signal $Q_k h_k$. We know that matched filtering maximizes SNR for such a channel and the signal to noise ratio is given by:

$$\mathsf{SNR} = \frac{P_k||Q_k h_k||^2}{N_o}$$

where P_k is the power allocated to the k^{th} stream, and N_o is the average noise power. The combination of the projection operation and matched filtering is usually called the *decorrelator*. Both operations involved in the decorrelator are linear, hence the decorrelator itself is a linear filter. The overall decorrelating filter is given by:

$$c_{decor_k} = (Q_k^* Q_k)h_k)$$

which is the projection of h_k onto the subspace V_k, expressed in terms of the original coordinates. In effect, the decorrelator maximizes the signal to noise ratio subject to the constraint that the interference from all other users are nulled. We will see the performance implications of this nulling in the next subsection. However, notice that $||Q_k h_k||$ will often be less than $||h_k||$, and we will face some penalty in nulling the interference.

The individual filters can be described by the above formula, however there is a compact form for a bank of decorrelators for each user given by:

$$\mathsf{H}^\dagger := (\mathsf{H}^*\mathsf{H})^{-1}\mathsf{H}^*$$

$\mathsf{H}^\dagger$ is called the pseudo-inverse of the matrix H. The decorrelator for the k^{th} stream is the column k of the pseudo-inverse of the matrix H. This matrix is also known as Moore-Penrose matrix inverse, and is only defined if $\mathsf{H}^*\mathsf{H}$ has an inverse. $\mathsf{H}^*\mathsf{H}$ is only invertible if and only if H has linearly independent columns [2]. This would always be the case for a full rank H, which would be the most relevant channel matrix type for our analysis as we would see later.

2.1.2 Performance

The performance of the decorrelator can be expressed in terms of the capacities of individual sub-streams belonging to each user. The output of each decorrelator bank is a Gaussian channel with SNR given as above. The capacity of such a channel is

$$C_k = \log(1 + \frac{P_k||Q_k h_k||^2}{N_o})$$

The capacity of the entire array is then just the sum of individual sub-streams. Notice that in general $||Q_k h_k|| \leq ||h_k||$, (specifically, when h_k's are not orthogonal, $||Q_k h_k|| < ||h_k||$ will always be true). Nulling the interference discards some information about the vector we are trying to estimate. This means that unless the spatial signatures are in totally different directions, there is no way to cancel interference without losing some portion of data. Indeed, intuitively this is what one would expect.

The above analysis was done for deterministic channel matrix H. For a stationary distribution of the channel matrix H, the achievable fast fading rate is simply the expected value of the expression above. The decorrelator bank performance for a i.i.d Rayleigh fading channel is then:

$$C_{decor} = \mathcal{E}[\sum_{i=1}^{\min(n_t,n_r)} \log(1 + \frac{\mathsf{SNR}}{n_t}||Q_k h_k||^2)]$$

The optimal covariance here is scaled version of the identity, and we have chosen to pour equal powers to each data stream. We can approximate the rate expression above at high SNR, to see the spatial degrees of freedom attained. At high SNR:

$$C_{decor} = \min(n_t, n_r) \log(\frac{\mathsf{SNR}}{n_t}) + \mathcal{E}[\sum_{i=1}^{n_t} \log(||Q_k h_k||^2)]$$

$$C_{decor} = \min(n_t, n_r) \log(\frac{\mathsf{SNR}}{n_t}) + O(1)$$

We see that the decorrelator bank achieves full spatial degrees of freedom of the MIMO channel, since the $\log \mathsf{SNR}$ term above has a $min(n_t, n_r)$ coefficient. Hence the decorrelator bank can reach the full spatial multiplexing gain offered by the channel.

To analyze the diversity gain, we start by restating the problem. We still have the channel model described in 2.1, i.e. we have the bank of decorrelators, belonging to each user and we encode each user's data independently at the transmitter. However, we assume a slow fading scenario, in which channel matrix H is random, but remains fixed once chosen. We will assume for the analysis to follow that H is Rayleigh faded.

Hence, our system looks like a V-BLAST system without successive interference cancellation. Our performance metric is the outage probability, the probability of the random channel capacity falling below some target data rate R. Since each of the sub-streams are independently coded at R, an outage occurs if any of the sub-streams cannot be decoded completely.

The equivalent parallel channel for our system in equation 2.1 can be written as follows [6]:

$$\mathbf{y}_i = \frac{\mathsf{SNR}}{n}\mathbf{g}_i\mathbf{x}_i + \mathbf{w}_i \text{ for } i = 1, \ldots, n_t \tag{2.2}$$

where $\mathbf{x}_i, \mathbf{y}_i, \mathbf{w}_i \in C^l$ are the transmitted, received signals and the noise for the i^{th} sub-stream; $\mathbf{g}_i$ is the gain of the i^{th} decorrelator, i.e. the square root of the signal to noise ratio at the output of i^{th} decorrelator.

Then, an outage happens if any of the scalar channels above falls below the target data rate R. We will assume that each sub-stream is assigned the same rate, $R_i = \frac{r}{n_t}\log\mathsf{SNR}$. The order in which the sub-streams are detected is not important. The receiver can choose arbitrarily, or use a fixed decoding order. Under these conditions, with a Rayleigh faded channel, each of the $\mathbf{g}_i$'s is chi-square distributed with one degree of freedom; $\mathbf{g}_i \sim \chi_2^2$. The constant degree of freedom in $\mathbf{g}_i$'s is due to lack of successive interference cancellation. If we had employed successive interference cancellation, each of the $\mathbf{g}_i's$ would have been chi-square distributed, with i degrees of freedom. Since each sub-stream passes through a scalar channel with gain $\mathbf{g}_i$, an error occurs at the i^{th} sub-stream with probability $P_e\left(\log(1 + \frac{\mathsf{SNR}}{n_t}\mathbf{g}_i^2) < \frac{r}{n}\log\mathsf{SNR}\right)$. When $\mathbf{g}_i$ is chi-squared distributed of order 1, this probability is given by:

$$P_e^i = \mathsf{SNR}^{(1-\frac{r}{n_t})}$$

Hence the diversity of order of the bank of decorrelators is just 1. In the above system, outage occurs whenever one of the sub-channels is in deep fade and cannot support the rate of the stream using that sub-channel. However, we can do better by coding across the sub-channels to provide reliable communication when

$$\sum_{i=1}^{n_t} \log(1 + \frac{\mathsf{SNR}}{n_t}\mathbf{g}_i^2) > R \tag{2.3}$$

The reason this scheme is superior is that each individual sub-stream passes through all the sub-channels; and hence an error in one of these sub-channels might not be fatal and the sub-stream might still be recovered.

2.1.3 Diversity Order: Decorrelator with Outer Code

To derive the diversity order of the bank of the decorrelators, we will need of find the probability that $\sum_{i=1}^{n_t} \log(1 + \frac{\mathsf{SNR}}{n_t}\mathbf{g}_i^2) < R$. Each of the terms in the sum are independent identically distributed variates with the following cumulative distribution function (assuming $n_r = n_t$):

$$F_R(R) = 1 - e^{-\frac{2^R - 1}{2\mathsf{SNR_n}}}$$

where $\mathsf{SNR_n} = \frac{\mathsf{SNR}}{n_t}$. To calculate the probability of sum rate falling below some arbitrary rate R, we need to know the probability density function (PDF) of the summation. We can calculate this PDF by using iterated convolutions, since each of the associated terms are independent. We differentiate the above CDF to get the PDF of the individual terms:

$$f_R(R) = \frac{\ln(2) 2^R e^{-\frac{2^R - 1}{2\mathsf{SNR_n}}}}{2\mathsf{SNR_n}} \tag{2.4}$$

The PDF for the sum is then just

$$f_t(t) = \int_0^t \int_0^{t_{n_t-1}} \cdots \int_0^{t_2} \int_0^{t_1} f_R(t-t_{n_t-1}) f_R(t_{n_t-1}-t_{n_t-2}) \ldots f_R(t_1-\tau) f_R(\tau) \quad \partial\tau\, \partial t_1 \, \ldots \, \partial t_{n_t-1} \tag{2.5}$$

It is difficult to evaluate this integral explicitly. For the outage event, we are only interested in the tail of the probability distribution of the sum. Note that, since each of the PDFs involved in the convolution are one sided, we can approximate these PDFs near their tails, and convolve the approximations to get an approximation to the tail of the PDF of the sum. Let us now substitute $R = \delta \log \mathsf{SNR}$ into 2.4 and take the taylor series expansion of the resulting expression near $\delta = 0$:

$$f_\delta(\delta) = \frac{\ln(2)}{2\mathsf{SNR}} \left(\mathsf{SNR}^\delta + \frac{\mathsf{SNR}^\delta - \mathsf{SNR}^{2\delta}}{2\mathsf{SNR}} + \frac{0.5\mathsf{SNR}^\delta - \mathsf{SNR}^{2\delta} + 0.5\mathsf{SNR}^{3\delta}}{4\mathsf{SNR}^2} + O(\mathsf{SNR}^{-3}) \right) \tag{2.6}$$

Now we can approximate 2.5 by using the first order terms of the Taylor series expansion 2.6:

$$\begin{aligned} f_t(t) &= \frac{(\ln(2))^{n_t}}{2\mathsf{SNR}^{n_t}} \int_0^t \int_0^{t_{n-1}} \cdots \int_0^{t_2} \int_0^{t_1} \mathsf{SNR}^\tau \mathsf{SNR}^{t_1-\tau} \ldots \mathsf{SNR}^{t_{n_t-1}-t_{n_t-2}} \mathsf{SNR}^{t-t_{n_t-1}} \quad \partial\tau\, \partial t_1 \, \ldots \, \partial t_{n_t-1} \\ &= \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{n_t} \frac{\mathsf{SNR}^t t^{n_t-1}}{n_t-1} \quad \textit{for } n_t \geq 2 \end{aligned} \tag{2.7}$$

The outage probability can be readily calculated using 2.7 when r is small ($r \ll 1$)

$$\begin{aligned} P\left(\sum_{i=1}^{n_t} \log(1 + \frac{\mathsf{SNR}}{n_t} \mathbf{g}_i^2) < r\right) &\approx \int_0^r \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{n_t} \frac{\mathsf{SNR}^t t^{n_t-1}}{n_t-1}\, \partial t \\ &= \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{n_t} \int_0^R \frac{\mathsf{SNR}^t t^{n_t-1}}{n_t-1}\, \partial t \\ &= \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{n_t} \left[\sum_{i=0}^{n_t-1} \frac{(-1)^i t^{n_t-i-1} \mathsf{SNR}^t}{(\ln \mathsf{SNR})^{i+1}} \frac{(n_t-1)!}{(n_t-i-1)!} \right]_{t=0}^r \end{aligned} \tag{2.8}$$

Leading term in the above summation is the one that has the lowest $\ln \mathsf{SNR}$ exponent in the denominator, i.e. $i = 0$. One easy way to see this is the following. If this was not the case, the above expression cannot be guaranteed to be positive and hence it would be useless as a probability approximation. As SNR goes to infinity, we can ignore all other terms and calculate $\frac{\log P_e}{\log(\mathsf{SNR})}$ using this term only. This gives us the diversity order we expect, which is n_t:

$$\begin{aligned}\frac{\ln P_e}{\ln(\mathsf{SNR})} &= \frac{\ln\left(\left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{n_t}\frac{r^{n_t-1}\mathsf{SNR}^r}{\ln\mathsf{SNR}}\right)}{\ln\mathsf{SNR}} \\ &= \frac{\ln\left(\left(\frac{\ln 2}{2}\right)^{n_t}\frac{r^{n_t-1}}{\ln\mathsf{SNR}}\mathsf{SNR}^{r-n_t}\right)}{\ln\mathsf{SNR}} \\ &= r - n_t + O(r^{n_t-1})\end{aligned}$$

2.1.4 Diversity Order: Decorrelator with SIC

Although the decorrelator bank receiver achieves maximal spatial multiplexing gain, it falls short of achieving the full diversity gain, $n_r \times n_t$, provided by the MIMO channel. Hence it is not diversity optimal. Before we move on to ways of improving this, we would like to briefly talk about why decorrelator bank performs so poorly. Each of the n_t data streams in the decorrelator bank see exactly one channel coefficient, since everything in the direction of other data streams is being discarded. Hence the decorrelator bank's diversity is lowered by the inter-stream interference. The nulling operation costs us a lot of fading coefficients. One way to increase the number of fading coefficients seen by each channel is to use successive cancellations.

The idea for successive cancellations is simple, once a data stream is successfully recovered, we can subtract it off from the received vector and decrease interference to the receivers of the remaining data streams. In this section we will again assume that we code across the transmit antennas (over the data streams), so that an outage happens only if the sum rate falls short of the target data rate. Hence our expression for outage event is still $\sum_{i=1}^{n_t}\log(1+\frac{\mathsf{SNR}}{n_t}\mathbf{g}_i^2) < R$, however now each of $\mathbf{g}_i$'s are chi-squared distributed with $n_r - (n_t - i)$ degrees of freedom. The streams that are decoded later have more tolerance to error. In effect, successive cancellations ensures that the later streams see more channel coefficients. $n_t - i$ is the number of interfering streams at the i^{th} stage, hence the loss of diversity at stage i is exactly the same as the number of interfering streams needed to be nulled out. The outage probability $P\left(\sum_{i=1}^{n_t} \log(1 + \frac{\mathsf{SNR}}{n_t}\mathbf{g}_i^2) < r \log \mathsf{SNR}\right)$ can be explicitly calculated as in section 2.1.3, and turns out to be :

$$P_e = \mathsf{SNR}^{(1-r)}\mathsf{SNR}^{(2-2r)}\mathsf{SNR}^{(3-3r)} \cdots \mathsf{SNR}^{(n-nr)} \tag{2.9}$$

for a Rayleigh fading square channel ($n_t = n_r = n$). Hence the maximal diversity reached by a decorrelator SIC system is $\frac{n(n+1)}{2}$. By increasing the number of fading coefficients seen by each sub-stream, the decorrelator-SIC receiver increases the diversity order significantly.

2.2 MMSE

2.2.1 Description

The decorrelator was motivated by the fact that it completely nulls out inter-stream interference; in fact it maximizes the SNR among all linear receivers that completely nulls out the interference. However, eliminating the interference causes some loss of the signal of interest as well. There is a natural tradeoff between eliminating the inter-stream interference and preserving as much energy in the signal as possible. The decorrelator performs very well at high SNR, where inter-stream interference is dominant over noise, but performs poorly at low SNR where noise is the primary factor impeding the performance. The MMSE filter optimally trades off inter-stream interference and background Gaussian noise [4]. The decorrelator maximizes SNR within all receivers that completely null out the interference, however intuitively we need a receiver that maximizes SINR for any value of SNR. Such a receiver would behave like a decorrelator when the inter-stream interference is large, and like a maximal ratio combiner when the inter-stream interference is small. The MMSE receiver was formulated in [4]. We use the same model as in section 2.1.1 equation 2.1, however we denote $z_k[m]$ as the noise plus interference faced by data stream k:

$$y[m] = h_k x_k[m] + z_k[m]$$

where $z_k = \sum_{i \neq k} h_i x_k[m] + w[m]$. The covariance of z_k is given by $K_{z_k} = N_0 I_{n_r} + \sum_{i \neq k}^{n_t} P_i h_i h_i^*$ where P_i is the power associated with the data stream i. The filter that will first whiten the noise, and then maximize the signal to noise ratio is simply the inverse of the noise plus interference covariance multiplied by the spatial signature of user k:

$$c_{mmse} = \left(N_0 I_{n_r} + \sum_{i \neq k}^{n_t} P_i h_i h_i^* \right)^{-1} h_k$$

The SINR of user k and the rate achieved is then given by:

$$\mathsf{SINR}_k = P_k \left(N_0 I_{n_r} + \sum_{i \neq k}^{n_t} P_i h_i h_i^* \right)^{-1} h_k$$

$$C = \sum_{i=1}^{n_t} \log(1 + \mathsf{SINR}_k)$$

2.2.2 Performance

At high SNR, the MMSE filter reduces to the decorrelator, so we readily know that it also achieves maximal spatial degrees of freedom offered by MIMO channel. At low SNR, the MMSE filter reduces to the maximal ratio combiner and therefore performs better than the decorrelator [4].

An outage probability calculation for a bank of MMSE's, like that of section 2.1.3, is difficult since SINR's of different sub-streams are correlated for the MMSE receiver. In [6], it is shown that MMSE filter and successive cancellations can achieve the entire optimal tradeoff curve. For any realization of channel matrix H, MMSE-SIC can achieve the mutual information of the channel, hence it can also reach optimal outage performance offered by the channel.

The reason MMSE performs better is that the channel gains $\mathbf{SNR}_i$ or $\mathbf{g}_i$'s are independent in the case of decorrelator, whereas SINR's are negatively correlated in the case of MMSE. The data streams in a MMSE-SIC see all the $n_t n_r$ channel coefficients offered by the MIMO channel.

To summarize, we discussed the decorrelator and MMSE transceiver architectures in this section. We have seen that when successive interference cancellations is not used, the outage performance of the bank of decorrelators or MMSE's is fairly poor.

Chapter 3

Beyond Diversity

In the last section, we have seen that in a Rayleigh faded $n \times n$ channel, the decorrelator and the MMSE can reach a maximal diversity of n in the absence of successive cancellations. This is quite small compared to the full diversity gain, n^2 offered by the MIMO channel. Augmented with SIC, the MMSE can reach the full diversity gain. The reason MMSE-SIC performs so well with appropriate coding is that all the n^2 channel coefficients are 'seen' by all the data streams. In a decorrelator or MMSE without SIC, each data stream is affected by only n coefficients, hence it is not possible to reach a higher diversity gain. However, it might be possible to increase the outage probability performance and still use a decorrelator or MMSE without SIC. This can be achieved in a variety of ways, but the underlying idea in all of them is to average over multiple path gains to increase reliability.

To facilitate how averaging over more channel gains might help, imagine a block fading scenario with multiple symbol transmissions happening within the same channel coherence block. In the fast fading scenario, we exploit the randomness of the channel by coding over many independent channel realizations in order to reach the ergodic capacity. We could do the same trick for the block-fading channel if the channel gains of different symbol transmissions were different. An extreme case of this would be to get a new channel every time we transmit a symbol. Most of the time, the coherence block is quite long, so we can segment it into equal length short sub-blocks. Then, by coding over N short sub-blocks within the original block, we get an error only if a codeword is confused with another codeword in all sub-blocks [6]. And hence, we can say that

$$P^N_{outage} \sim (P^1_{outage})^N$$

where P^1_{outage} is the outage probability within a single sub-block (see Appendix A). Hence, the diversity gain we can expect from a bank of decorrelators is Nn, since the maximal diversity of single sub-block decorrelator bank is just n. That is, if we can afford to increase our code length, we can operate at a higher rate by reducing our diversity gain. The diversity performance offered by coding over N sub-blocks is quite impressive. This scheme provides an order of N improvement over the single block scheme evaluated in Chapter 2 equation 2.8. The improvement is caused by having N independently faded blocks. As N approaches ∞, the outage (error) probability becomes vanishingly small and we can communicate reliably at the ergodic capacity.

Now let us modify the above scenario a bit. We assume we do not have N independently faded sub-blocks (hence we can never reach the gain above), and that our coherence block is long enough so that we can transmit multiple times within a block as before. Even though we cannot have independent and hence perfectly uncorrelated channel gains for each sub-block, we can have somewhat uncorrelated channel gains in the case of suboptimal receivers. By coding over these correlated (in fact highly correlated for the most part) channel replicas, it is possible to decrease the outage probability and increase the reliability of the system by having each data sub-stream see more fading coefficients. In effect, some of the fading coefficients that are not seen in the original system are brought out by having each sub-stream go through many correlated channel gains. In the remainder of this chapter, we give an example of a technique that can exploit more of the fading coefficients to decrease the outage probability compared to decorrelator and MMSE transceivers described in the previous chapter. The technique we describe uses unitary rotations with joint-coding in order to improve the diversity performance of these systems. The rotations merely serve as bringing out some of the fading coefficients that are lost due to nulling the

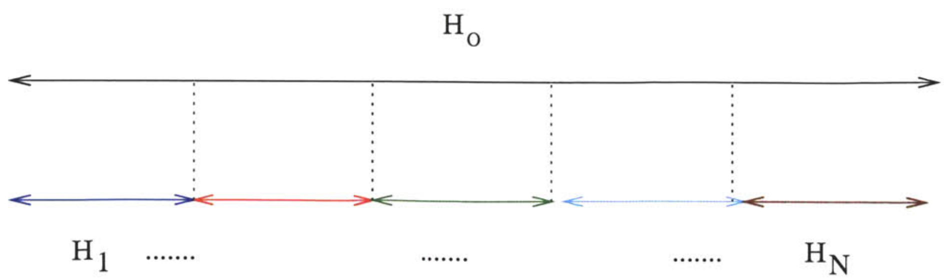

Figure 3-1: Depiction of sub-blocks

interference from other sub-streams. The technique we describe cannot be used for the optimal transceivers, since such transceivers already achieve the best diversity performance offered by a Rayleigh faded MIMO channel. For a proof of this, see Appendix B.

In this thesis, we focus on correlated channel coefficients with suboptimal transceiver architectures unlike the previous work in the area [6]. This problem is challenging to work with for the following reasons:

- Due to the correlation of channel coefficients, the outage probability cannot be calculated explicitly.
- The improvement observed in the outage probability, is usually too small to be captured by diversity analysis but could be important in practice. Therefore a more precise formulation that goes beyond the diversity analysis is required to observe the improvement.

3.1 Formulation

Motivated by coding over independent sub-blocks described earlier, we can generate highly correlated (but not identical) sub-blocks from a given channel matrix and code over the correlated sub-blocks. We propose randomly rotating the input vector each time we transmit within each coherence block. We assume the receiver and the transmitter agree on k rotations at the beginning of transmission, and also that k

symbols can be transmitted within the coherence time of each block. We will denote each of the k rotation matrices with Q_m, with m being the index of a particular rotation within the rotation set $(1 < m < k)$.

During each coherence block, the transmitter encodes n_t data streams independently, and transmits kn_t symbols, n_t symbols a time in k transmissions. The symbol vector $x[m]\ \epsilon \mathcal{C}^{n_t}$ at time m contains the symbol for each data stream in one entry. Then we rotate $x[m]$ with the m^{th} rotation matrix Q_m. The receiver upon receiving the vector $y[m] \epsilon \mathcal{C}^{n_r}$ of entire k sub-blocks, then decodes each data stream using either the MMSE or the decorrelator. Since the receiver also knows the order of rotations, it can perfectly recover each of the sub-streams.

For simplicity, we propose choosing k unitary rotation matrices uniformly at random. We could have chosen to draw the rotation matrices from any arbitrary distribution. However, we argue, without proof, that a uniform distribution is at least optimal in the asymptotic sense. Intuitively, without any prior information about the channel state, any direction of rotation is as good of a choice as the other. As long as we do not choose our rotations to be very similar, this scheme works fine.

However choosing k rotation matrices independently at random will not give the optimal performance for finite k in almost all cases. Since with independent choice of rotations, there always is the chance that two rotations are arbitrarily close to each other, hence we cannot get different coefficients from these two rotations. In other words, one can always come up with better set of rotations, i.e. one that covers more of the rotation space than the one generated randomly. For example, in the all real case, we should choose the angle of rotation uniformly between 0 and 2π, as rotating the channel in any direction is good for our purposes. However, when chosen randomly, any one of these rotation matrices might be similar to another one, as our selection scheme does not guarantee that this would never happen and fine tuning on the set of rotation matrices chosen this way is likely to improve the performance. When k goes to infinity, then our fine tuning will be unable to find a better set of rotations, hence random generation of rotation matrices will be as good as fine tuning. These issues will not be discussed in further detail in this thesis, but are interesting problems to look at.

By using rotations, we divide a single block into k sub-blocks. In effect, we create the new channel picture depicted in figure 3-1. However, note that now the H_i's are not independent from each other whereas before we assumed their independence. We now give the following example to illustrate how rotations create sub-blocks with different channel gains.

3.2 An Example: Decorrelator Bank

As an example, consider a decorrelator bank system that does not employ successive interference cancellation of sub-streams. The channel model is identical to the one used in section 2.1.1 of Chapter 2. We will assume we have a 2×2 square channel, with all real and deterministic entries. We will later relax this constraint and return to our original channel.

$$y = \mathsf{H}x + w$$

where $\mathsf{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$, w being Gaussian noise with covariance I_2.

We further assume that the input data is divided into 2 independent sub-streams which are transmitted on different antennas. Then the equivalent channel is given by:

$$y_i = \frac{\sqrt{\mathsf{SNR}}}{2}\mathbf{g}_i x_i + w, i = 1, 2$$

Each $\mathbf{g}_i^2$ is the magnitude squared of projection of each user's channel onto a subspace perpendicular to channels of all other users. For $n = 2$, with $h_1 = [h_{11}h_{21}]^\dagger$, $h_2 = [h_{12}h_{22}]^\dagger$, $\mathbf{g}_i$'s are defined as $\mathbf{g}_1^2 = ||h_1 - \frac{<h1,h2>}{||h2||}h_2||^2$ and $\mathbf{g}_2^2 = ||h_2 - \frac{<h2,h1>}{||h1||}h_1||^2$. Now consider rotating the input x with a unitary rotation matrix $Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}$. The rotated channel will then be :

$$HQ = \begin{bmatrix} \sum_{i=1}^{2} h_{1i}q_{i1} & \sum_{i=1}^{2} h_{1i}q_{i2} \\ \sum_{i=1}^{2} h_{2i}q_{i1} & \sum_{i=1}^{2} h_{2i}q_{i2} \end{bmatrix}$$

Note that channel gains for the rotated channel are different from the channel gains for the original channel. To see this more carefully note that $< h_1^{'}, h_2^{'} >$ with $h_1^{'} = [\sum_{i=1}^{2} h_{1i}q_{i1} \ \sum_{i=1}^{2} h_{2i}q_{i1}]^{\dagger}$ and $h_2^{'} = [\sum_{i=1}^{2} h_{1i}q_{i2} \ \sum_{i=1}^{2} h_{2i}q_{i2}]^{\dagger}$for the rotated channel is the following

$$< h_1^{'}, h_2{'} >= \sum_{j=1}^{2}\sum_{i=1}^{2}(h_{1i}h_{1j} + h_{2i}h_{2j})q_{i1}q_{j2}$$

$$\frac{< h_1^{'}, h_2{'} >}{< h_2^{'}, h_2^{'} >} = \frac{\sum_{j=1}^{2}\sum_{i=1}^{2}(h_{1i}h_{1j} + h_{2i}h_{2j})q_{i1}q_{j2}}{\sum_j \sum_i (h_{1i}h_{1j} + h_{2i}h_{2j})q_{i2}q_{j2}}$$

User 1 will have $h_{1\perp 2}^{'}$ as its effective channel:

$$h_{1\perp 2}^{'} = \begin{bmatrix} \sum_{i=1}^{2} h_{1i}(q_{i1} - \zeta q_{i2}) \\ \sum_{i=1}^{2} h_{2i}(q_{i1} - \zeta q_{i2}) \end{bmatrix}$$

with $\zeta = \frac{<h_1^{'}, h_2{'}>}{<h_2^{'}, h_2^{'}>}$. Note that $h_{1\perp 2}$ will be equal to $h_{1\perp 2}^{'}$ if $Q = \mathbf{I}_n$. In general, for different choices of Q, $h_{1\perp 2}^{'}$ and $h_{2\perp 1}^{'}$ will be different from $h_{1\perp 2}$ and $h_{2\perp 1}$ and hence the channel capacity obtained will be different.

As seen from this example, the rotated channel's capacity will be different from the original channel's capacity for different choices of rotation matrix Q. It can be lower or higher than the capacity given by the original channel. The reasoning behind the different capacity we obtain for the decorrelator bank is because each sub-stream in the rotated channel has different channel gains. The channel gains for each sub-stream in the rotated channel are different since each sub-stream sees a different set of n coefficients than the original channel. If we can transmit many times with a fixed channel realization, maybe with appropriate coding, we can recover some of the diversity advantage by having each sub-stream see more channel coefficients.

3.3 Outage Analysis

The independence of the H_i's in figure 3-1 allows us to use a random Gaussian code to code over different blocks to achieve the ergodic capacity of the channel. For the correlated sub-blocks, we can still use a random Gaussian code for each sub-stream, however, we can do better by using some other code than the Gaussian, since the channel gains of different sub-blocks are correlated.

For the analysis to follow, we will assume that we will be using random Gaussian code in calculating the capacity of the channel since this simplifies the capacity calculations and gives a lower bound on the performance. We will not discuss the problem of finding a better code that probably works better with a particular set of rotations, although it is a quite challenging and interesting problem.

Assuming a $n \times n$ channel, by coding over k sub-blocks, we can achieve the average capacity of a block:

$$C^k = \frac{1}{k} \sum_{j=1}^{k} \sum_{i=1}^{n} \log \left(1 + \frac{\mathsf{SNR}}{n} \mathbf{g}_{ij}^2\right) \tag{3.1}$$

(i.e. C^k represents the average capacity with the original block broken into k sub-blocks.) However, $\mathbf{g}_{ij}$'s are correlated now, and outage analysis is not straightforward. Calculating the probability density function of the C^k is difficult due to the correlation of $\mathbf{g}_{ij}$'s. We can approximate $P(C^k < R)$ using an upper bound, and see how it compares to outage performance when rotations are not employed $P(C^1 < R)$:

$$P(C^k < R) \leq P(C_1 < kR)\ P(C_2 < kR)\ \cdots\ P(C_k < kR)$$

with $R = r \log \mathsf{SNR}$, $C^k = \frac{1}{k} \sum_{j=1}^{k} C_j$ and $C_j = \sum_{i=1}^{n} \left(\log(1 + \frac{\mathsf{SNR}}{n} \mathbf{g}_i^2\right)$. C_i's above are identically distributed, and $P(C_i < kR)$ can be calculated explicitly. Then, $P(C^k < R) \leq (P(C_1 < kR))^k$. Note that with $C^1 = C_i$ for any i. We claim that $P(C^k < R)$ is less than $P(C^1 < R)$ for sufficiently small R. That is, when operating at very low rates, the outage probability of a system that employs rotations over k sub-blocks is lower than that of one that does not employ rotations. The argument is straightforward for the decorrelator bank, and we provide in the next section.

3.3.1 Decorrelator: Outage with Rotations

Assuming the receiver uses the decorrelator to recover each sub-stream and assuming the square channel model of section 2.1.1 , we need to calculate $P(C^1 < R)$ and $P(C^1 < kR)^k$.

$P(C^1 < R)$ and $P(C^1 < kR)$, with $R = r \log \mathsf{SNR}$, can be estimated by the approximation derived in equation 2.8, provided that kr is sufficiently small. This implies that if the number of sub-blocks, k, is small, then

$$P(C^1 < kR) \approx \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^n \frac{(kr)^{n-1}\mathsf{SNR}^{kr}}{\ln \mathsf{SNR}} \tag{3.2}$$

Now for $(P(C^1 < kR))^k$ will be less than $P(C^1 < R)$ if we are operating at the outage tail. By equating $P(C^1 < kR)^k$ to $P(C^1 < R)$, and solving for r we find that

$$P(C^1 < kR)^k < P(C^1 < R) \tag{3.3}$$

if $0 \leq r \leq e^{-W(k,n,\mathsf{SNR})}$ where $W(z)$ is a Lambert's W function. $W(k, n, \mathsf{SNR})$ will be finite and positive for finite k, n and SNR, but will tend to infinity for arbitrarily larger k, n or SNR. This means that as SNR goes to infinity, our approximation and probability inequality 3.3 is valid in a vanishingly small region.

Assuming we are operating in outage tail defined above and k is small, the outage probability for the rotation system given in equation 3.2 will have a diversity order of kn. This is a very promising gain, however there are number of things that need to emphasized. The diversity order kn will be valid for only small k, e.g. $k = 2$, and it would not scale linearly with increasingly large k. The reasoning for this is simple. The probability approximation we used $P(C^1 < kR)$ will only be valid if kr is not significantly larger than r, as the probability approximation we calculated is valid for $r \approx 0$. In other words, the performance improvement will not scale linearly, but settle after a small number of rotations have been used. For example, in an actual system, using 4 rotations will have some performance benefit, whereas using 64 rotations will improve very little compared to 4 rotations.

The analogous analysis for the MMSE transceiver is difficult. The channel gains. $\mathbf{g}_i's$, that belong to each rotated channel replica are already correlated with each other. Hence an analysis like the above gives only a very coarse estimate, and cannot capture the additional benefit of using the MMSE transceivers. A more rigorous analysis involving Chebyshev type of tail bounds did not give enough accuracy to capture the effect of the MMSE rotations. Therefore, we omit the analysis here and provide a set of simulations demonstrating the reduction in the outage probability for the MMSE transceivers.

3.3.2 Simulation Results

In this section, we provide MATLAB simulation results for the decorrelator and the MMSE transceiver architectures. From here on we will refer to a system that uses n sub-blocks and hence rotations as an n-rotation system. In our simulations, we used the channel model of equation 2.1, and assumed H to be consisting of circularly symmetric Gaussian entries with unit variance (i.e. $h_{kj} = x_{kj} + \mathbf{i}\, y_{kj}$ with $x_{kj}, y_{kj} \sim \mathcal{N}\left(0, \frac{\sqrt{2}}{2}\right)$). We also assumed unit noise covariance, $\mathcal{E}\left[ww^{\dagger}\right] = \mathbf{I}_{\mathbf{n_t}}$. The n unitary rotations used were generated uniformly in random and same set of rotations was used for the MMSE and the decorrelator.

Figure 3.3.2 and 3.3.2 show the distribution of a 4-rotation system capacity compared to the distribution of a no rotation system capacity for the decorrelator and the MMSE transceivers respectively. Note that for both the decorrelator and the MMSE transceivers, the distribution of 4-rotation system lies under the no-rotation curve near $R \approx 0$.

Figure 3.3.2 shows the 1% outage capacity for 4-rotation MMSE and decorrelator systems. 1% outage capacity refers to the capacity point when the outage probability is $\frac{1}{100}$. The improvement we see in the MMSE transceiver is higher at high SNR, and almost negligible at low SNR. We will discuss the reasons for this in the next chapter.

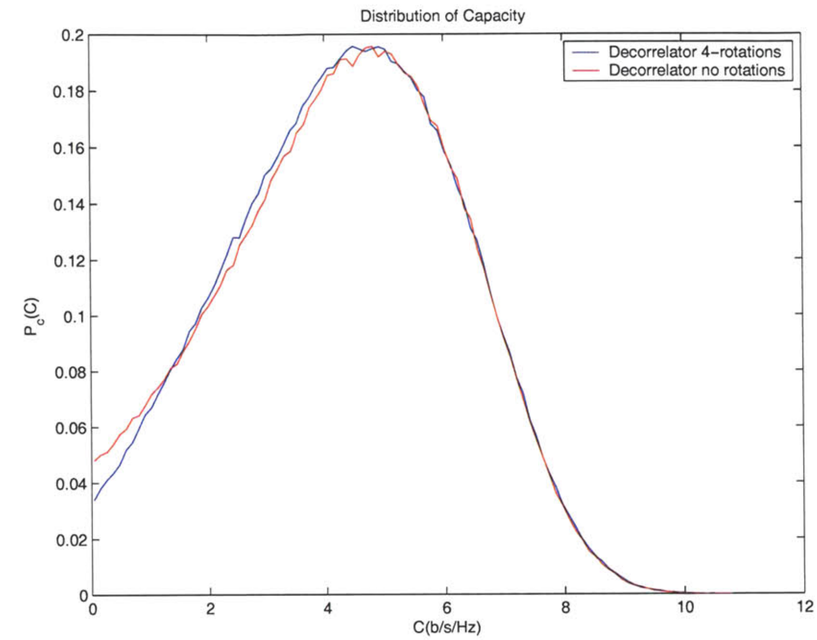

Figure 3-2: Decorrelator 4-rotation capacity distribution

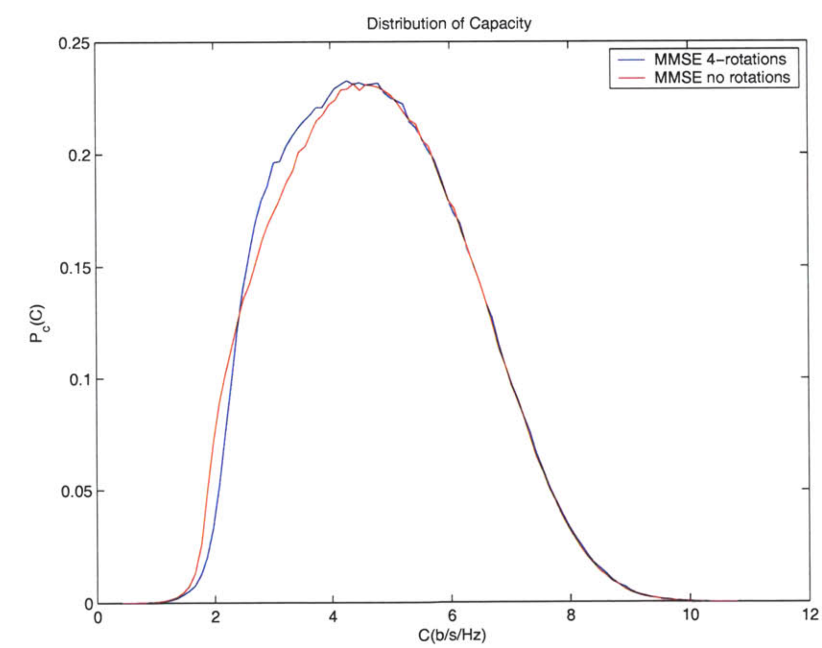

Figure 3-3: MMSE 4-rotation capacity distribution

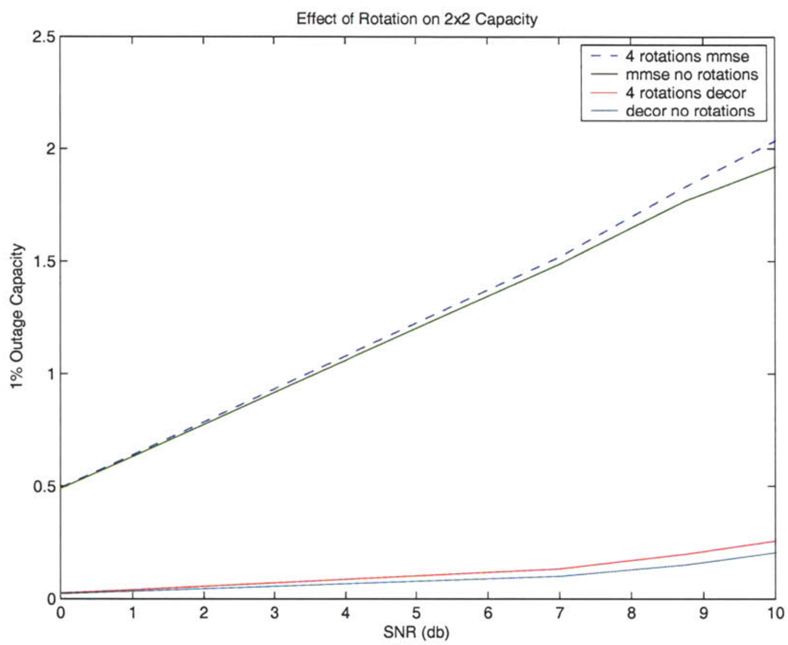

Figure 3-4: 1% Outage Capacity for MMSE and Decorrelator

Chapter 4

Geometric Analysis

The previous section laid out the foundations of rotation scheme and demonstrated its impact on the outage probability. If we look at equation 3.1, we see that for a rotation system to perform better than a regular system, for a particular realization in outage, at least one of our rotations should give us a sufficiently high capacity so that the average capacity is greater than the target data rate. Put slightly differently, for a rotation to help, we must be facing a channel realization, that supports the target data R for sure with optimal transceivers and furthermore also supports the data rate with a right rotation with sub-optimal transceivers, but is in outage due to poor arrangement of individual channels. If the channel is in outage because of deep fade, there is nothing a rotation can do to get it out of outage. In this section, we will be using the term "channels" to refer to the sub-channels belonging to virtual users, i.e the columns of channel matrix H.

This raises an interesting question, for a particular realization of channel matrix, how many of k rotations might take us out of outage? More importantly, what kind of channel realizations can be improved by rotations? That is, with some channel realizations no matter how the channel is rotated, there will be outage, whereas with some of them almost all of the rotations will help. Are these channel events different for the MMSE and the decorrelator?

In this section, we describe in which channel events rotations become useful. Our treatment will be focused on the decorrelator and the MMSE architectures. There are several reasons why rotations can help in moving a channel out of outage. The decorrelator projects the signal of each user, into a subspace perpendicular to all other users. The obvious way a rotation might help, is to make the channels of each user perpendicular to other channels. In that way, less of the signal content will be thrown away, and a higher capacity would be reached. The less obvious is how a rotation might effect capacity when the channels of users are almost perpendicular to each other. This happens by redistribution of channel gains between channels, and can have some capacity impact as we describe later on. With the first case we mentioned, we would expect most of the k rotations to improve our performance, as the rotations will change the relative directions of the channels. However, with the second case, when the channels are almost perpendicular, we need a 'lucky' rotation to tune the channel coefficients in the right way in order to maximize the capacity achieved. Therefore, we would expect this to be a much less likely event, considering that we have finite amount of rotations.

The above statements are valid for the MMSE transceiver as well at relatively high SNR, but how does a rotation-augmented MMSE transceiver behave when operating at low SNR?

With the equivalent channel model given in chapter 2, we have the following maximization problem:

$$\max_{\mathbf{Q}\mathbf{Q}^{\dagger}=\mathbf{Q}\mathbf{Q}^{-1}=\mathbf{I_n}} \sum_{i=1}^{n} \left(\log(1+\mathsf{SNR}\mathbf{g}_i^2)\right)$$

with $\mathbf{g}_i^2$ are channel gains derived from rotated channel matrix $\mathsf{H}\mathbf{Q}$ as given in Chapter 2. The solution to the above problem can be found by setting up the appropriate Lagrangian. The solution turns out to be very similar to the water-filling solution for the MIMO channel, with the obvious modification that we provide equal power to each eigenmode. The maximizing rotation $\mathbf{Q}_{\mathbf{max}}$ turns out be related to the singular value decomposition of channel matrix H:

$$\mathsf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^{\dagger}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{\Lambda}$ is a real diagonal matrix. Then we have $\mathbf{Q_{max}} = \mathbf{V}$. By a rotation of $\mathbf{V}$ we create n sub-channels which are all orthogonal to each other. To see this note that the columns of matrix $\mathbf{U\Lambda}$ are identical to the columns of $\mathbf{U}$, except a scaling of λ_i, where λ_i's are the singular values of H and are the diagonal entries of the matrix $\mathbf{\Lambda}$. The columns of matrix $\mathbf{U}$ are orthogonal to each other as $\mathbf{U}$ is unitary, and scaling the columns with a real constant will not change their orthogonality. It is also true that any matrix $\mathbf{A}$ that has its columns orthogonal to each other, can be represented as a product of a unitary matrix $\mathbf{U}$ and a diagonal matrix $\mathbf{\Lambda}$ that contains the singular values of $\mathbf{A}$. Hence, $\mathbf{U\Lambda}$ always has orthogonal columns.

We can set up the same problem, to minimize the capacity expression, which would give us the worst case Q. The actual solution of the above problem is of little relevance. The important thing is that there is a maximizing rotation, $\mathbf{Q_{max}}$, with an associated maximum capacity $C_{\mathbf{Q_{max}}}$, and a minimizing rotation, $\mathbf{Q_{min}}$ with an associated minimum capacity $C_{\mathbf{Q_{min}}}$. Surely, these maximum and minimum capacities are with respect to the sub-optimal architectures being used and do not represent the true capacity of the channel at hand. Note that it is always true that $C_{\mathbf{Q_{min}}} \leq C_{\mathbf{Q_{max}}} \leq C_{optimal}$, where $C_{optimal}$ is the capacity of an optimal system.

When $\mathbf{Q_{max}} = \mathbf{I_n}$, then our original channel is already good, and most of our rotations will end up giving a smaller channel capacity. Similarly, with $\mathbf{Q_{min}} = \mathbf{I_n}$, our original channel is the worst possible and we can only improve by using rotations.

A channel realization that can be improved by rotations necessarily has $C_{\mathbf{Q_{min}}} < R < C_{\mathbf{Q_{max}}}$. The probability of improvement via rotations increases as $C_{\mathbf{Q_{max}}} - R$ increases at the expense of $R - C_{\mathbf{Q_{min}}}$. If $|C_{\mathbf{Q_{max}}} - R| > |R - C_{\mathbf{Q_{min}}}|$, then with our k rotations we have more chance of falling above outage than falling below on average. One can think of this as a dart game, as the target gets bigger the chances of hitting it increases. Note that, with $C_{\mathbf{Q_{max}}} \approx R$, we can almost be certain that the rotations will not be helpful, as it is almost sure that the particular rotation $\mathbf{Q_{max}}$ is not within our original set of rotations. This is depicted in figure 4-1. The pie chart on the left represents the scenario where the channel is above the outage rate only in

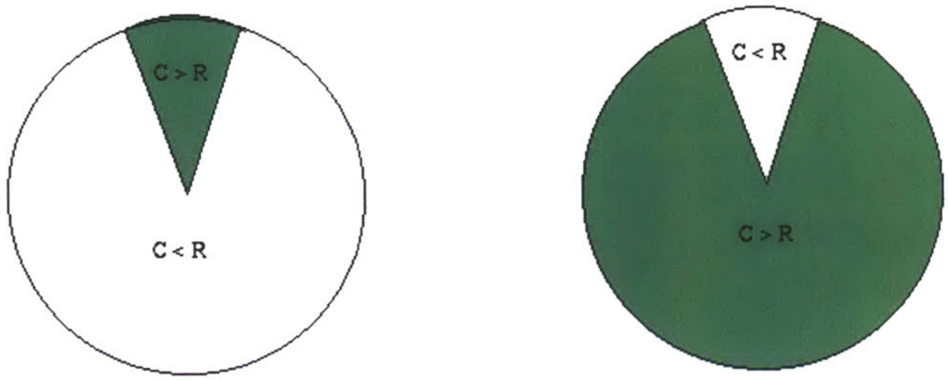

Figure 4-1: Depiction of outage event on rotation space

a small portion of the rotation space. With a uniformly distributed rotation set, we are unlikely to improve under these conditions. The chart on the right represents the opposite scenario where the channel is above outage rate almost in all of the rotation space. Under such conditions, almost of the rotations will bring the channel out of outage.

4.1 Decorrelator

Now we turn to the decorrelator. For the decorrelator, when we say that a channel realization is in outage, we mean that the data rate rate supported by the channel with the sub-optimal decorrelator transceiver is less than some target rate R. For rotations to help increase the rate supported, the channels belonging to virtual users must almost be parallel to each other. A rotation which makes these channels less parallel will surely improve the rate supported by the channel. The reasoning follows from the fact that maximizing rotation $\mathbf{Q_{max}}$ makes the channel look like n orthogonal channels. Hence if the channels are not orthogonal to begin with, using the right set of rotations we can improve.

For channels that are perfectly parallel, the gain would be the most, as then most of the rotations will make them non-parallel, and there is some room for improvement.

The amount of improvement depends on how good a channel realization we are facing. If $C_{\mathbf{Q}_{\max}} \approx R$, then our rotations will most likely not help, as each of the terms making up the average capacity can only get as large as R, and with a well-chosen rotation set we would expect each rotation to give a different rate. With an evenly spaced rotation set, the aim is to hit the channel realizations with their unrotated capacity somewhere between $C_{\mathbf{Q}_{\min}}$ and $C_{\mathbf{Q}_{\max}}$. Therefore, by bringing such channel realizations out of outage, we reduce the number of channel realizations that are in outage with the original channel.

For channel realizations that have their columns perfectly orthogonal, rotations will most likely hurt. The maximizing rotation in such cases is identity matrix, and therefore almost all the rotations within our set will decrease individually, and naturally, the average rate of the block will also decrease. However, this is not performance reducing in terms of outage probability. Since our goal is to get a channel realization, which is already in outage, out of outage, we do not affect the number of channel realizations that are in outage in. Channel realizations, with orthogonal columns, that were already in outage, stay in outage.

In the earlier sections we argued that rotations can improve some channel realizations. Following the same logic, one can argue that rotations can hurt in the same way. However, even though this statement is true for some channel realizations, it is a secondary effect in terms of the number of channel realizations that fall into this category. These are the channel realizations for which $\mathbf{Q}_{\max}$ is identity, and $C_{\mathbf{Q}_{\max}} \approx R$. The probability of facing such a channel realization is low compared to the probability of facing a channel realization, whose columns are almost parallel. Hence, even though, we lose some channel realizations to outage with rotations, the number of realizations we gain is much larger.

This is illustrated in the scatter plot of figure 4-2. The simulation is carried out in MATLAB, with 2×2 channel matrix H having all real entries that are i.i.d Gaussian with unit variance. The covariance matrix of noise was also assumed to be identity. The x-axis on the plot, is the angle between columns of H (in degrees), and y-axis represents the capacity of a 4-rotation system. Each of the stars and diamonds

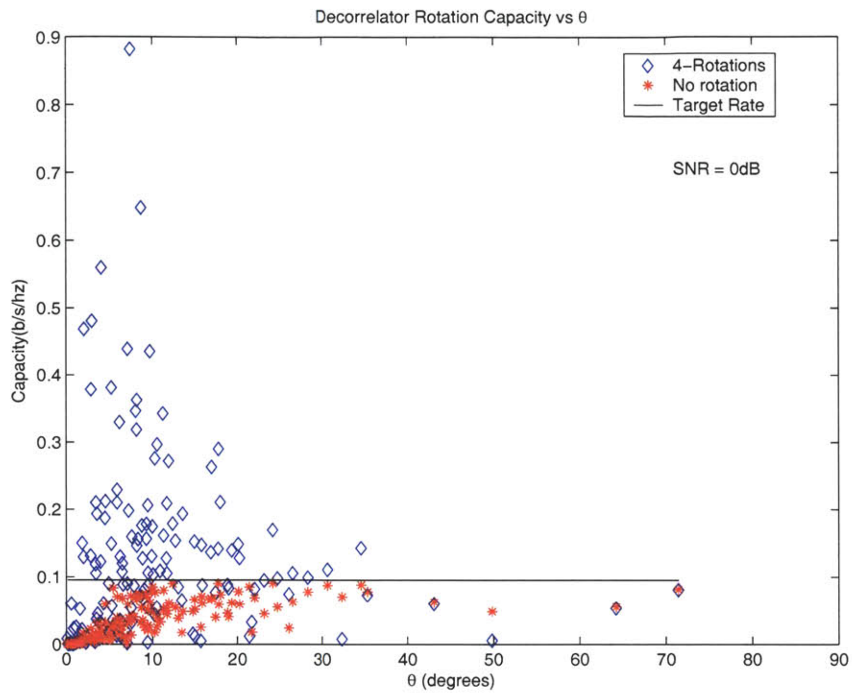

Figure 4-2: Decorrelator rotation capacity versus θ

represents a channel realization, with the x-coordinate representing the angle between the columns of H prior to rotation. The stars represent the capacity of the channel realization prior to rotation, and diamonds represent the capacity after 4-rotations and averaging. The black line represent the target data rate R. As it can be seen, most of the channel realizations that are brought above the target rate are those that have their columns almost parallel to each other. There is almost no improvement when the columns are perpendicular.

4.2 MMSE

The MMSE transceiver maximizes SINR ratio for each user. It takes into account how strong the interference from other streams are, and if the interference is not strong it does not ignore the signal content in the direction of interference. For the decorrelator transceiver the relative directions of the signals belonging to each user is the most important channel realization characteristic determining the channel capacity. However, for the MMSE transceiver this effect is not as significant, as the MMSE also checks how strong of an interference is faced by each stream prior to nulling operation. Therefore, although rotations have some impact on the channel capacity of a realization, through changing the relative directions of streams, this effect is not as significant as the one we observe in the decorrelator. Hence the improvement we see on outage probability due to rotations and averaging is not as significant in the MMSE transceiver, since the channel gains do not show much variation for different rotations. The MMSE extracts almost all the gain from the channel realization. The rotations are merely providing some of the gain lost due to the successive interference cancellation. As we have seen MMSE-SIC achieves the mutual information of the channel for every realization of the channel matrix, hence rotations cannot change the capacity of sub-blocks at all.

As the SINR for each sub-stream increases, the effect of rotations become more important for the MMSE transceiver. In general, as $\mathsf{SNR} \to \infty$, the MMSE starts behaving more like the decorrelator, and hence all the effects described in the previous section about the decorrelator applies to the MMSE. As we lower the SINR the gain we can get from using an n-rotation system decreases, and this is due to the MMSE transceiver already doing a good job in extracting almost all the gain from the channel realization.

This is illustrated in the scatter plot of figure 4-3. The simulation is carried out in MATLAB, with 2×2 channel matrix H having all real entries that are i.i.d Gaussian with unit variance. The covariance matrix of noise was also assumed to be identity. The x-axis on the plot, is the angle between two columns of H (in degrees), and y-axis represents the capacity of the system. The black line represent the target data rate R. As it can be seen, rotations do not improve many channel realizations for MMSE at low SNR. The scatter plots of figure 4-4, show the effect of increasing SNR per sub-stream. As it can be seen, the MMSE scatter plots start looking more like the scatter plots of the decorrelator as SNR is increased.

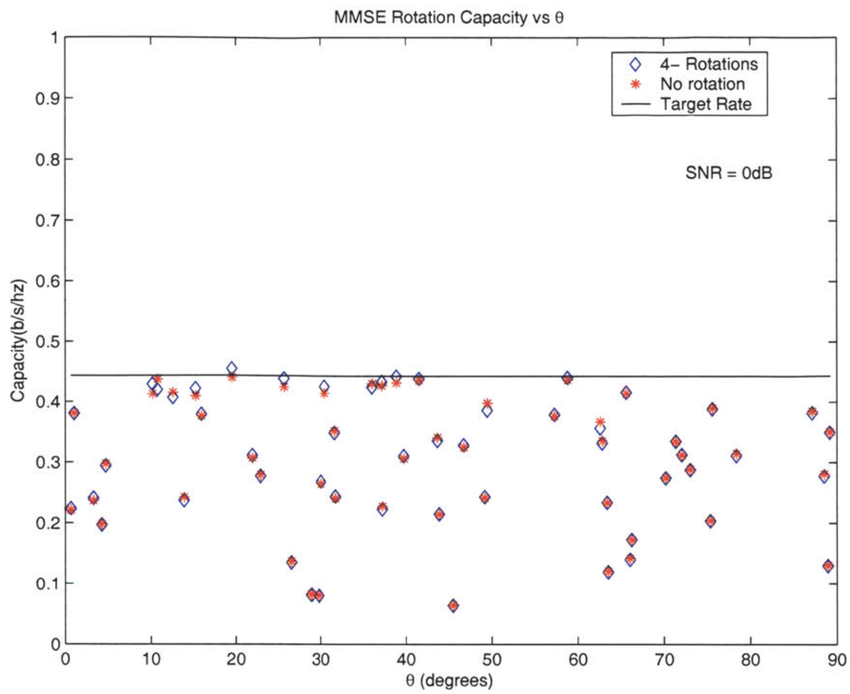

Figure 4-3: MMSE rotation capacity versus θ

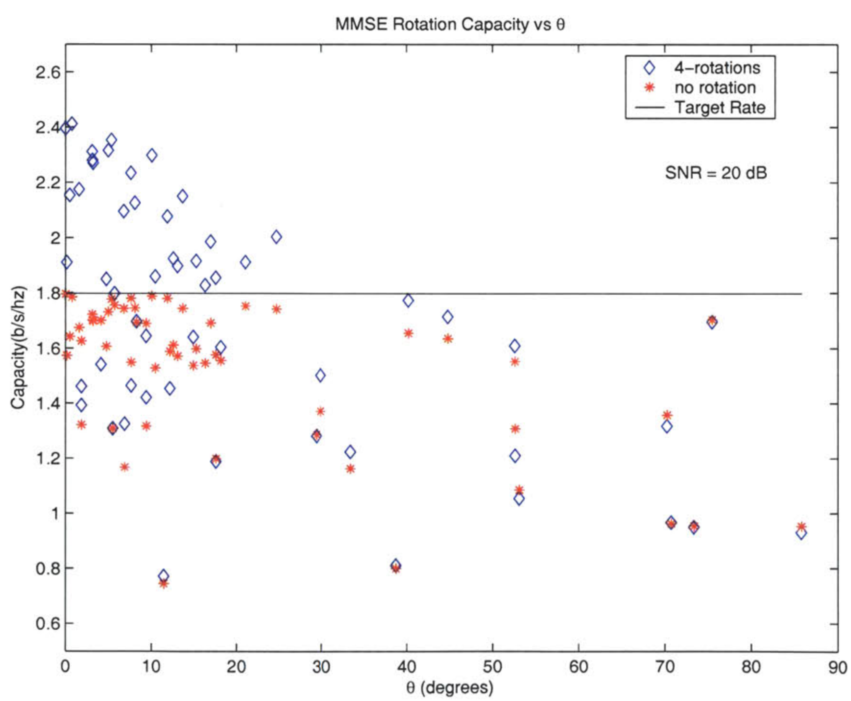

Figure 4-4: MMSE rotation capacity versus θ (varying SNR)

Chapter 5

Conclusions

We introduced a scheme that reduces the probability of outage for sub-optimal tranceiver architectures in slow fading channels. The scheme we describe uses unitary rotations combined with coding over correlated channel replicas in order to decrease the outage probability. The exact characterization of the amount of diversity gain is difficult due to the correlated channel gains. We provided simple approximations in order to quantify this gain.

We also explained how a rotation system helps in the case of the decorrelator and the MMSE tranceivers. In the case of the decorrelator, the gain achieved by the rotation system is higher compared to the MMSE.

Appendix A

Decorrelator Diversity

A.1 Diversity: Sub-block coding

As described in chapter 3, if we code over N sub-blocks the channel will only be in outage if the overall sum capacity of N sub-blocks is less than some rate NR (the rate is normalized here since we are transmitting N times). The average capacity of a sub-block is then:

$$C_m = \frac{1}{N}\sum_{i=1}^{N}\sum_{i=1}^{n}\log\left(1+\mathsf{SNR}\mathbf{g}_i^2\right)$$

Assuming that the channel gains are independent for different sub-blocks, we can reduce the capacity expression to a single summation. The outage probability is then $P\left(\frac{1}{N}\sum_{i=1}^{Nn}\log\left(1+\mathsf{SNR}\mathbf{g}_i^2\right)<R\right)$ with $R = r\log\mathsf{SNR}$. We have already calculated the outage probability of such a channel in chapter 2. The outage probability is given by 2.8:

$$P_{outage}^{N} = \left(\frac{\ln 2}{2\mathsf{SNR}}\right)^{Nn}\frac{(Nr)^{Nn-1}\mathsf{SNR}^{Nr}}{\ln\mathsf{SNR}}$$

and hence the maximal diversity order reached is Nn.

Appendix B

Optimal Receivers with Rotations

To see that the capacity of a MIMO channel using an ideal receiver architecture is invariant to rotations or permutations when the channel matrix H is circularly symmetric, we note that the mutual information between input and output of the MIMO channel can be written as :

$$I(\mathbf{x};\mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x})$$

assuming $\mathbf{x}$ is zero-mean and satisfies $\mathcal{E}[\mathbf{x}^\dagger\mathbf{x}] \leq P$ and that $\mathcal{E}[\mathbf{x}\mathbf{x}^\dagger] = Q$, then $\mathbf{y}$ is zero-mean with covariance $\mathcal{E}[\mathbf{y}\mathbf{y}^\dagger] = HQH^\dagger + I_{n_r}$. The choice of $\mathbf{y}$ that maximizes the entropy is circularly symmetric complex Gaussian [3] . The mutual information is then given by:

$$I(\mathbf{x};\mathbf{y}) = \log\left(\det(I_{n_r} + \mathsf{H}Q\mathsf{H}^\dagger)\right) \tag{B.1}$$

This is the channel capacity for a $n_r \times n_t$ MIMO channel with channel matrix H. For general distributions of H, we do not know the optimal distribution of Q. However, it turns out for *isotropic* H, the optimal distribution of Q turns out to be identity $Q = \frac{P}{n_t} I_{n_t}$ in the lack of CSI. So the mutual information is thus $\log\det(I + \frac{P}{n_t}\mathsf{H}\mathsf{H}^\dagger)$. Now to prove that rotations cannot improve performance for optimal receivers for Rayleigh fading channels, it suffices to note that $(\mathsf{H}\Pi)(\mathsf{H}\Pi)^\dagger = \mathsf{H}\mathsf{H}^\dagger$ for any unitary matrix Π. This also makes sense from an intuitive point of view. Since the channel is symmetric in each direction and since we do not know the realization of channel, there is no benefit one can expect from rotating the channel in any particular direction. So any scheme that achieves $\log(\det(I + \frac{P}{n_t}\mathsf{H}\mathsf{H}^\dagger))$ for each realization of H, cannot benefit in any way from rotating the input vector. In fact, such schemes are called outage optimal, since they can reach the full diversity advantage provided by the MIMO channel. For other distributions of H this might or might not be true.

Bibliography

[1] G. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[2] G. Strang. *Introduction to Linear Algebra*, chapter 4.3. Wellesley-Cambridge Press, Box 812060, Wellesley, MA, 1998.

[3] I. E. Telatar. Capacity of multi-antenna gaussian channels. *European Trans. on Telecommunications*, 10:585–595, November 1999.

[4] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*, chapter 4.3. Preprint, 2003.

[5] P. Wolniansky, G. J. Foschini, G. D. Golden, and R. A. Valenzuela. V-BLAST: An architecture for realizing very high data rates over the rich scattering wireless channel. In *Proc. ISSSE-98*, 1988.

[6] L. Zheng and D. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *IEEE Trans. on Info. Theory*, 49(5):1073–1096, May 2003.