Multiple Antennas in Wireless Communications: Array Signal Processing and Channel Capacity

by

Mahesh Godavarti

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2001

Doctoral Committee:

Professor Alfred O. Hero, III, Chair
Associate Professor Kamal Sarabandi
Professor Wayne E. Stark
Associate Professor Kim Winick

ABSTRACT

Multiple Antennas in Wireless Communications: Array Signal Processing and Channel Capacity

by
Mahesh Godavarti

Chair: Alfred O. Hero, III

We investigate two aspects of multiple-antenna wireless communication systems in this thesis: 1) deployment of an adaptive beamformer array at the receiver; and 2) space-time coding for arrays at the transmitter and the receiver. In the first part of the thesis, we establish sufficient conditions for the convergence of a popular least mean squares (LMS) algorithm known as the sequential Partial Update LMS Algorithm for adaptive beamforming. Partial update LMS (PU-LMS) algorithms are reduced complexity versions of the full update LMS that update a subset of filter coefficients at each iteration. We introduce a new improved algorithm, called Stochastic PU-LMS, which selects the subsets at random at each iteration. We show that the new algorithm converges for a wider class of signals than the existing PU-LMS algorithms.

The second part of this thesis deals with the multiple-input multiple-output (MIMO) Shannon capacity of multiple antenna wireless communication systems under the average energy constraint on the input signal. Previous work on this problem has concentrated on capacity for Rayleigh fading channels. We investigate the more general case of Rician fading. We derive capacity expressions, optimum transmit signals as well as upper and lower bounds on capacity for three Rician fading models. In the first model the specular component is a dynamic isotropically distributed random process. In this case, the optimum transmit signal structure is the same as that for Rayleigh fading. In the second model the specular component is a static isotropically distributed random process unknown to the transmitter, but known to the receiver. In this case the transmitter has to design the transmit signal to guarantee a certain rate independent of the specular component. Here also, the optimum transmit signal structure, under the constant magnitude constraint, is the same as that for Rayleigh fading. In the third model the specular component is deterministic and known to both the transmitter and the receiver. In this case the optimum transmit signal and capacity both depend on the specular component. We show that for low signal to noise ratio (SNR) the specular component completely determines the the signal structure whereas for high SNR the specular component has no effect. We also show that training is not effective at low SNR and give expressions for rate-optimal allocation of training versus communication.

To my family and friends

ACKNOWLEDGEMENTS

I would like to extend my sincere thanks to my advisor Prof. Alfred O. Hero-III for letting me dictate the pace of my research, giving me freedom to a large extent on the choice of topics, his invaluable guidance and encouragement, and for giving me invaluable insight into my research.

I would also like to extend my thanks to Dr. Thomas Marzetta at Bell Labs, Murray Hill, NJ for supervising my summer internship and helping me develop a significant part of my thesis.

I am grateful to my committee members, Professor Wayne Stark, Associate Professor Kim Winick and Associate Professor Kamal Sarabandi for their valuable inputs. I am also thankful to them, especially Dr. Sarabandi, for being gracious enough to accommodate me and making the final oral defense possible on the day of my choice.

I would like to thank my friends, Olgica Milenkovic, Dejan Filipovic, Tara Javidi, Victoria Yee, Riten Gupta, Robinson Piramuthu, Bing Ma, Tingfang Ji, John Choi, Ketan Patel, Sungill Kim, Selin Aviyente, Jia Li, Ali Ozgur and Rajiv Vijayakumar for making my experiences in the department pleasant; Zoya Khan and Madhumati Ramesh for being great neighbours; Sivakumar Santhanakrishnan for being an understanding roommate; and Bhavani Raman and Muralikrishna Datla for making my experience in SPICMACAY rewarding.

A special thanks to Navin Kashyap for being a sounding board.

My sincere thanks to Department of Defense Research & Engineering (DDR&E)

Multidisciplinary University Research Initiative (MURI) on "Low Energy Electronics Design for Mobile Platform" managed by the Army Research Office (ARO) for supporting this research through ARO grant DAAH04-96-1-0337. My thanks to the Department of Electrical Engineering as well for supporting this research through various fellowships and Assistantships.

Finally, I would like to thank my parents, brother and sister who have always encouraged me in my quest for higher education and especially my wife and friend of many years Kavita Raman for encouraging me and listening to my occasional laments and for supporting my many decisions.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

CHAPTER 1

Introduction

This thesis deals with theory underlying the deployment of multiple antennas, i.e. antenna arrays, at transmitter and receiver for the purpose of improved communication, reliability and performance. The thesis can be divided into two main parts. The first part deals with conditions for convergence of adaptive receiver arrays and a new reduced complexity beamformer algorithm. The second part deals with channel capacity for a Rician fading multiple-input multiple-output (MIMO) channel. More details are given in the rest of this chapter.

Wireless communications have been gaining popularity because of better antenna technologies, lower costs, easier deployment of wireless systems, greater flexibility, better reliability and the need for mobile communication. In some cases, like in very remote areas, wireless connections may be the only option.

Even though the popularity of mobile wireless telephony and paging is a recent phenomenon, fixed-wireless systems have a long history. Point-to-point microwave connections have long been used for voice and data communications, generally in backhaul networks operated by phone companies, cable TV companies, utilities, railways, paging companies and government agencies, and will continue to be an important part of the communications infrastructure. Improvements in technology have allowed higher frequencies and thus smaller antennas to be used resulting in lower costs and easier-to-deploy systems.

Another reason for the popularity of wireless systems is that consumers demand for data rates has been insatiable. Wireline models have topped off at a rate of 56Kbps and end-users have been looking for integrated digital subscriber network (ISDN) and digital subscriber line (DSL) connections. Companies with T1 connections of 1.54Mbps have found the connections inadequate and are turning to T3 optical fiber connections. The very expensive deployment of fiber connections, however, has caused companies to turn to fixed wireless links.

This has resulted in the application of wireless communications to a host of applications ranging from: fixed microwave links; wireless local area networks (LANs); data over cellular networks; wireless wide area networks (WANs); satellite links; digital dispatch networks; one-way and two-way paging networks; diffuse infrared; laser-based communications; keyless car entry; the Global Positioning System (GPS); mobile cellular communications; and indoor-radio.

One challenge in wireless systems not present in wireline systems is the issue of fading. Fading arises due to the possible existence of multiple paths from the transmitter to the receiver with destructive combination at the receiver output. There are many models describing fading in wireless channels [20]. The classic models being Rayleigh and Rician flat fading models. Rayleigh and Rician models are typically applied to narrowband signals and do not include the doppler shift induced by the motion of the transmitter or the receiver.

In wireless systems, there are three different ways to combat fading: 1) frequency diversity; 2) time diversity; and 3) spatial diversity. Frequency diversity makes use of the fact that multipath structure in different frequency bands is different. This fact can be exploited to mitigate the effect of fading. But, the positive effects of frequency diversity are limited due to bandwidth limitations. Wireless communication uses radio spectrum, a finite resource. This limits the number of wireless users and the amount of spectrum available to any user at any moment in time. Time diversity makes use of the fact that fading over different time intervals is different. By using channel coding the effect of bad fading intervals can be mitigated by good fading intervals. However, due to delay constraints time diversity is difficult to exploit.

Spatial diversity exploits multiple antennas either separated in space or differently polarized [7, 23, 24]. Different antennas see different multipath characteristics or different fading characteristics and this can be used to generate a stronger signal. Spatial diversity techniques do not have the drawbacks associated with time diversity and frequency diversity techniques. The one drawback of spatial diversity is that it involves deployment of multiple antennas at the transmitter and the receiver which is not always feasible.

In this thesis, we will concentrate on spatial diversity resulting from deployment of multiple antennas. Spatial diversity, at the receiver (multiple antennas at the receiver) or at the transmitter (multiple antennas at the transmitter), can improve link performance in the following ways [31]

1. Improvements in spectrum efficiency: Multiple antennas can be used to accommodate more than one user in a given spectral bandwidth.

2. Extension of range coverage: Multiple antennas can be used to direct the energy of a signal in a given direction and hence minimize leakage of signal energy.

3. Tracking of multiple mobiles: The outputs of antennas can be combined in different ways to isolate signals from each and every mobile.

4. Increases in channel reuse: Improving spectral efficiency can allow more than one user to operate in a cell.

5. Reductions in power usage: By directing the energy in a certain direction and increasing range coverage lesser energy can be used to reach a user at a given distance.

6. Generation of multiple access: Appropriately combining the outputs of the antennas can selectively provide access to users.

7. Reduction of co-channel interference

8. Combating of fading

9. Increase in information channel capacity: Multiple antennas have been used to increase the maximum achievable data rates.

Traditionally, all the gains listed above have been realized by explicitly directing the receive or transmit antenna array to point in specific directions. This process is called beamforming. For receive antennas, beamforming can be achieved electronically by appropriately weighting the antenna outputs and combining them to make the antenna response to energy emanating from certain directions more sensitive than others. Until recently most of the research on antenna arrays for beamforming has dealt with beamformers at the receiver. Transmit beamformers behave differently and require different algorithms and hardware [32].

Methods of beamforming at the receive antenna array currently in use are based on array processing algorithms for signal copy, direction finding and signal separation [32]. These include Applebaum/Frost beamforming, null steering beamforming, optimal beamforming, beam-space Processing, blind beamforming, optimum combining and maximal ratio combining [15, 32, 64, 68, 69, 74, 80]. Many of these beamformers require a reference signal and use adaptive algorithms to optimize beamformer weights with respect to some beamforming performance criterion [32, 78, 79, 81].

Adaptive algorithms can also be used without training for tracking a time varying mobile user, tracking multiple users, or tracking time varying channels. Popular examples [32] are the Least Mean Squares Algorithm (LMS), Constant Modulus Algorithm (CMA) and the Recursive Least Squares (RLS) algorithm. The algorithm of interest in this work is the LMS Algorithm because of its ease of implementation and low complexity.

Another research topic in the field of beamforming that has generated much interest is the effect of calibration errors in direction finding and signal copy problems [25, 49, 65, 66, 67, 83, 84]. An array with Gaussian calibration errors operating in a non-fading environment has the same model as a Rician fading channel. Thus the work done in this thesis can be easily translated to the case of array calibration errors.

Beamforming at the receiver is one way of exploiting receive diversity. Most previous work (1995 and earlier) in the literature concentrates on this kind of diversity. Another way to exploit diversity is to perform beamforming at the transmitter, i.e. transmit diversity. Beamforming at the transmitter increases the signal to noise ratio (SNR) at the receiver by focusing the transmit energy in the directions that ensures the strongest reception at the receiver. Exploitation of transmit diversity can involve [71] using the channel state information obtained via feedback for reassigning energy at different antennas via waterpouring, linear processing of signals to spread the information across transmit antennas and using channel codes and transmitting the codes using different antennas in an orthogonal manner.

An early use of the multiple transmit antennas was to obtain diversity gains by sending multiple copies of a signal over orthogonal time or frequency slices (repetition code). This of course, incurs a bandwidth expansion factor equal to the number of antennas. A transmit diversity technique without bandwidth expansion was first suggested by Wittenben [82]. Wittenben's diversity technique of sending time-delayed copies of a common input signal over transmit multiple antennas was also independently discovered by Seshadri and Winters [58] and by Weerackody [77]. An information theoretic approach to designing transmit diversity schemes was undertaken by Narula [53, 54]. The authors design schemes that maximize the mutual information between the transmitter and the receiver.

These methods correspond to beamforming where knowledge of the channel is available at the transmitter and receiver, for example by training and feedback. In such a case the strategy is to mitigate the effect of multipath by spatial diversity and focusing the channel to a single equivalent direct path channel by beamforming.

Recently, researchers have realized that beamforming may not be the optimal way to increase data rates. The BLAST project showed that multipaths are not as harmful as previously thought and that the multiple diversity can be exploited to increase capacity even when the channel is unknown [23, 50]. This has given rise to research on space-time codes [8, 42, 43, 47, 48, 52, 70, 71]. Space-time coding is a coding technique that is designed for use with multiple transmit antennas. One of the first papers in this area is by Alamouti [5] who designed a simple scheme for a two-antenna transmit system. The codes are designed to induce spatial and temporal correlations into signals that are robust to unknown channel variations and can be exploited at the receiver. Space-time codes are simply a systematic way to perform beneficial space-time processing of signals before transmission [52].

Design of space-time codes has taken many forms. Tarokh et. al. [70, 71] have taken an approach to designing space-time codes for both Rayleigh and Rician fading channels with complete channel state information at the receiver that maximizes a pairwise codeword distance criterion. The pairwise distance criterion was derived from an upper bound on probability of decoding error. There have also been code designs where the receiver has no knowledge about the MIMO channel. Hero and Marzetta [42] design space-time codes with a design criterion of maximizing the cut-off rate for the MIMO Rayleigh fading channel. Hochwald et. al [43, 44] propose a design based on signal structures that asymptotically achieve capacity in the non-coherent case for MIMO Rayleigh fading channels. Hughes [47, 48] considered the design of space-time based on the concept of Group codes. The codes can be viewed as an extended version of phase shift keying for the case of multiple antenna communications. In [47] the author independently proposed a scheme similar to that proposed by Hochwald and Marzetta in [43]. More recent work in this area has been by Hassibi on linear dispersion codes [38] and fixed-point free codes [40] and by Shokrollahi on double diagonal space-time codes [61] and unitary space-time codes [62].

The research reported in this dissertation concentrates on adaptive beamforming receivers for fixed deterministic channels and channel capacity of multiple antennas in the presence of Rician fading. We will elaborate more on the research contributions in the following sections.

1.1 Partial Update LMS Algorithms

The LMS algorithm is a popular algorithm for adaptation of weights in adaptive beamformers using antenna arrays and for channel equalization to combat intersymbol interference. Many others application areas of LMS include interference cancellation, echo cancellation, space time modulation and coding and signal copy in surveillance. Although there exist algorithms with faster convergence rates like RLS, LMS is very popular because of ease of implementation and low computational costs.

One of the variants of LMS is the Partial Update LMS (PU-LMS) Algorithm. Some of the applications in wireless communications like channel equalization and echo cancellation require the adaptive filter to have a very large number of coefficients. Updating of the entire coefficient set might be beyond the ability of the mobile units. Therefore, partial updating of the LMS adaptive filter has been proposed to further reduce computational costs [30, 33, 51]. In this era of mobile computing and communications, such implementations are also attractive for reducing power consumption. However, theoretical performance predictions on convergence rate and steady state tracking error are more difficult to derive than for standard full update LMS. Accurate theoretical predictions are important as it has been observed that the standard LMS conditions on the step size parameter fail to ensure convergence of the partial update algorithm.

Two of the partial update algorithms prevalent in the literature have been described in [18]. They are referred to as the "Periodic LMS algorithm" and the "Sequential LMS algorithm". To reduce computation by a factor of P, the Periodic LMS algorithm (P-LMS) updates all the filter coefficients every P^{th} iteration instead of every iteration. The Sequential LMS (S-LMS) algorithm updates only a fraction of coefficients every iteration.

Another variant referred to as "Max Partial Update LMS algorithm" (Max PU-LMS) has been proposed in [16, 17] and [3]. In this algorithm, the subset of coefficients to be updated is dependent on the input signal. The subset is chosen so as to minimize the increase in the mean squared error due to partial as opposed to full updating. The input signals multiplying each coefficient are ordered according to their magnitude and the coefficients corresponding to the largest $\frac{1}{P}$ of input signals are chosen for update in an iteration. Some analysis of this algorithm has been done in [17] for the special case of one coefficient per iteration but, analysis for more general cases still needs to be completed. The results on stochastic updating in Chapter 3 provide a small step in this direction.

1.2 Multiple-Antenna Capacity

Shannon in his famous paper [59] showed that it is possible to communicate over a noisy channel with arbitrary reliability provided that the amount of information communicated (bits/channel use) is less than a constant. This constant is known as the channel capacity. Shannon showed that the channel capacity can be computed by maximizing the mutual information between the input and the output over all possible input distributions. The channel capacity for a range of channels like the binary symmetric channel, the additive white Gaussian noise (AWGN) channel have already been computed in the literature [13, 26]. Computing the capacity for more complicated channels like Rayleigh fading and Rician fading channels is in general a difficult problem.

The seminal paper by Foschini et. al. [23, 24] showed that a significant gain in capacity can be achieved by using multiple antennas in the presence of Rayleigh fading. Let M be the number of antennas at the transmitter and N be the number of antennas at the receiver. Foschini and Telatar showed [73] that with perfect channel knowledge at the receiver, for high SNR a capacity gain of $\min(M, N)$ bits/second/Hz can be achieved with every 3 dB increase in SNR. Channel knowledge at the receiver however requires that the time between different fades be sufficiently large to enable the receiver to learn the channel via training. This might not be true in the case of fast mobile receivers and large numbers of transmit antennas. Furthermore, the use of training is an overhead which reduces the attainable capacity.

Following Foschini [23], there have been many papers written on the subject of calculating capacity for a MIMO channel [7, 10, 12, 28, 29, 34, 35, 50, 60, 72, 76]. Others have studied the achievable rate regions for the MIMO channel in terms of cut-off rate [42] and error exponents [1].

Marzetta and Hochwald [50] considered a Rayleigh fading MIMO channel when neither the receiver nor the transmitter has any knowledge of the fading coefficients. In their model the fading coefficients remain constant for T symbol periods and instantaneously change to new independent complex Gaussian realizations every T symbol periods. They established that to achieve capacity it is sufficient to use $M = T$ antennas at the transmitter and that the capacity achieving signal matrix consists of a product of two independent matrices, a $T \times T$ isotropically random unitary matrix and a $T \times M$ real nonnegative diagonal matrix. Hence, it is sufficient to optimize over the density of a smaller parameter set of size $\min\{M, T\}$ instead of the original parameter set of size $T \cdot M$.

Zheng and Tse [85] derived explicit capacity results for the case of high SNR in the case of no channel knowledge at the transmitter or receiver. They showed that the number of degrees of freedom for non-coherent communication is $M^*(1 - M^*/T)$ where $M^* = \min\{M, N, T/2\}$ as opposed to $\min\{M, N\}$ in the case of coherent communications.

The literature cited above has limited its attention to Rayleigh fading channel models for computing capacity of multiple-antenna wireless links. However, Rayleigh fading models are inadequate in describing the many fading channels encountered in practice. Another popular model used in the literature to fill this gap is the Rician fading channel. Rician fading is a more accurate model when there are some direct paths present between the transmitter and the receiver along with the diffuse multipath (Figure 1.1). Rician fading components traditionally have been modeled

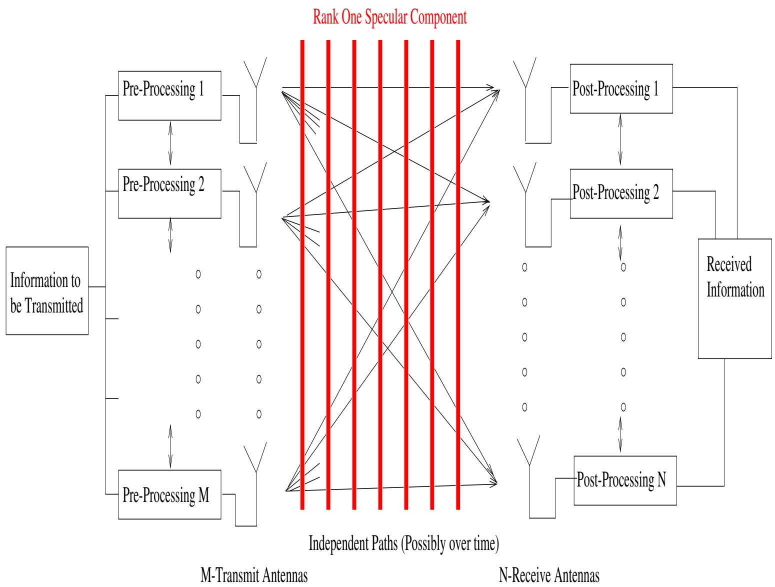


Figure 1.1: Diagram of a multiple antenna communication system

as independent Gaussian components with a deterministic non-zero mean [9, 19, 21, 56, 57, 71]. Farrokhi et. al. [21] used this model to analyze the capacity of a MIMO channel with a single specular component. In their paper they assumed that the specular component is static and unknown to the transmitter but known to the receiver. They also assumed that the receiver has complete knowledge about the fading coefficients (i.e. the Rayleigh and specular components are completely known). They work with the premise that since the transmitter has no knowledge about the specular component the signaling scheme has to be designed to guarantee a given rate irrespective of the value of the specular component. They conclude that the signal matrix has to be composed of independent circular Gaussian random variables of mean 0 and equal variance in order to maximize the rate and achieve capacity.

1.3 Organization of the Dissertation and Significant Contributions

In this work, we have made the following contributions. These contributions are divided into two fields: 1) LMS algorithm convergence for adaptive arrays at the receiver; and 2) evaluation of Shannon capacity for multiple antennas at the transmitter and the receiver in the presence of Rician fading.

1. In Chapter 2 we analyze the Sequential PU-LMS for stability and come up with more stringent conditions on stability than were previously known. We illustrate our findings via simulations.

   - **Contributions**: Derived conditions ensuring the stability of the Sequential PU-LMS algorithm for stationary signals without the restrictive assumptions of [18]. The analysis of the algorithm for cyclo-stationary signals establishes that the deterministic sequences of updates is the reason behind the algorithm's poor convergence. This motivates a new Stochastic PU-LMS, a more stable algorithm.

2. Chapter 3 analyzes the Stochastic Partial Update LMS algorithm where the coefficients to be updated in an iteration are chosen at random. This generalizes the previous PU-LMS methods. We derive conditions for stability and also analyze the algorithm for performance. We demonstrate the effectiveness of our analysis via simulations.

   - **Contributions**: Proposed a new Partial Update algorithm with better convergence properties than those of existing Partial Update LMS algorithms. The convergence of Stochastic PU-LMS is better than the existing PU-LMS algorithms for the case of non-stationary signals and similar to the existing algorithms for the case of stationary signals. The analysis elucidates the role of parameters which determine the convergence or divergence of PU-LMS algorithms.

Contributions reported in the subsequent chapters have been in the area of computing capacity for a more general fading model than the Rayleigh model. This we consider is the first significant step towards computing capacities for more realistic models for MIMO systems.

3. In Chapter 4, we introduce a MIMO Rician fading where the specular component is also modeled as dynamic and random but, with an isotropically uniform density [50]. With this model the channel capacity can be easily characterized. We also derive a lower bound to capacity which is useful to establish achievable rate regions as the calculation of the exact capacity is difficult even for this model.

    - **Contributions**: Proposed a new tractable model for analysis enabling characterization of MIMO capacity achieving signals and also derived a useful lower bound on channel capacity for Rician fdaing. This bound is also applicable to the case of Rayleigh fading. Showed that the optimum signal structure for Rician fading is the same as that of Rayleigh fading channel. Therefore the space-time codes developed so far can be used directly for the model described above.

4. In Chapter 5, we study MIMO capacity for the case of static and constant

(persistent) specular component. In this case the channel is non-ergodic and the channel capacity is not defined. We therefore maximize the worst possible rate available for communication over the ensemble of values of the specular component under a constant specular norm constraint. This rate is the *min-capacity.*

- **Contributions**: Proposed a tractable formulation of the problem and derived capacity expressions, lower bound on capacity and characterized the properties of capacity achieving signals. The results show that a large class of space-time codes developed so far for MIMO Rayleigh fading channel can be directly applied to Rician fading with a persistent isotropic specular component.

5. In Chapter 6, we evaluate MIMO capacity for the same Rician model as in Chapter 5 but we assume that both the transmitter and the receiver have complete knowledge concerning the specular component. In this case, the channel is ergodic and the channel capacity in terms of Shannon theory is well defined.

   - **Contributions**: Derived coherent and non-coherent capacity expressions in the low and high SNR regimes for the standard Rician fading model. The analysis shows that for low SNR the optimal signaling is beamforming whereas for high SNR it is diversity signaling. For low SNR we demonstrated that the Rician channel provides as much capacity as an AWGN channel. Also, characterized the optimum training signal, training duration and power allocation for training in the case of the standard Rician fading model. Established that for low SNR, training is not required.

6. In Chapter 7, we introduce rigorous definitions for two quantities of interest,

diversity and degrees of freedom, that are used to quantify the advantages of a multiple antenna MIMO system when compared to a single input single output (SISO) system. We verify the effectiveness of the definitions by computing the quantities of interest for various existing examples.

- **Contributions**: Gave an intuitive interpretation for diversity and degrees of freedom which helps in qualifying the advantages of a MIMO system. Provided rigorous definitions for the quantities of interest in a more general setting which will allow computation of these quantities for systems other than multiple antenna MIMO systems.

CHAPTER 2

Sequential Partial Update LMS Algorithm

2.1 Introduction

The least mean-squares (LMS) algorithm is an approximation of the steepest descent algorithm used to arrive at the Weiner-Hopf solution for computing the weights (filter coefficients) of a finite impulse response (FIR) filter. The filter coefficients are computed so as to produce the closest approximation in terms of mean squared error to a *desired output*, which is stochastic in nature from the input to the filter, which is also stochastic in nature. The Weiner-Hopf solution involves an inversion of the input signal correlation matrix. The steepest descent algorithm avoids this inversion by recursively computing the filter coefficients using the gradient computed using the input signal correlation matrix. The LMS algorithm differs from the steepest algorithm in that it uses a "stochastic gradient" as opposed to the exact gradient. Knowledge of the exact input signal correlation matrix is not required for the algorithm to function. The reduction in complexity of the algorithm comes at an expense of greater instability and degraded performance in terms of final mean squared error. Therefore, the issues with the LMS algorithm are "filter stability", "final misadjustment" and "convergence rate" [41, 46, 63].

Partial update LMS algorithms are reduced complexity versions of LMS as described in section 1.1. The gains in complexity reduction arising from updates of only a subset of coefficients at an iteration are significant when there are a large number of weights in the filter. For example, in channel equalization and in fixed "repeater" links with large baseline and large number of array elements.

In [18], a condition for convergence in mean for the Sequential Partial Update LMS (S-LMS) algorithm was derived under the assumption of small step-size parameter (μ). This condition turned out to be the same as that for the standard LMS algorithm for wide sense stationary (W.S.S.) signals. In this chapter, we prove a stronger result: for arbitrary $\mu > 0$, and for W.S.S. signals, convergence in mean of the regular LMS algorithm guarantees convergence in mean of S-LMS.

We also derive bounds on the step-size parameter μ for S-LMS Algorithm which ensures convergence in mean for the special case involving alternate even and odd coefficient updates. The bounds are based on extremal properties of the matrix 2-norm. We derive bounds for the case of stationary and cyclo-stationary signals. For simplicity we make the standard independence assumptions used in the analysis of LMS [6].

The organization of the chapter is as follows. First in section 2.2, a brief description of the sequential partial update algorithm is given. The algorithm with arbitrary sequence of updates is analyzed for the case of stationary signals in section 2.3. This is followed by the analysis of the even-odd update algorithm for cyclo-stationary signals in section 2.4. In section 2.5 an example is given to illustrate the usefulness of the bounds on step-size μ derived in section 2.4. Finally, conclusions and directions for future work are indicated in section 2.6.

2.2 Algorithm Description

The block diagram of S-LMS for a N-tap LMS filter with alternating even and odd coefficient updates is shown in Figure 2.1. We refer to this algorithm as even-odd S-LMS.

It is assumed that the LMS filter is a standard FIR filter of even length, N. For convenience, we start with some definitions. Let $\{x_{i,k}\}$ be the input sequence and let $\{w_{i,k}\}$ denote the coefficients of the adaptive filter. Define

$$\begin{aligned} W_k &= [w_{1,k}\ w_{2,k}\ \ldots\ w_{N,k}]^\tau \\ X_k &= [x_{1,k}\ x_{2,k}\ x_{3,k}\ \ldots\ x_{N,k}]^\tau \end{aligned}$$

where the terms defined above are for the instant k and τ denotes the transpose operator. In addition, Let d_k denote the desired response. In typical applications d_k is a known training signal which is transmitted over a noisy channel with unknown FIR transfer function.

In this paper we assume that d_k itself obeys an FIR model given by $d_k = W_{opt}^\dagger X_k + n_k$ where W_{opt} are the coefficients of an FIR model given by $W_{opt} = [w_{1,opt}\ \ldots\ w_{N,opt}]^\tau$ and $\dagger$ denotes the hermitian operator. Here $\{n_k\}$ is assumed to be a zero mean i.i.d sequence that is independent of the input sequence X_k.

For description purposes we will assume that the filter coefficients can be divided into P mutually exclusive subsets of equal size, i.e. the filter length N is a multiple of P. For convenience, define the index set $S = \{1, 2, \ldots, N\}$. Partition S into P mutually exclusive subsets of equal size, $S_1,\ S_2, \ldots, S_P$. Define I_i by zeroing out the j^{th} row of the identity matrix I if $j \notin S_i$. In that case, $I_i X_k$ will have precisely $\frac{N}{P}$ non-zero entries. Let the sentence "choosing S_i at iteration k" stand to mean "choosing the weights with their indices in S_i for update at iteration k".

The S-LMS algorithm is described as follows. At a given iteration, k, one of the sets S_i, $i = 1, \ldots, P$, is chosen in a pre-determined fashion and the update is performed.

$$w_{k+1,j} = \begin{cases} w_{k,j} + \mu e_k^* x_{k,j} & \text{if } j \in S_i \\ w_{k,j} & \text{otherwise} \end{cases}$$

where $e_k = d_k - W_k^\dagger X_k$. The above update equation can be written in a more compact form in the following manner

$$W_{k+1} = W_k + \mu e_k^* I_i X_k \tag{2.1}$$

In the special case of even and odd updates, $P = 2$ and S_1 consists of all even indices and S_2 of all odd indices as shown in Figure 2.1.

We also define the coefficient error vector as

$$V_k \quad = \quad W_k - W_{opt}$$

which leads to the following coefficient error vector update for S-LMS when k is odd

$$\begin{aligned} V_{k+2} \quad &= \quad (I - \mu I_2 X_{k+1} X_{k+1}^\dagger)(I - \mu I_1 X_k X_k^\dagger) V_k + \\ &\qquad \mu(I - \mu I_2 X_{k+1} X_{k+1}^\dagger) n_k I_1 X_k + \mu n_{k+1} I_2 X_{k+1}, \end{aligned}$$

and the following when k is even

$$\begin{aligned} V_{k+2} \quad &= \quad (I - \mu I_1 X_{k+1} X_{k+1}^\dagger)(I - \mu I_2 X_k X_k^\dagger) V_k + \\ &\qquad \mu(I - \mu I_1 X_{k+1} X_{k+1}^\dagger) n_k I_2 X_k + \mu n_{k+1} I_1 X_{k+1}. \end{aligned}$$

2.3 Analysis: Stationary Signals

Assuming that d_k and X_k are jointly WSS random sequences, we analyze the convergence of the mean coefficient error vector $E[V_k]$. We make the standard assumptions that V_k and X_k are mutually uncorrelated and that X_k is independent of X_{k-1} [6] which is not an unreasonable assumption for the case of antenna arrays. For regular full update LMS algorithm the recursion for $E\left[V_k\right]$ is given by

$$E\left[V_{k+1}\right] = (I - \mu R)E\left[V_k\right] \tag{2.2}$$

where I is the N-dimensional identity matrix and $R = E\left[X_k X_k^\dagger\right]$ is the input signal correlation matrix. The necessary and sufficient condition for stability of the recursion is given by

$$0 < \mu < 2/\lambda_{max}(R) \tag{2.3}$$

where $\lambda_{max}(R)$ is the maximum eigen-value of the input signal correlation matrix R.

Taking expectations under the same assumptions as above, using the independence assumption on the sequences X_k, n_k, the mutual independence assumption on X_k and V_k, and simplifying we obtain for even-odd S-LMS when k is odd

$$E\left[V_{k+2}\right] = (I - \mu I_2 R)(I - \mu I_1 R)E[V_k] \tag{2.4}$$

and when k is even

$$E\left[V_{k+2}\right] = (I - \mu I_1 R)(I - \mu I_2 R)E[V_k]. \tag{2.5}$$

It can be shown that under the above assumptions on X_k, V_k and d_k, the convergence conditions for even and odd update equations are identical. We therefore focus on (2.4). Now to ensure stability of (2.4), the eigenvalues of $(I - \mu I_2 R)(I - \mu I_1 R)$ should lie inside the unit circle. We will show that if the eigenvalues of $I - \mu R$ lie inside the unit circle then so do the eigenvalues of $(I - \mu I_2 R)(I - \mu I_1 R)$.

Now, if instead of just two partitions of even and odd coefficients ($P = 2$) we have any number of arbitrary partitions ($P \geq 2$) then the update equations can be similarly written as above with $P > 2$. Namely,

$$E[V_{k+P}] = \prod_{i=1}^{P}(I - \mu I_{(i+k)\%P} R)E[V_k]$$

where $(i+k)\%P$ stands for $(i+k)$ modulo P. I_i, $i = 1, \ldots, P$ is obtained from I, the identity matrix of dimension $N \times N$, by zeroing out some rows in I such that $\sum_{i=1}^{P} I_i$ is positive definite.

We will show that for any arbitrary partition of any size ($P \geq 2$); S-LMS converges in the mean if LMS converges in the mean(Theorem 2.2). The case $P = 2$ follows as a special case. The intuitive reason behind this fact is that both the algorithms try to minimize the mean squared error $V_k^{\dagger} R V_k$. This error term is a quadratic bowl in the V_k co-ordinate system. Note that LMS moves in the direction of the negative gradient $-RV_k$ by retaining all the components of this gradient in the V_k co-ordinate system whereas S-LMS discards some of the components at every iteration. The resulting gradient vector (the direction in which S-LMS updates its weights) obtained from the remaining components still points towards the bottom of the quadratic bowl and hence if LMS reduces the mean squared error then so does S-LMS.

We will show that if R is a positive definite matrix of dimension $N \times N$ with eigenvalues lying in the open interval $(0, 2)$ then $\prod_{i=1}^{P}(I - I_i R)$ has eigenvalues inside the unit circle.

The following theorem is used in proving the main result in Theorem 2.2.

Theorem 2.1. *[45, Prob. 16, page 410] Let B be an arbitrary $N \times N$ matrix. Then $\rho(B) < 1$ if and only if there exists some positive definite $N \times N$ matrix A such that $A - B^{\dagger} A B$ is positive definite. $\rho(B)$ denotes the spectral radius of B ($\rho(B) = \max_{1,\ldots,N} |\lambda_i(B)|$).*

Theorem 2.2. *Let R be a positive definite matrix of dimension $N \times N$ with $\rho(R) = \lambda_{max}(R) < 2$ then $\rho(\prod_{i=1}^{P}(I - I_i R)) < 1$ where I_i, $i = 1, \ldots, P$ are obtained by zeroing out some rows in the identity matrix I such that $\sum_{i=1}^{P} I_i$ is positive definite. Thus if X_k and d_k are jointly W.S.S. then S-LMS converges in the mean if LMS converges* *in the mean.*

Proof: Let $\mathbf{x}_0 \in \mathbb{C}^N$ be an arbitrary non-zero vector of length N. Let $\mathbf{x}_i = (I - I_i R)\mathbf{x}_{i-1}$. Also, let $\mathbf{P} = \prod_{i=1}^{P}(I - I_i R)$.

First we will show that $\mathbf{x}_i^\dagger R\mathbf{x}_i \leq \mathbf{x}_{i-1}^\dagger R\mathbf{x}_{i-1} - \alpha\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1}$, where $\alpha = \frac{1}{2}(2 - \lambda_{max}(R)) > 0$.

$$\begin{aligned}
\mathbf{x}_i^\dagger R\mathbf{x}_i &= \mathbf{x}_{i-1}^\dagger(I - RI_i)R(I - I_iR)\mathbf{x}_{i-1} \\
&= \mathbf{x}_{i-1}^\dagger R\mathbf{x}_{i-1} - \alpha\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1} - \\
&\quad \beta\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1} + \mathbf{x}_{i-1}^\dagger RI_iRI_iR\mathbf{x}_{i-1}
\end{aligned}$$

where $\beta = 2 - \alpha$. If we can show $\beta RI_iR - RI_iRI_iR$ is positive semi-definite then we are done. Now

$$\beta RI_iR - RI_iRI_iR = \beta RI_i(I - \frac{1}{\beta}R)I_iR.$$

Since $\beta = (1 + \lambda_{max}(R)/2) > \lambda_{max}(R)$ it is easy to see that $I - \frac{1}{\beta}R$ is positive definite. Therefore, $\beta RI_1R - RI_1RI_1R$ is positive semi-definite and

$$\mathbf{x}_i^\dagger R\mathbf{x}_i \leq \mathbf{x}_{i-1}^\dagger R\mathbf{x}_{i-1} - \alpha\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1}.$$

Combining the above inequality for $i = 1, \ldots, P$, we note that $\mathbf{x}_P^\dagger R\mathbf{x}_P < \mathbf{x}_0^\dagger R\mathbf{x}_0$ if $\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1} > 0$ for at least one i, $i = 1, \ldots, P$. We will show by contradiction that is indeed the case.

Suppose not, then $\mathbf{x}_{i-1}^\dagger RI_iR\mathbf{x}_{i-1} = 0$ for all i, $i = 1, \ldots, P$. Since, $\mathbf{x}_0^\dagger RI_1R\mathbf{x}_0 = 0$ this implies $I_1R\mathbf{x}_0 = \mathbf{0}$. Therefore, $\mathbf{x}_1 = (I - I_1R)\mathbf{x}_0 = \mathbf{x}_0$. Similarly, $\mathbf{x}_i = \mathbf{x}_0$ for all i, $i = 1, \ldots, P$. This in turn implies that $\mathbf{x}_0^\dagger RI_iR\mathbf{x}_0 = 0$ for all i, $i = 1, \ldots, P$ which is a contradiction since $R(\sum_{i=1}^{P} I_i)R$ is a positive-definite matrix and $0 = \sum_{i=1}^{P}\mathbf{x}_0^\dagger RI_iR\mathbf{x}_0 = \mathbf{x}_0^\dagger R(\sum_{i=1}^{P} I_i)R\mathbf{x}_0 \neq 0$.

Finally, we conclude that

$$\begin{aligned} \mathbf{x}_0^\dagger \mathbf{P}^\dagger R \mathbf{P} \mathbf{x}_0 &= \mathbf{x}_P^\dagger R \mathbf{x}_P \\ &< \mathbf{x}_0^\dagger R \mathbf{x}_0. \end{aligned}$$

Since $\mathbf{x}_0$ is arbitrary we have $R - \mathbf{P}^\dagger R \mathbf{P}$ to be positive definite so that applying Theorem 2.1 we conclude that $\rho(\mathbf{P}) < 1$.

Finally, if LMS converges in the mean we have $\rho(I - \mu R) < 1$ or $\lambda_{max}(\mu R) < 2$. Which from the above proof is sufficient for concluding that $\rho(\prod_{i=1}^{P}(I - \mu I_i R)) < 1$. Therefore, S-LMS also converges in the mean. ■

2.4 Analysis: Cyclo-stationary Signals

Next, we consider the case when X_k and d_k are jointly cyclo-stationary with covariance matrix R_k. We limit our attention to even-odd S-LMS as shown in Figure 2.1. Let X_k be a cyclo-stationary signal with period L. i.e, $R_{i+L} = R_i$. For simplicity, we will assume L is even. For the regular LMS algorithm we have the following L update equations

$$E\left[V_{k+L}\right] = \prod_{i=0}^{L-1} (I - \mu R_{i+d}) E\left[V_k\right]$$

for $d = 1,\ 2,\ \ldots, L$, in which case we would obtain the following sufficient condition for convergence

$$0 < \mu < \min_i \{2/\lambda_{max}(R_i)\}$$

where $\lambda_{max}(R_i)$ is the largest eigenvalue of the matrix R_i.

Define $A_k = (I - \mu I_1 R_k)$ and $B_k = (I - \mu I_2 R_k)$ then for the partial update algorithm the $2L$ valid update equations are

$$E\left[V_{k+L}\right] = \left(\prod_{i=0}^{\frac{L-1}{2}} B_{2*i+1+d} A_{2*i+d} \right) E\left[V_k\right] \tag{2.6}$$

for $d = 1,\ 2, \ldots, L$ and odd k and

$$E\left[V_{k+L}\right] = \left(\prod_{i=0}^{\frac{L-1}{2}} A_{2*i+1+d} B_{2*i+d} \right) E\left[V_k\right] \tag{2.7}$$

for $d = 1,\ 2, \ldots, L$ and even k.

Let $\|A\|$ denote the spectral norm $\lambda_{max}(AA^{\dagger})^{1/2}$ of the matrix A. Since $\rho(A) \leq \|A\|$ and $\|\prod A_i\| \leq \prod \|A_i\|$, for ensuring the convergence of the iteration (2.6) and (2.7) a sufficient condition is

$$\|B_{i+1}A_i\| < 1 \quad \text{and} \quad \|A_{i+1}B_i\| < 1 \quad \text{for } i = 1,\ 2,\ \ldots, L.$$

Since we can write $B_{i+1}A_i$ as

$$B_{i+1}A_i = (I - \mu R_i) + \mu I_2(R_i - R_{i+1}) + \mu^2 I_2 R_{i+1} I_1 R_i$$

and $A_{i+1}B_i$ as

$$A_{i+1}B_i = (I - \mu R_i) + \mu I_1(R_i - R_{i+1}) + \mu^2 I_1 R_{i+1} I_2 R_i$$

we have the the following expression which upper bounds both $\|B_{i+1}A_i\|$ and $\|A_{i+1}B_i\|$

$$\|I - \mu R_i\| + \mu\|R_{i+1} - R_i\| + \mu^2\|R_{i+1}\|\|R_i\|.$$

This tells us that the sufficient condition to ensure convergence of both (2.6) and (2.7) is

$$\|I - \mu R_i\| + \mu\|R_{i+1} - R_i\| + \mu^2\|R_{i+1}\|\|R_i\| < 1 \tag{2.8}$$

for $i = 1, \ldots, L$.

If we make the assumption that

$$\mu < \min_i\left\{\frac{2}{\lambda_{max}(R_i) + \lambda_{min}(R_i)}\right\} \tag{2.9}$$

and

$$\delta_i = \|R_{i+1} - R_i\| < \max\{\lambda_{min}(R_i), \lambda_{min}(R_{i+1})\} = \eta_i \tag{2.10}$$

for $i = 1, 2, \ldots, L$ then (2.8) translates to

$$1 - \mu\eta_i + \mu\delta_i + \mu^2\lambda_{max}(R_i)\lambda_{max}(R_{i+1}) < 1$$

which gives

$$0 < \mu < \min_{i=1}^{L}\{\frac{\eta_i - \delta_i}{\lambda_{max}(R_i)\lambda_{max}(R_{i+1})}\}. \tag{2.11}$$

Equation (2.11) is the sufficient condition for convergence of even-odd S-LMS with cyclostationary signals.

Therefore, we have the following theorem.

Theorem 2.3. *Let X_k and d_k be jointly cyclostationary. Let R_i, $i = 1, \ldots, L$ denote the L covariance matrices corresponding to the period L of cyclo-stationarity. If we assume X_k is slowly varying in the sense given by (2.10) and μ is small enough given by (2.9) then the sufficient condition on μ for the convergence of iterations (2.6) and (2.7) is given by (2.11)*

2.5 Example

The usefulness of the bound on step-size for the cyclo-stationary case can be gauged from the following example. Consider a 2-tap filter and a cyclo-stationary $\{x_{i,k} = x_{k-i+1}\}$ with period 2 having the following auto-correlation matrices

$$R_1 = \begin{bmatrix} 5.1354 & -0.5733 - 0.6381i \\ -0.5733 + 0.6381i & 3.8022 \end{bmatrix}$$

$$R_2 = \begin{bmatrix} 3.8022 & 1.3533 + 0.3280i \\ 1.3533 - 0.3280i & 5.1354 \end{bmatrix}$$

For this choice of R_1 and R_2, η_1 and η_2 turn out to be 3.38 and we have $\|R_1 - R_2\| = 2.5343 < 3.38$. Therefore, R_1 and R_2 satisfy the assumption made for analysis. Now, $\mu = 0.33$ satisfies the condition for the regular LMS algorithm but, the eigenvalues of B_2A_1 for this value of μ have magnitudes 1.0481 and 0.4605. Since one of the eigenvalues lies outside the unit circle the recursion (2.6) is unstable for this choice of μ. Where as the bound (2.11) gives $\mu = 0.0254$. For this choice of μ the eigenvalues of B_2A_1 turn out to have magnitudes 0.8620 and 0.8773. Hence (2.6) is stable.

We have plotted the evolution trajectory of the 2-tap filter with input signal satisfying the above properties. We chose $W_{opt} = [0.4\ 0.5]$ in Figures 2.2 and 2.3. For Figure 2.2 μ was chosen according to be 0.33 and for Figure 2.3 μ was chosen to be 0.0254. For simulation purposes we set $d_k = W_{opt}^{\dagger} S_k + n_k$ where $S_k = [s_k\ s_{k-1}]^{\tau}$ is a vector composed of the cyclo-stationary process $\{s_k\}$ with correlation matrices given as above, and $\{n_k\}$ is a white sequence, with variance equal to 0.01, independent of $\{s_k\}$. We set $\{x_k\} = \{s_k\} + \{v_k\}$ where $\{v_k\}$ is a white sequence, with variance equal to 0.01, independent of $\{s_k\}$.

2.6 Conclusion

We have analyzed the alternating odd/even partial update LMS algorithm and we have derived stability bounds on step-size parameter μ for wide sense stationary and cyclo-stationary signals based on extremal properties of the matrix 2-norm. For the case of wide sense stationary signals we have shown that if the regular LMS algorithm converges in mean then so does the sequential LMS algorithm for the general case of arbitrary but fixed ordering of the sequence of partial coefficient updates. For cyclo-stationary signals the bounds derived may not be the weakest possible bounds but they do provide the user with a useful sufficient condition on μ which ensures convergence in the mean. We believe the analysis undertaken in this thesis is the first step towards deriving concrete bounds on step-size without making small μ assumptions. The analysis also leads directly to an estimate of mean convergence rate.

In the future, it would be useful to analyze the partial update algorithm, without the assumption of independent snapshots and also, if possible, perform a second order analysis (mean square convergence). Furthermore, as S-LMS exhibits poor convergence in non-stationary signal scenarios (illustrative example given in the following chapter) it is of interest to develop new partial update algorithms with better convergence properties. One such algorithm based on randomized partial updating of filter coefficients is described in the following chapter (chapter 3).

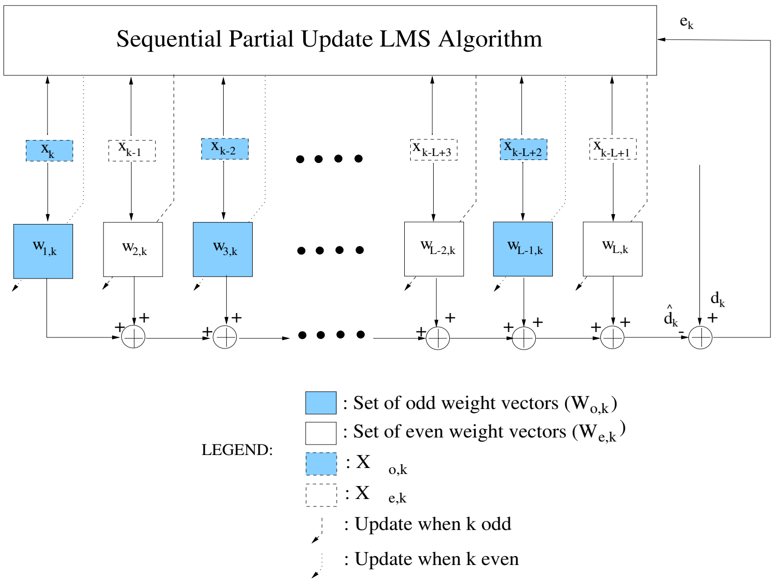


Figure 2.1: Block diagram of S-LMS for the special case of alternating even/odd coefficient update

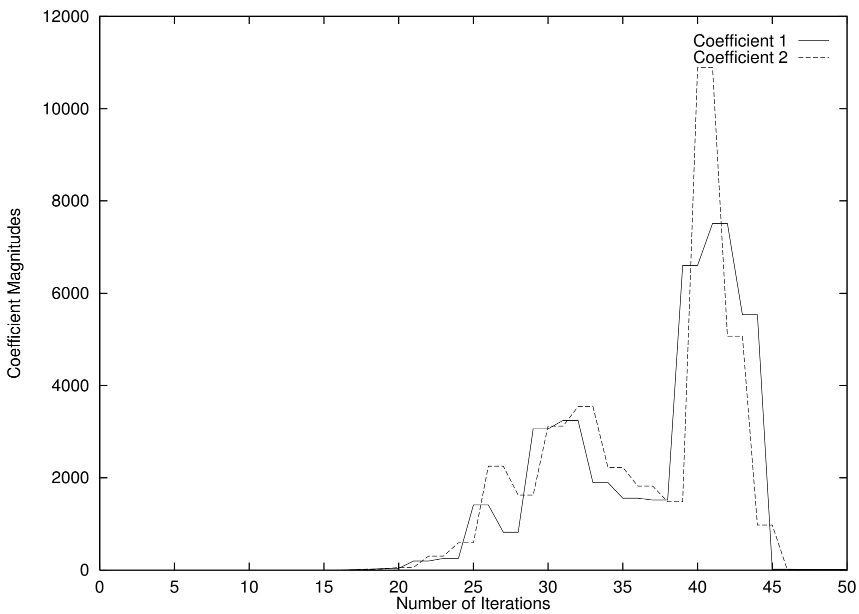


Figure 2.2: Trajectory of $w_{1,k}$ and $w_{2,k}$ for $\mu = 0.33$

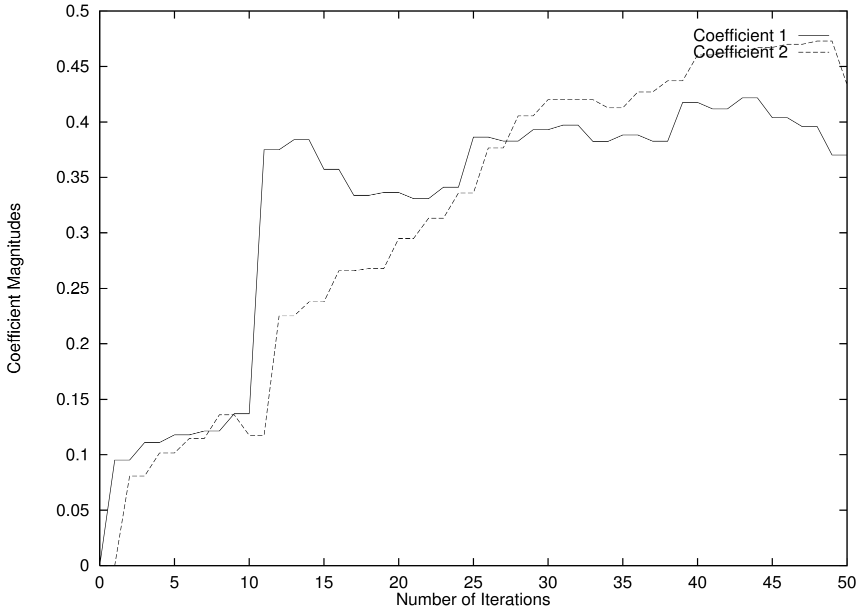


Figure 2.3: Trajectory of $w_{1,k}$ and $w_{2,k}$ for $\mu = 0.0254$

CHAPTER 3

Stochastic Partial Update LMS Algorithm

3.1 Introduction

An important characteristic of the partial update algorithms described in section 1.1 is that the coefficients to be updated at an iteration are pre-determined. It is this characteristic which renders P-LMS (see 1.1) and S-LMS unstable for certain signals and which makes random coefficient updating attractive. The algorithm proposed in this chapter is similar to S-LMS except that the subset of the filter coefficients that are updated each iteration is selected at random. The algorithm, referred to as Stochastic Partial Update LMS algorithm (SPU-LMS), involves selection of a subset of size $\frac{N}{P}$ coefficients out of P possible subsets from a fixed partition of the N coefficients in the weight vector. For example, filter coefficients can be partitioned into even and odd subsets and either even or odd coefficients are randomly selected to be updated in each iteration. In this chapter we derive conditions on the step-size parameter which ensures convergence in the mean and in mean square for stationary signals, generic signals and deterministic signals.

The organization of the chapter is as follows. First, a brief description of the algorithm is given in section 3.2 followed by analysis of the stochastic partial update algorithm for the stationary stochastic signals in section 3.3, deterministic signals in section 3.4 and for generic signals in 3.5. Section 3.6 gives a description of of the existing Partial Update LMS algorithms. This is followed by section 3.8 consisting of examples. In section 3.3 verification of theoretical analysis of the new algorithm is carried out via simulations and examples are given to illustrate the advantages of SPU-LMS. In sections 3.8.1 and 3.8.2 techniques developed in section 3.5 are used to show that the performance of SPU-LMS is very close to that of LMS in terms of final misconvergence. Finally conclusions and directions for future work are indicated in section 3.9.

3.2 Algorithm Description

Unlike in the standard LMS algorithm where all the filter taps are updated every iteration the algorithm proposed in this chapter updates only a subset of coefficients at each iteration. Furthermore, unlike other partial update LMS algorithms the subset to be updated is chosen in a random manner so that eventually every weight is updated.

The description of SPU-LMS is similar to that of S-LMS (section 2.2). The only difference is as as follows. At a given iteration, k, for S-LMS one of the sets S_i, $i = 1, \ldots, P$ is chosen in a *pre-determined fashion* whereas for SPU-LMS, one of the sets S_i are sampled at *random* from $\{S_1,\ S_2, \ldots,\ S_P\}$ with probability $\frac{1}{P}$ and subsequently the update is performed. i.e.

$$w_{k+1,j} = \begin{cases} w_{k,j} + \mu e_k^* x_{k,j} & \text{if } j \in S_i \\ w_{k,j} & \text{otherwise} \end{cases} \tag{3.1}$$

where $e_k = d_k - W_k^\dagger X_k$. The above update equation can be written in a more compact form

$$W_{k+1} = W_k + \mu e_k^* I_i X_k \tag{3.2}$$

where I_i now is a randomly chosen matrix.

3.3 Analysis of SPU-LMS: Stationary Stochastic Signals

In the stationary signal setting the offline problem is to choose an optimal W such that

$$\begin{aligned}\xi^{(W)} &= E\left[(d_k - y_k)(d_k - y_k)^*\right] \\ &= E\left[(d_k - W^\dagger X_k)(d_k - W^\dagger X_k)^*\right]\end{aligned}$$

is minimized, where a^* denotes the complex conjugate of a. The solution to this problem is given by

$$W_{opt} = R^{-1} r \tag{3.3}$$

where $R = E[X_k X_k^\dagger]$ and $r = E[d_k^* X_k]$. The minimum attainable mean square error $\xi^{(W)}$ is given by

$$\xi_{min} = E[d_k d_k^*] - r^\dagger R^{-1} r.$$

For the following analysis, we assume that the desired signal, d_k satisfies the following relation [1][18]

$$d_k = W_{opt}^\dagger X_k + n_k \tag{3.4}$$

where X_k is a zero mean complex circular Gaussian[2] random vector and n_k is a zero mean circular complex Gaussian (not necessarily white) noise, with variance ξ_{min}, uncorrelated with X_k.

[1]Note: the model assumed for d_k is same as assuming d_k and X_k are jointly Gaussian sequences. Under this assumption d_k can be written as $d_k = W_{opt}^\dagger X_k + m_k$, where W_{opt} is as in (3.3) and $m_k = d_k - W_{opt}^\dagger X_k$. Since $E[m_k X_k] = E[X_k d_k] - E[X_k X_k^\dagger] W_{opt} = 0$ and m_k and X_k are jointly Gaussian we conclude that m_k and X_k are independent of each other which is same as model (3.4).

[2]A complex circular Gaussian random vector consists of Gaussian random variables whose marginal densities depend only on their magnitudes. For more information see [55, p. 198] or [50, 73]

We also make the independence assumption used in the analysis of standard LMS [6] which is reasonable for the present application of adaptive beamforming. We assume that X_k is a Gaussian random vector and that X_k is independent of X_j for $j < k$. We also assume that I_i and X_k are mutually independent.

For convergence-in-mean analysis we obtain the following update equation conditioned on a choice of S_i.

$$E[V_{k+1}|S_i] = (I - \mu I_i R)E[V_k|S_i]$$

which after averaging over all choices of S_i gives

$$E[V_{k+1}] = (I - \frac{\mu}{P}R)E[V_k]. \tag{3.5}$$

To obtain the above equation we have made use of the fact that the choice of S_i is independent of V_k and X_k. Therefore, μ has to satisfy $0 < \mu < \frac{2P}{\lambda_{max}}$ to guarantee convergence in mean.

For convergence-in-mean square analysis we are interested in the convergence of $E[e_k e_k^*]$. Under the assumptions we obtain $E[e_k e_k^*] = \xi_{min} + \mathrm{tr}\{RE[V_k V_k^\dagger]\}$ where ξ_{min} is as defined earlier.

We have followed the procedure of [46] for our mean-square analysis. First, conditioned on a choice of S_i, the evolution equation of interest for $\mathrm{tr}\{RE[V_k V_k^\dagger]\}$ is given by

$$\begin{aligned} RE[V_{k+1}V_{k+1}^\dagger|S_i] &= RE[V_k V_k^\dagger|S_i] - 2\mu RI_i RE[V_k V_k^\dagger|S_i] + \\ &\quad \mu^2 I_i RI_i E[X_k X_k^\dagger A_k X_k X_k^\dagger|S_i] + \mu^2 \xi_{min} RI_i RI_i \end{aligned}$$

where $A_k = E[V_k V_k^\dagger]$. For simplicity, consider the case of block diagonal R satisfying $\sum_{i=1}^{P} I_i R I_i = R$. Then, we obtain the final equation of interest for convergence-in-mean square to be

$$G_{k+1} \quad = \quad (I - \frac{2\mu}{P}\Lambda + \frac{\mu^2}{P}\Lambda^2 + \frac{\mu^2}{P}\Lambda^2\mathbf{1}\mathbf{1}^\tau)G_k + \frac{\mu^2}{P}\xi_{min}\Lambda^2\mathbf{1} \tag{3.6}$$

where G_k is a vector of diagonal elements of $\Lambda E[U_k U_k^\dagger]$ where $U_k = QV_k$ with Q such that $QRQ^\dagger = \Lambda$. It is easy to obtain the following necessary and sufficient conditions (see Appendix A.1) for convergence of the SPU-LMS algorithm

$$\begin{aligned} 0 < \mu < \tfrac{2}{\lambda_{max}} \\ \eta(\mu) \stackrel{\text{def}}{=} \textstyle\sum_{i=1}^{N} \tfrac{\mu\lambda_i}{2-\mu\lambda_i} < 1 \end{aligned} \tag{3.7}$$

which is independent of P and identical to that of LMS.

We use the integrated MSE difference $J = \sum_{k=0}^{\infty}[\xi_k - \xi_\infty]$ introduced in [22] as a measure of the convergence rate and $M(\mu) = \frac{\xi_\infty - \xi_{min}}{\xi_{min}}$ as a measure of misadjustment. The misadjustment factor is simply (see Appendix A.3)

$$M(\mu) = \frac{\eta(\mu)}{1 - \eta(\mu)} \tag{3.8}$$

which is the same as that of the standard LMS. Thus, we conclude that random update of subsets has no effect on the final excess mean-squared error.

Finally, it is straightforward to show (see Appendix A.2) the integrated MSE difference is

$$J = P\ \mathrm{tr}\{[2\mu\Lambda - \mu^2\Lambda^2 - \mu^2\Lambda^2\mathbf{1}\mathbf{1}^\tau]^{-1}(G_0 - G_\infty)\} \tag{3.9}$$

which is P times the quantity obtained for standard LMS algorithm. Therefore, we conclude that for block diagonal R, random updating slows down convergence by a factor of P without affecting the misadjustment. Furthermore, it can be easily verified that $0 < \mu < \frac{1}{\mathrm{tr}\{R\}}$ is a sufficient region for convergence of SPU-LMS and the standard LMS algorithm.

The Max PU-LMS described in Section 1.1 is similar SPU-LMS in the sense that the coefficient subset chosen to be updated at an iteration are also random. However, update equations (3.5) and (3.6) are not valid for Max PU-LMS as we can no longer assume that X_k and I_i are independent since the coefficients to be updated in an iteration explicitly depend on X_k.

3.4 Analysis SPU-LMS: Deterministic Signals

Here we followed the analysis given in [63, pp. 140–143] which can be extended to SPU-LMS with complex signals in a straightforward manner. We assume that the input signal X_k is bounded, that is $\sup_k(X_k^\dagger X_k) \leq B < \infty$ and that the desired signal d_k follows the model

$$d_k = W_{opt}^\dagger X_k$$

which is different from (3.4) in that we assume that there is no noise present at the output.

Define $V_k = W_k - W_{opt}$ and $e_k = d_k - W_k^\dagger X_k$.

Lemma 3.1. *If $\mu < 2/B$ then $\overline{e_k^2} \to 0$ as $k \to \infty$. Here, $\overline{\{\cdot\}}$ indicates statistical expectation over all possible choices of S_i, where each S_i is chosen uniformly from $\{S_1, \ldots, S_P\}$.*

Proof: See Appendix A.4

Theorem 3.1. *If $\mu < 2/B$ and the signal satisfies the following persistence of excitation condition:*

For all k, there exist $K < \infty$, $\alpha_1 > 0$ and $\alpha_2 > 0$ such that

$$\alpha_1 I < \sum_{i=k}^{k+K} X_i X_i^\dagger < \alpha_2 I \tag{3.10}$$

then $\overline{V_k}^\dagger \overline{V_k} \to 0$ and $\overline{V_k^\dagger V_k} \to 0$ exponentially fast.

Proof: See Appendix A.4

Condition (3.10) is identical to the persistence of excitation condition for standard LMS. Therefore, the sufficient condition for exponential stability of LMS is enough to guarantee exponential stability of SPU-LMS.

3.5 General Analysis of SPU-LMS

In this section, we analytically compare the performance of LMS and SPU-LMS in terms of stability and misconvergence when the independent snapshots assumption is invalid. For this we employ the theory developed in [37] and [4]. Even though the theory developed is for the case of real random variables it can easily be adapted to the case of complex circular random variables.

In this section, results for stability and performance for the case of SPU-LMS are developed for describing the performance hit taken when going from LMS to SPU-LMS. One of the important results obtained is that for stability LMS and SPU-LMS have the same necessary and sufficient conditions. The theory used for stability analysis and performance analysis follows along [37] and [4], respectively.

3.5.1 Stability Analysis

Notations are the same as those used in [37]. $\|A\|_p$ is used to denote the L_p-norm of a random matrix A given as $\|A\|_p \stackrel{\text{def}}{=} \{E\|A\|^p\|^{1/p}$ for $p \geq 1$ where $\|A\| \stackrel{\text{def}}{=} \{\sum_{i,j} |a|_{ij}^2\}^{1/2}$ is the Euclidean norm of the matrix A. Note that in [37], $\|A\| \stackrel{\text{def}}{=} \{\lambda_{max}(AA^\dagger)\}^{1/2}$. Since the two norms are related by a constant the results in [37] could as well have been stated with the definition used here. Our definition is identical to the norm defined in [4].

A process X_k is said to be ϕ-mixing if there is a function $\phi(m)$ such that $\phi(m) \to 0$

as $m \to \infty$ and

$$\sup_{A \in \mathcal{M}_{-\infty}^{k}(X), B \in \mathcal{M}_{k+m}^{\infty}(X)} |P(B|A) - P(B)| \leq \phi(m), \forall m \geq 0, k \in (-\infty, \infty)$$

where $\mathcal{M}_i^j(X)$, $-\infty \leq i \leq j \leq \infty$ is the σ-algebra generated by $\{X_k\}$, $i \leq k \leq j$

For any random matrix sequence $F = \{F_k\}$, define $\mathcal{S}_p(\alpha, \mu^*)$ for $\mu* > 0$ and $0 < \alpha < 1/\mu^*$ by

$$\begin{aligned} \mathcal{S}_p(\alpha, \mu^*) \;\; = \;\; & \left\{ F : \left\| \prod_{j=i+1}^{k} (I - \mu F_j) \right\|_p \leq K_{\alpha,\mu^*}(F)(1-\mu\alpha)^{k-i} \right. \\ & \left. \qquad \forall \mu \in (0, \mu^*], \forall k \geq i \geq 0 \right\} \end{aligned}$$

$\mathcal{S}_p(\alpha, \mu^*)$ is the family of L_p-stable random matrices.

Similarly, the averaged exponentially stable family is defined as $\mathcal{S}(\alpha, \mu^*)$ for $\mu* > 0$ and $0 < \alpha < 1/\mu^*$ by

$$\begin{aligned} \mathcal{S}(\alpha, \mu^*) \;\; = \;\; & \left\{ F : \left\| \prod_{j=i+1}^{k} (I - \mu E[F_j]) \right\|_p \leq K_{\alpha,\mu^*}(E[F])(1-\mu\alpha)^{k-i} \right. \\ & \left. \qquad \forall \mu \in (0, \mu^*], \forall k \geq i \geq 0 \right\}. \end{aligned} \tag{3.11}$$

We also define $\mathcal{S}_p$ and $\mathcal{S}$ as $\mathcal{S}_p \stackrel{\text{def}}{=} \cup_{\mu^* \in (0,1)} \cup_{\alpha \in (0,1/\mu^*)} \mathcal{S}_p(\alpha, \mu^*)$ and $\mathcal{S} \stackrel{\text{def}}{=} \cup_{\mu^* \in (0,1)} \cup_{\alpha \in (0,1/\mu^*)} \mathcal{S}(\alpha, \mu^*)$.

Let X_k be the input signal vector generated from the following process

$$X_k = \sum_{j=-\infty}^{\infty} A(k,j)\epsilon_{k-j} + \psi_k \tag{3.12}$$

with $\sum_{j=-\infty}^{\infty} \sup_k \|A(k,j)\| < \infty$. $\{\psi_k\}$ is a d-dimensional deterministic process, and $\{\epsilon_k\}$ is a general m-dimensional ϕ-mixing sequence. The weighting matrices $A(k,j) \in \mathcal{R}^{d \times m}$ are assumed to be deterministic.

Define the index set $S = \{1, 2, \ldots, N\}$. Partition S into P mutually exclusive subsets of equal size, S_1, $S_2, \ldots, S_P$. Define $\mathcal{I}_i$ by zeroing out the j^{th} row of the identity matrix I if $j \notin S_i$. Let I_j be a sequence of i.i.d $d \times d$ masking matrices chosen with equal probability from $\mathcal{I}_i$, $i = 1, \ldots, P$.

Then, we have the following theorem which is similar to Theorem 2 in [37].

Theorem 3.2. *Let X_k be as defined above with $\{\epsilon_k\}$ a ϕ-mixing sequence such that it satisfies for any $n \geq 1$ and any increasing integer sequence $j_1 < j_2 < \ldots < j_n$*

$$E\left[\exp\left(\alpha \sum_{i=1}^{n} \|\epsilon_{j_i}\|^2\right)\right] \leq M \exp(Kn) \tag{3.13}$$

where α, M, and K are positive constants. Then for any $p \geq 1$, there exist constants $\mu* > 0$, $M > 0$, and $\alpha \in (0, 1)$ such that for all $\mu \in (0, \mu*]$ and for all $t \geq k \geq 0$

$$\left[E\left\|\prod_{j=k+1}^{t}(I - \mu I_j X_j X_j^{\dagger})\right\|^p\right]^{1/p} \leq M(1 - \mu\alpha)^{t-k}$$

if and only if there exists an integer $h > 0$ and a constant $\delta > 0$ such that for all $k \geq 0$

$$\sum_{i=k+1}^{k+h} E[X_i X_i^{\dagger}] \geq \delta I. \tag{3.14}$$

Proof: For proof see Appendix A.5. ■

Note that the LMS algorithm has the same necessary and sufficient condition for convergence (Theorem 2 in [37]). Therefore, SPU-LMS behaves exactly like LMS in this respect.

Finally, Theorem 2 in [37] follows from 3.2 by setting $I_j = I$ for all j.

3.5.2 Analysis of SPU-LMS for Random Mixing Signals

For performance analysis, we assume that

$$d_k = X_k^{\dagger} W_{opt,k} + n_k$$

$W_{opt,k}$ varies as follows $W_{opt,k+1} - W_{opt,k} = w_{k+1}$, where w_{k+1} is the lag noise. Then for LMS we can write the evolution equation for the tracking error $V_k \stackrel{\text{def}}{=} W_k - W_{opt,k}$

as

$$V_{k+1} = (I - \mu X_k X_k^\dagger)V_k + \mu X_k n_k - w_{k+1}$$

and for SPU-LMS the corresponding equation can be written as

$$V_{k+1} = (I - \mu I_k X_k X_k^\dagger)V_k + \mu X_k n_k - w_{k+1}$$

Now, V_{k+1} can be decomposed [4] as $V_{k+1} = {}^uV_k + \mu^n V_k + {}^wV_k$ where

$$\begin{aligned}
{}^uV_{k+1} &= (I - \mu P_k X_k X_k^\dagger)^u V_k, \quad {}^uV_0 = V_0 = -W_{opt,0} \\
{}^nV_{k+1} &= (I - \mu P_k X_k X_k^\dagger)^n V_k + P_k X_k n_k, \quad {}^nV_0 = 0 \\
{}^wV_{k+1} &= (I - \mu P_k X_k X_k^\dagger)^w V_k - w_{k+1}, \quad {}^nV_0 = 0
\end{aligned}$$

where $P_k = I$ for LMS and $P_k = I_k$ for SPU-LMS. $\{{}^uV_k\}$ denotes the unforced term, reflecting the way the successive estimates of the filter coefficients forget the initial conditions. $\{{}^nV_k\}$ accounts for the errors introduced by the measurement noise, n_k and $\{{}^vV_k\}$ accounts for the errors associated with the lag-noise $\{w_k\}$.

In general ${}^nV_k$ and ${}^wV_k$ obey the following inhomogeneous equation

$$\delta_{k+1} = (I - \mu F_k)\delta_k + \xi_k, \quad \delta_0 = 0$$

δ_k can be represent by a set of recursive equations as follows

$$\delta_k = J_k^{(0)} + J_k^{(1)} + \ldots + J_k^{(n)} + H_k^{(n)}$$

where the processes $J_k^{(r)}, 0 \le r < n$ and $H_k^{(n)}$ are described by

$$\begin{aligned}
J_{k+1}^{(0)} &= (I - \mu \bar{F}_k)J_k^{(0)} + \xi_k; J_0^{(0)} = 0 \\
J_{k+1}^{(r)} &= (I - \mu \bar{F}_k)J_k^{(r)} + \mu Z_k J_k^{(r-1)}; \quad J_k^{(r)} = 0, 0 \le k < r \\
H_{k+1}^{(n)} &= (I - \mu F_k)H_k^{(n)} + \mu Z_k J_k^{(n)}; \quad H_k^{(n)} = 0, 0 \le k < n
\end{aligned}$$

where $Z_k = F_k - \bar{F}_k$ and $\bar{F}_k$ is an appropriate deterministic process, usually chosen as $\bar{F}_k = E[F_k]$. In [4] under appropriate conditions it was shown that there exists some constant $C < \infty$ and $\mu_0 > 0$ such that for all $0 < \mu \leq \mu_0$, we have

$$\sup_{k \geq 0} \|H_k^{(n)}\|_p \leq C\mu^{n/2}.$$

Now, we modify the definition of weak dependence as given in [4] for circular complex random variables. The theory developed in [4] can be easily adapted for circular random variables using this definition. Let $q \geq 1$ and $X = \{X_n\}_{n \geq 0}$ be a $(l \times 1)$ matrix valued process. Let $\beta = (\beta(r))_{r \in N}$ be a sequence of positive numbers decreasing to zero at infinity. The complex process $X = \{X_n\}_{n \geq 0}$ is said to be (δ, q)-weak dependent if there exist finite constants $C = \{C_1, \ldots, C_q\}$, such that for any $1 \leq m < s \leq q$ and m-tuple $k_1, \ldots, k_m$ and any $(s-m)$-tuple $k_{m+1}, \ldots, k_s$, with $k_1 \leq \ldots \leq k_m < k_m + r \leq k_{m+1} \leq \ldots \leq k_s$, it holds that

$$\begin{aligned} \sup_{1 \leq i_1, \ldots, i_s \leq l, f_{k_1,i_1}, f_{k_2,i_2} \cdots f_{k_m,i_m}} \quad & \Big| \mathrm{cov} \Big(f_{k_1,i_1}(\tilde{X}_{k_1,i_1}) \cdot \ldots \cdot f_{k_m,i_m}(\tilde{X}_{k_m,i_m}), \\ & f_{k_{m+1},i_{m+1}}(\tilde{X}_{k_{m+1},i_{m+1}}) \cdot \ldots \cdot f_{k_s,i_s}(\tilde{X}_{k_s,i_s}) \Big) \Big| \quad \leq C_s \beta(r) \end{aligned}$$

where $\tilde{X}_{n,i}$ denotes the i-th component of $X_n - E(X_n)$ and the set of functions $f_{n,i}()$ that the sup is being taken over are given by $f_{n,i}(\tilde{X}_{n,i}) = \tilde{X}_{n,i}$ and $f_{n,i}(\tilde{X}_{n,i}) = \tilde{X}^*_{n,i}$.

Define $\mathcal{N}(p)$ from [4] as follows

$$\begin{aligned} \mathcal{N}(p) = \quad & \Big\{ \epsilon : \big\| \textstyle\sum_{k=s}^{t} D_k \epsilon_k \big\|_p \leq \rho_p(\epsilon) \left(\sum_{k=s}^{t} |D_k|^2 \right)^{1/2} \ \forall 0 \leq s \leq t \\ & \text{and } \forall D = \{D_k\}_{k \in N} (q \times l) \text{ deterministic matrices } \Big\} \end{aligned}$$

where $\rho_p(\epsilon)$ is a constant depending only on the process ϵ and the number p.

F_k can be written as $F_k = P_k X_k X_k^\dagger$ where $P_k = I$ for LMS and $P_k = I_k$ for SPU-LMS. It is assumed that the following hold true for F_k. For some $r, q \in N$, $\mu_0 > 0$ and $0 < \alpha < 1/\mu_0$

- $\mathbf{F1}(r, \alpha, \mu_0)$: $\{F_k\}_{k\geq 0}$ is in $\mathcal{S}(r, \alpha, \mu_0)$ that is $\{F_k\}$ is L_r-exponentially stable.
- $\mathbf{F2}(\alpha, \mu_0)$: $\{E[F_k]\}_{k\geq 0}$ is in $\mathcal{S}(\alpha, \mu_0)$, that is $\{E[F_k]\}_{k\geq 0}$ is averaged exponentially stable.

Conditions **F3** and **F4** stated below are trivially satisfied for $P_k = I$ and $P_k = I_k$.

- $\mathbf{F3}(q, \mu_0)$: $\sup_{k\in N} \sup_{\mu\in(0,\mu_0]} \|P_k\|_q < \infty$ and $\sup_{k\in N} \sup_{\mu\in(0,\mu_0]} |E[P_k]| < \infty$
- $\mathbf{F4}(q, \mu_0)$: $\sup_{k\in N} \sup_{\mu\in(0,\mu_0]} \mu^{-1/2}\|P_k - E[P_k]\|_q < \infty$

The excitation sequence $\xi = \{\xi_k\|_{k\geq 0}$ [4] is assumed to be decomposed as $\xi_k = M_k\epsilon_k$ where the processes $M = \{M_k\}_{k\geq 0}$ is a $d \times l$ matrix valued process and $\epsilon = \{\epsilon_k\}_{k\geq 0}$ is a $(l \times 1)$ vector-valued process that verifies the following assumptions

- **EXC1**: $\{M_k\}_{k\in Z}$ is $\mathcal{M}_0^k(X)$-adapted[3] and $\mathcal{M}_0^k(\epsilon)$ and $\mathcal{M}_0^k(X)$ are independent.
- $\mathbf{EXC2}(r, \mu_0)$: $\sup_{\mu\in(0,\mu_0]} \sup_{k\geq 0} \|M_k\|_r < \infty$, $(r > 0, \mu_0 > 0)$
- $\mathbf{EXC3}(p, \mu_0)$: $\epsilon = \{\epsilon_k\}_{k\in N}$ belongs to $\mathcal{N}(p)$, $(p > 0, \mu_0 > 0)$

The following theorems from [4] are relevant.

Theorem 3.3 (Theorem 1 in [4]). *Let $n \in N$ and let $q \geq p \geq 2$. Assume* **EXC1**, $\mathbf{EXC2}(pq/(q-p), \mu_0)$ *and* $\mathbf{EXC3}(p, \mu_0)$*. For $a, b, \alpha > 0$, $a^{-1} + b^{-1} = 1$, and some $\mu_0 > 0$, assume in addition* $\mathbf{F2}(\alpha, \mu_0)$, $\mathbf{F4}(aqn, \mu_0)$ *and*

- *$\{G_k\}_{k\geq 0}$ is $(\beta, (q+2)n)$ weakly dependent and $\sum (r+1)^{((q+2)n/2)-1}\beta(r) < \infty$*
- $\sup_{k\geq 0} \|G_k\|_{bqn} < \infty$

Then, there exists a constant $K < \infty$ (depending on $\beta(k)$, $k \geq 0$ and on the numerical constants $p, q, n, q, b, \mu_0, \alpha$ but not otherwise on $\{X_k\}$, $\{\epsilon_k\}$ or on μ), such

[3] A sequence of random variables, X_i is called adapted with respect to a sequence of σ-fields $\mathcal{F}_i$ if X_i is $\mathcal{F}_i$ measurable [11].

that for all $0 < \mu \leq \mu_0$*, for all* $0 \leq r \leq n$

$$\sup_{s \geq 1} \|J_s^{(r)}\|_p \leq K\rho_p(\epsilon) \sup_{k \geq 0} \|M_k\|_{pq/(q-p)} \mu^{(r-1)/2}.$$

Theorem 3.4 (Theorem 2 in [4]). *Let* $p \geq 2$ *and let* $a, b, c > 0$ *such that* $1/a + 1/b + 1/c = 1/p$*. Let* $n \in N$*. Assume* $\mathbf{F1}(a, \alpha, \mu_0)$ *and*

- $\sup_{s \geq 0} \|Z_s\|_b < \infty$
- $\sup_{s \geq 0} \|J_s^{(n+1)}\|_c < \infty$

Then there exists a constant $K' < \infty$ *(depending on the numerical constants* $a, b, c,$ α, μ_0, n *but not on the process* $\{\epsilon_k\}$ *or on the stepsize parameter* μ*), such that for all* $0 < \mu \leq \mu_0$*,*

$$\sup_{s \geq 0} \|H_s^{(n)}\|_p \leq K' \sup_{s \geq 0} \|J_s^{(n+1)}\|_c.$$

We next show that if LMS satisfies the assumptions above (assumptions in section 3.2 in [4]) then so does SPU-LMS. Conditions **F1** and **F2** follow directly from Theorem 3.2. It is easy to see that **F3** and **F4** hold easily for LMS and SPU-LMS.

Lemma 3.2. *The constant* K *in Theorem 3.3 calculated for LMS can also be used for SPU-LMS.*

Proof: Here all that is needed to be shown is that if LMS satisfies the conditions **(EXC1)**, **(EXC2)** and **(EXC3)** then so does SPU-LMS. Moreover, the upper bounds on the norms for LMS are also upper bounds for SPU-LMS. That easily follows because $M_k^{LMS} = X_k$ whereas $M_k^{SPU-LMS} = I_k X_k$ and $\|I_k\| \leq 1$ for any norm $\|\cdot\|$. ∎

Lemma 3.3. *The constant* K' *in Theorem 3.4 calculated for LMS can also be used for SPU-LMS.*

Proof: First we show that if for LMS $\sup_{s\geq 0}\|Z_s\|_b < \infty$ then so it is for SPU-LMS. First, note that for LMS we can write $Z_s^{LMS} = X_sX_s^\dagger - E[X_sX_s^\dagger]$ whereas for SPU-LMS

$$\begin{aligned} Z_s^{SPU-LMS} &= I_sX_sX_s^\dagger - \frac{1}{P}E[X_sX_s^\dagger] \\ &= I_sX_sX_s^\dagger - I_sE[X_sX_s^\dagger] + (I_s - \frac{1}{P}I)E[X_sX_s^\dagger] \end{aligned}$$

That means $\|Z_s^{SPU-LMS}\|_b \leq \|I_s\|_b\|Z_s^{LMS}\|_b + \|I_s - \frac{1}{P}I\|_b\|E[X_sX_s^\dagger]\|_b$. Therefore, since $\sup_{s\geq 0}\|_bE[X_sX_s^\dagger]\|_b < \infty$ and $\sup_{s\geq 0}\|Z_s^{LMS}\|_b < \infty$ we have

$$\sup_s \|Z_s^{SPU-LMS}\|_b < \infty.$$

Since all conditions for Theorem 2 have been satisfied by SPU-LMS in a similar manner the constant obtained is also the same. ■

The two lemmas states that the error terms are bounded above by same constants.

3.6 Periodic and Sequential LMS Algorithms

For P-LMS, the update equation can be written as follows

$$W_{k+P} = W_k + \mu e_k^* X_k$$

For the Sequential LMS algorithm the update equation is same as (3.2) except that the choice of I_i is no longer random. The sequence of I_i as k progresses is pre-determined and fixed.

For the P-LMS algorithm, using the method of analysis described in [46] we conclude that the conditions for convergence are identical to standard LMS. That is (3.7) holds also for P-LMS. Also, the misadjustment factor remains the same. The only difference between LMS and P-LMS is that the measure J for P-LMS is P times that of LMS. Therefore, we see that the behavior of SPU-LMS and P-LMS algorithms is very similar for stationary signals.

The difference between P-LMS and SPU-LMS becomes evident for deterministic signals. From [18] we conclude that the persistence of excitation condition for P-LMS is stricter than that for SPU-LMS. In fact, in the next section we construct signals for which P-LMS is guaranteed not to converge whereas SPU-LMS will converge.

The convergence of Sequential LMS algorithm has been analyzed using the small μ assumption in [18]. Theoretical results for this algorithm are not presented here. However, we show through simulation examples that this algorithm diverges for certain signals and therefore should be employed with caution.

3.7 Simulation of an Array Examples to Illustrate the advantage of SPU-LMS

We simulated an m-element uniform linear antenna array operating in a multiple signal environment. Let A_i denote the response of the array to the i^{th} plane wave signal: $A_i = [e^{-\mathbf{j}(\frac{m}{2}-\tilde{m})\omega_i}\ e^{-\mathbf{j}(\frac{m}{2}-1-\tilde{m})\omega_i}\ \ldots\ e^{\mathbf{j}(\frac{m}{2}-1-\tilde{m})\omega_i}\ e^{\mathbf{j}(\frac{m}{2}-\tilde{m})\omega_i}]^\tau$ where $\tilde{m} = (m+1)/2$ and $\omega_i = \frac{2\pi D \sin\theta_i}{\lambda}$, $i = 1, \ldots, M$. θ_i is the broadside angle of the i^{th} signal, D is the inter-element spacing between the antenna elements and λ is the common wavelength of the narrowband signals in the same units as D and $\frac{2\pi D}{\lambda} = 2$. The array output at the k^{th} snapshot is given by $X_k = \sum_{i=1}^{M} A_i s_{k,i} + n_k$ where M denotes the number of signals, the sequence $\{s_{k,i}\}$ the amplitude of the i^{th} signal and n_k the noise present at the array output at the k^{th} snapshot. The objective, in both the examples, is to maximize the SNR at the output of the beamformer. Since the signal amplitudes are random the objective translates to obtaining the best estimate of $s_{k,1}$, the amplitude of the desired signal, in the MMSE sense. Therefore, the desired signal is chosen as $d_k = s_{k,1}$.

In the first example (Figure 3.1), the array has 4 elements and a single planar waveform with amplitude, $s_{k,1}$ propagates across the array from direction angle,

$\theta_1 = \frac{\pi}{2}$. The amplitude sequence $\{s_{k,1}\}$ is a binary phase shifty keying (BPSK) signal with period four taking values on $\{-1, 1\}$ with equal probability. The additive noise n_k is circular Gaussian with variance 0.25 and mean 0. In all the simulations for SPU-

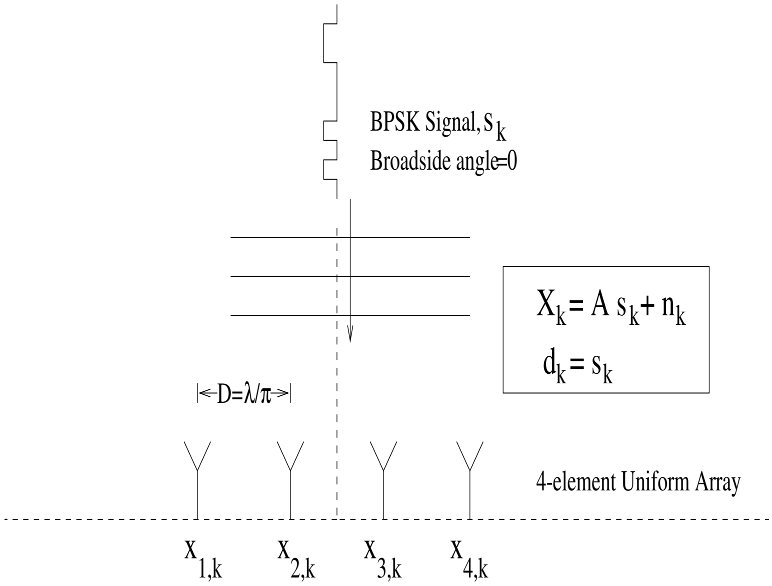


Figure 3.1: Signal Scenario for Example 1

LMS, P-LMS, and S-LMS the number of subsets for partial updating, P was chosen to be 4. It can be easily determined from (3.7) that for Gaussian and independent signals the necessary and sufficient condition for convergence of LMS and SPU-LMS is $\mu < 0.67$. Figure 3.2 shows representative trajectories of the empirical mean-squared error for LMS, SPU-LMS, P-LMS and S-LMS algorithms averaged over 100 trials for $\mu = 0.6$ and $\mu = 1.0$. All algorithms were found to be stable for the BPSK signals even for μ values greater than 0.67. It was only as μ approached 1 that divergent behavior was observed. As expected, LMS and SPU-LMS were observed to have similar μ regions of convergence. It is also clear from Figure 3.2, that as, expected SPU-LMS, P-LMS, and S-LMS take roughly 4 times longer to converge than LMS.

In the second example, we consider an 8-element uniform linear antenna array

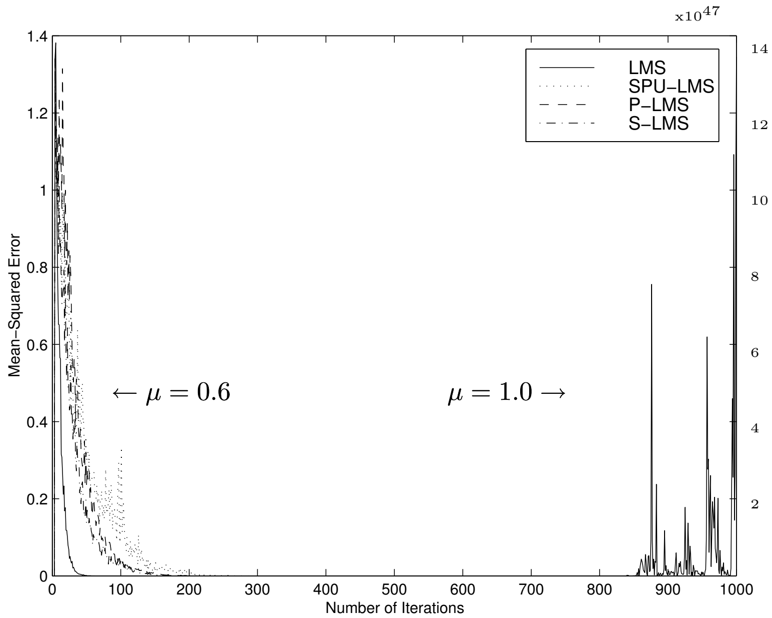


Figure 3.2: Trajectories of MSE for Example 1

with one signal of interest propagating at angle θ_1 and 3 interferers propagating at angles θ_i, $i = 2, 3, 4$. The array noise n_k is again mean 0 circular Gaussian but with variance 0.001. We generated signals, such that $s_{k,1}$ is stationary and $s_{k,i}$, $i = 2, 3, 4$ are cyclostationary with period four, which make both S-LMS and P-LMS non-convergent. All the signals were chosen to be independent from time instant to time instant. First, we found signals for which S-LMS doesn't converge by the following procedure. Make the small μ approximation $I - \mu \sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^\dagger]$ to the transition matrix $\prod_{i=1}^{P}(I - \mu I_i E[X_{k+i} X_{k+i}])$ and generate sequences $s_{k,i}$, $i = 1, 2, 3, 4$ such that $\sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^\dagger]$ has roots in the negative left half plane. This ensures that $I - \mu \sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^\dagger]$ has roots outside the unit circle. The sequences found in this manner were then verified to cause the roots to lie outside the unit circle for all μ. One such set of signals found was: $s_{k,1}$ is equal to a BPSK signal with period one taking values in $\{-1, 1\}$ with equal probability. The interferers, $s_{k,i}$, $i = 2, 3, 4$ are cyclostationary BPSK type signals taking values in $\{-1, 1\}$ with the restriction

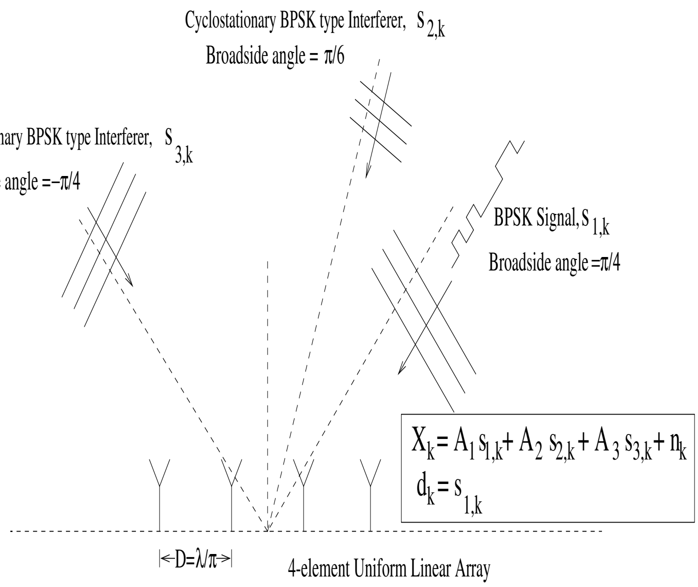


Figure 3.3: Signal Scenario for Example 2

that $s_{k,2} = 0$ if $k \% 4 \neq 1$, $s_{k,3} = 0$ if $k \% 4 \neq 2$ and $s_{k,4} = 0$ if $k \% 4 \neq 3$. Here $a \% b$ stands for a modulo b. θ_i, $i = 1, 2, 3, 4$ are chosen such that $\theta_1 = 1.0388$, $\theta_2 = 0.0737$, $\theta_3 = 1.0750$ and $\theta_4 = 1.1410$. These signals render the S-LMS algorithm unstable for all μ.

The P-LMS algorithm also fails to converge for the signal set described above irrespective of μ and the choice of θ_1, θ_2, θ_3, and θ_4. Since P-LMS updates the coefficients every 4^{th} iteration it sees at most one of the three interfering signals throughout all its updates and hence can place a null at atmost one signal incidence angle θ_i. Figure 3.4 shows the envelopes of the e_k^2 trajectories of S-LMS and P-LMS for the signals given above with the representative value $\mu = 0.03$. As can be seen P-LMS fails to converge whereas S-LMS shows divergent behavior. SPU-LMS and LMS were observed to converge for the signal set described above when $\mu = 0.03$.

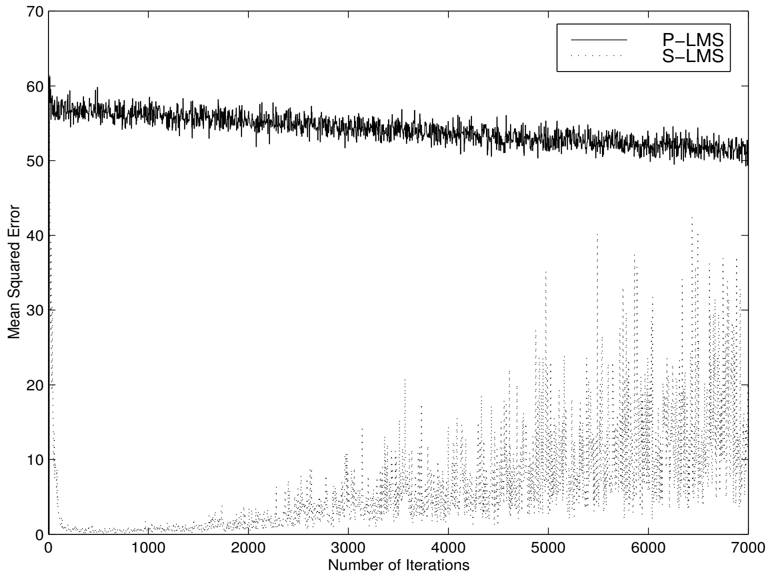


Figure 3.4: Trajectories of MSE for Example 2

3.8 Examples

3.8.1 I.i.d Gaussian Input Sequence

In this section, we assume that $X_k = [x_k \; x_{k-1} \; \ldots \; x_{k-N+1}]^\tau$ where N is the length of the vector X_k. $\{x_k\}$ is a sequence of zero mean i.i.d Gaussian random variables. We assume that $w_k = 0$ for all $k \geq 0$. In that case

$$V_{k+1} = (I - \mu P_k X_k X_k^\dagger)V_k + X_k n_k \quad V_0 = -W_{opt,0} = W_{opt}$$

where for LMS we have $P_k = I$ and $P_k = I_k$ in case of SPU-LMS. We assume n_k is a white i.i.d. Gaussian noise with variance σ_v^2. We see that since the conditions (3.13) and (3.14) are satisfied for theorem 3.2 both LMS and SPU-LMS are exponentially stable. In fact both have the same α exponent of decay. Therefore, conditions **F1** and **F2** are satisfied.

We rewrite $V_k = J_k^{(0)} + J_k^{(1)} + J_k^{(2)} + H_k^{(2)}$. Choosing $\bar{F}_k = E[F_k]$ we have $E[P_k X_k X_k^\dagger] = \sigma^2 I$ in the case of LMS and $\frac{1}{P}\sigma^2 I$ in the case of SPU-LMS. By Theorems 3.3 and 3.4 and Lemmas 3.2 and 3.3 we can upperbound both $|J_k^{(2)}|$ and $|H_k^{(2)}|$

by exactly the same constants for LMS and SPU-LMS. In particular, there exists some constant $C < \infty$ such that for all $\mu \in (0, \mu_0]$, we have

$$\begin{aligned}
\sup_{t\geq 0} \left| E[J_t^{(1)}(J_t^{(2)} + H_t^{(2)})^\dagger] \right| &\leq C\|X_0\|_{r(r+\delta)/\delta}\rho_r^2(v)\mu^{1/2} \\
\sup_{t\geq 0} \left| E[J_t^{(0)} H_t^{(2)}] \right| &\leq C\rho_r(v)\|X_0\|_{r(r+\delta)/\delta}\mu^{1/2}.
\end{aligned}$$

Next, for LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1 - \mu\sigma^2)J_k^{(0)} + X_k n_k \\
J_{k+1}^{(1)} &= (1 - \mu\sigma^2)J_k^{(1)} + \mu(\sigma^2 I - X_k X_k^\dagger)J_k^{(0)}
\end{aligned}$$

and for SPU-LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1 - \frac{\mu}{P}\sigma^2)J_k^{(0)} + I_k X_k n_k \\
J_{k+1}^{(1)} &= (1 - \frac{\mu}{P}\sigma^2)J_k^{(1)} + \mu(\frac{\sigma^2}{P} I - I_k X_k X_k^\dagger)J_k^{(0)}.
\end{aligned}$$

Solving (see Appendix A.6), we obtain for LMS

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] &= \frac{\sigma_v^2}{\mu(2 - \mu\sigma^2)} I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] &= 0 \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] &= 0 \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] &= \frac{N\sigma^2\sigma_v^2}{(2 - \mu\sigma^2)^2} I \\
&= \frac{N\sigma^2\sigma_v^2}{4} I + O(\mu) I
\end{aligned}$$

which yields $\lim_{k\to\infty} E[V_k V_k^\dagger] = \frac{\sigma_v^2}{2\mu} I + \frac{N\sigma^2\sigma_v^2}{4} I + O(\mu^{1/2}) I$ and for SPU-LMS we obtain

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] &= \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P}\sigma^2)} I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] &= 0 \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] &= 0 \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] &= \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{(2-\frac{\mu}{P}\sigma^2)^2} I \\
&= \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{4} I + O(\mu) I
\end{aligned}$$

which yields $\lim_{k\to\infty} E[V_k V_k^\dagger] = \frac{\sigma_v^2}{2\mu} I + \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{4} I + O(\mu^{1/2}) I$. Therefore, we see that SPU-LMS is marginally worse than LMS in terms of misadjustment.

3.8.2 Temporally Correlated Spatially Uncorrelated Array Output

In this section we consider X_k given by

$$X_k = \kappa X_{k-1} + \sqrt{1-\kappa^2} U_k$$

where U_k is a vector of circular Gaussian random variables with unit variance. Similar to section 3.8.1, we rewrite $V_k = J_k^{(0)} + J_k^{(1)} + J_k^{(2)} + H_k^{(2)}$. Since, we have chosen $\bar{F}_k = E[F_k]$ we have $E[P_k X_k X_k^\dagger] = I$ in the case of LMS and $\frac{1}{P} I$ in the case of SPU-LMS. Again, conditions **F1** and **F2** are satisfied because of Theorem 3.2. By [4] and Lemmas 1 and 2 we can upperbound both $J_k^{(2)}$ and $H_k^{(2)}$ by exactly the same constants for LMS and SPU-LMS. By Theorems 3.3 and 3.4 and Lemmas 3.2 and 3.3 we have that there exists some constant $C < \infty$ such that for all $\mu \in (0, \mu_0]$, we have

$$\begin{aligned}
\sup_{t\geq 0} \left| E[J_t^{(1)}(J_t^{(2)} + H_t^{(2)})^\dagger] \right| &\leq C \|X_0\|_{r(r+\delta)/\delta} \rho_r^2(v) \mu^{1/2} \\
\sup_{t\geq 0} \left| E[J_t^{(0)} H_t^{(2)}] \right| &\leq C \rho_r(v) \|X_0\|_{r(r+\delta)/\delta} \mu^{1/2}.
\end{aligned}$$

Next, for LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1-\mu)J_k^{(0)} + X_k n_k \\
J_{k+1}^{(1)} &= (1-\mu)J_k^{(1)} + \mu(I - X_k X_k^\dagger)J_k^{(0)}
\end{aligned}$$

and for SPU-LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1-\frac{\mu}{P})J_k^{(0)} + I_k X_k n_k \\
J_{k+1}^{(1)} &= (1-\frac{\mu}{P})J_k^{(1)} + \mu(\frac{1}{P}I - I_k X_k X_k^\dagger)J_k^{(0)}.
\end{aligned}$$

Solving (see Appendix A.7), we obtain for LMS

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] &= \frac{\sigma_v^2}{\mu(2-\mu)}I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] &= -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)}I + O(\mu)I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] &= \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)}I + O(\mu)I \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] &= \frac{(1+\kappa^2)\sigma_v^2 N}{4(1-\kappa^2)}I + O(\mu)I
\end{aligned}$$

which leads to $\lim_{k\to\infty} E[V_k V_k^\dagger] = \frac{\sigma_v^2}{2\mu}I + \frac{N\sigma_v^2}{4}I + O(\mu^{1/2})I$ and for SPU-LMS we obtain

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] &= \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P})}I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] &= -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)P}I + O(\mu)I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] &= \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)P}I + O(\mu)I \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] &= \frac{\sigma_v^2}{4}[\frac{N}{P}\frac{1+\kappa^2}{1-\kappa^2} + (N+1)\frac{P-1}{P}]I + O(\mu)I
\end{aligned}$$

which leads to $\lim_{k\to\infty} E[V_k V_k^\dagger] = \frac{\sigma_v^2}{2\mu}I + \frac{\sigma^2}{4}[N+1-\frac{1}{P}]I + O(\mu^{1/2})I$. Again, SPU-LMS is marginally worse than LMS in terms of misadjustment.

3.9 Conclusion and Future Work

We have proposed a new algorithm based on randomization of filter coefficient subsets for partial updating of filter coefficients. The conditions on step-size for convergence-in-mean and mean-square were shown to be equivalent to those of standard LMS. It was verified by theory and by simulation that LMS and SPU-LMS have similar regions of convergence. We also have shown that the Stochastic Partial Update LMS algorithm has the same performance as the Periodic LMS algorithm for stationary signals but, can have superior performance for some cyclo-stationary and deterministic signals.

The idea of random choice of subsets proposed in the chapter can be extended to include arbitrary subsets of size $\frac{N}{P}$ and not just subsets from a particular partition. No special advantage is immediately evident from this extension though.

CHAPTER 4

Capacity: Isotropically Random Rician Fading

4.1 Introduction

In this chapter, we analyze MIMO channel capacity under a Rayleigh/Rician fading model with average energy constraint on the input. The model consists of a line of sight component (specular component) and a diffuse component (Rayleigh component) both changing over time. We model the specular component as isotropically random statistically independent of the Rayleigh component. This model could apply to the situation where we have either the transmitter or the receiver in motion resulting in variable diffuse and specular components.

Traditionally, in a Rician model the fading coefficients are modeled as Gaussian with non-zero mean. We depart from the traditional model in the sense that we model the mean (specular component) as time-varying and stochastic. The specular component is modeled as an isotropic rank one matrix with the specular component staying constant for T symbol durations and taking independent values every T^{th} instant. We establish similar properties for this Rician model as those shown by Marzetta and Hochwald [50] for the Rayleigh fading model. In particular, it is sufficient to optimize over a smaller parameter set of size $\min\{T, M\}$ of real valued magnitudes of the transmitted signals instead of $T \cdot M$ complex valued symbols.

Furthermore, the capacity achieving signal matrix is shown to be the product of two independent matrices, a $T \times T$ isotropically random unitary matrix and a $T \times M$ real nonnegative matrix.

The isotropically random Rician fading model is described in detail in section 4.2. In section 4.4, we derive a new lower bound on capacity. The lower bound also holds for the case of a purely Rayleigh fading channel. In section 4.5 we show the utility of this bound by computing capacity regions for both Rayleigh and Rician fading channels.

4.2 Signal Model

The fading channel is assumed to stay constant for T channel uses and then take on a completely independent realization for the next T channel uses. Let there be M transmit antennas and N receive antennas. We transmit a $T \times M$ signal matrix S and receive a $T \times N$ signal matrix X which are related as follows

$$X = \sqrt{\frac{\rho}{M}} SH + W \tag{4.1}$$

where the elements, w_{tn} of W are independent circularly symmetric complex Gaussian random variables with mean 0 and variance 1 ($\mathcal{CN}(0,1)$) and ρ is the average signal to noise ratio present at each of the receive antennas.

The only difference between the Rayleigh model and the Rician model considered here involves the statistics of the fading matrix H. In the case of the Rayleigh model the elements h_{mn} of H are modeled as independent $\mathcal{CN}(0,1)$ random variables. For the isotropically random Rician fading model, the matrix H is modeled as

$$H = \sqrt{1-r}G + \sqrt{rNM} v \alpha \beta^\dagger$$

where G consists of independent $\mathcal{CN}(0,1)$ random variables, v is a real random variable such that $E[v^2] = 1$ and α and β are independent isotropically random unit magnitude vectors of length M and N, respectively. G, α and β take on independent values every T^{th} symbol period and remain unchanging in between. The parameter r ranges between zero and one, with the limits corresponding respectively to purely Rayleigh or purely specular propagation. Irrespective of the value of r, the average variance of the components of H is equal to one, $E[\text{tr}\{HH^\dagger\}] = M \cdot N$.

An M-dimensional unit vector α is isotropically random if its probability density is invariant to pre-multiplication by an $M \times M$ deterministic unitary matrix, that is $p(\Psi\alpha) = p(\alpha)$, $\forall \Psi : \Psi^\dagger\Psi = I_M$ [50]. The isotropic density is $p(\alpha) = \frac{\Gamma(M)}{\pi^M}\delta(\alpha^\dagger\alpha - 1)$ where $\Gamma(M) = (M-1)!$.

The choice of the density $p(v)$ of v is not clear. One choice is for $p(v)$ to maximize the entropy of $R = v\alpha\beta^\dagger$ corresponding to the worst case scenario (capacity = minimum) for the channel capacity. One might expect that R corresponding to maximum entropy would be Gaussian distributed. However, that this is not the case follows from the proposition below.

Proposition 4.1. *There is no distribution $p(v)$ such that the elements of $R = v\alpha\beta^\dagger$ have a joint Gaussian distribution, where α and β are isotropically random unitary vectors, and v, α, and β are mutually independent.*

Proof: Proof is by contradiction. Consider the covariance of the elements, R_{mn} of R.

$$\begin{aligned} E[R_{m_1n_1}R^*_{m_2n_2}] &= E[v^2]E[\alpha_{m_1}\alpha^*_{m_2}]E[\beta_{n_1}\beta^*_{n_2}] \\ &= E[v^2]\frac{1}{M}\delta_{m_1m_2}\delta_{n_1n_2}. \end{aligned}$$

If elements of R were jointly Gaussian then they must be independent of each other which contradicts the assumption that R is of rank one. ∎

From now on we will assume that v is identically equal to 1. In that case, the conditional probability density function of the measurements at the receiver is given by

$$p(X|S) = E_\alpha E_\beta \left[\frac{e^{-\mathrm{tr}\{[I_T+(1-r)\frac{\rho}{M}SS^\dagger]^{-1}(X-\sqrt{\rho r N}S\alpha\beta^\dagger)(X-\sqrt{\rho r N}S\alpha\beta^\dagger)^\dagger\}}}{\pi^{TN}\det^N[I_T+(1-r)\frac{\rho}{M}SS^\dagger]} \right]$$

where E_α denotes the expectation over the density of α.

Irrespective of whether the fading is Rayleigh or Rician, we have $p(\Psi^\dagger H) = p(H)$ for any $M \times M$ unitary matrix Ψ. In the rest of the section we will deal with H satisfying this property and refer to Rayleigh and Rician fading as special cases of this channel. In that case, the condition probability density of the received signals has the following properties

1. For any $T \times T$ unitary matrix Φ

$$p(\Phi X|\Phi S) = p(X|S)$$

2. For any $M \times M$ unitary matrix Ψ

$$p(X|S\Psi) = p(X|S)$$

Lemma 4.1. *If $p(X|S)$ satisfies property 2 defined above then the transmitted signal can be written as ΦV where Φ is a $T \times T$ unitary matrix and V is a $T \times M$ real nonnegative diagonal matrix.*

Proof: Let the input signal matrix S have the SVD $\Phi V \Psi^\dagger$ then the channel can be written as

$$X = \sqrt{\frac{\rho}{M}}\Phi V \Psi^\dagger H + W.$$

Now, consider a new signal S_1 formed by multiplying Φ and V and let X_1 be the corresponding received signal. Then

$$X_1 = \sqrt{\frac{\rho}{M}}\Phi V H + W.$$

Note that X_1 and X have exactly the same statistics since $p(\Psi^\dagger H) = p(H)$. Therefore, one might as well send ΦV instead of $\Phi V \Psi^\dagger$. ■

Corollary 4.1. *If $M > T$ then power should be transmitted only through T of the antennas.*

Proof: Note that the V in the signal transmitted, ΦV is $T \times M$. It means that $V = [V_T | \mathbf{0}]$ where V_T is $T \times T$ and $\mathbf{0}$ is $T \times (T - M)$. That means

$$X = \sqrt{\frac{\rho}{M}} \Phi V_T H_T + W$$

where H_T is the matrix of first T rows in H. That means power is transmitted via only through T transmit antennas instead of through all M. ■

Even though the power is transmitted only through the first T transmit antennas when $M > T$ the capacity corresponding to this case is not the same as the case when we have $M = T$ antennas. This is one drawback of the model. The model obtained by starting with T antennas in the first place is not consistent with the model as starting with $M > T$ antennas and then discarding $M - T$ of them. To avoid this inconsistency we will assume $M \leq T$ from now on.

In the case of Rayleigh Fading however there is no such inconsistency and Lemma 4.1 gives rise to a stronger result [50]

Theorem 4.1. *For any coherence interval T and any number of receiver antennas, the capacity obtained with $M > T$ transmitter antennas is the same as the capacity obtained with $M = T$ antennas.*

4.3 Properties of Capacity Achieving Signals

Marzetta and Hochwald [50] have established several results for the case of a purely Rayleigh fading channel whose proofs were based only on the fact that the conditional probability density satisfies Properties 1 and 2 as stated in section 5.2. Therefore, the results of [50] are also applicable to the case of the isotropically random Rician fading channel discussed in this chapter.

Assume the power constraint $E[\mathrm{tr}\{SS^\dagger\}] \leq TM$.

Lemma 4.2. *: Suppose that S has a probability density $p_0(S)$ that generates some mutual information I_0. Then, for any $M \times M$ unitary matrix Ψ and for any $T \times T$ unitary matrix Φ, the "rotated" probability density, $p_1(S) = p_0(\Phi^\dagger S \Psi)$, also generates I_0.*

Proof: (For more details refer to the proof for Lemma 1 in [50].) The proof hinges on the fact that Jacobian determinant of any unitary transformation is one, $p(\Phi X|\Phi S) = p(X|S)$, $p(X|S\Psi^\dagger) = p(X|S)$ and $E[\mathrm{tr}\{SS^\dagger\}]$ is invariant to pre- and post-multiplication of S by unitary matrices. ■

Lemma 4.3. *: For any transmitted signal probability density $p_0(S)$, there is a probability density $p_1(S)$ that generates at least as much mutual information and is unchanged by rearrangements of the rows and columns of S.*

Proof: (For more details refer to the proof for Lemma 2 in [50].) From Lemma 4.2 it is evident any density obtained from the original density on S, $p_0(S)$ by pre- and post-multiplying S by any arbitrary permutation matrices P_{Tk}, $k = 1, \ldots, T!$ (there are $T!$ permutations of the rows) and P_{Ml}, $l = 1, \ldots, M!$ (there are $M!$ permutations of the columns), generates the same mutual information. Since mutual information is a concave functional of the input signal density a mixture input density, $p_1(S)$ formed by taking the average over all densities obtained by permuting S generates a mutual information at least as large as that of the original density. Note that the new mixture density satisfies the same power constraint as the original density since

$E[\mathrm{tr}\{SS^\dagger\}]$ is invariant to permutations of S. ■

Corollary 4.2. *: The following power constraints all yield the same capacity of the isotropically random Rician fading channel.*

- $E|s_{tm}|^2 = 1, \quad m = 1, \ldots, M, \; t = 1, \ldots, T$
- $\frac{1}{M}\sum_{m=1}^{M} E|s_{tm}|^2 = 1, \quad t = 1, \ldots, T$
- $\frac{1}{T}\sum_{t=1}^{T} E|s_{tm}|^2 = 1, \quad m = 1, \ldots, M$
- $\frac{1}{TM}\sum_{t=1}^{T}\sum_{m=1}^{M} E|s_{tm}|^2 = 1$

Basically, the corollary tells us that many different types of power constraints result in the same channel capacity.

Theorem 4.2. *The signal matrix that achieves capacity can be written as $S = \Phi V$, where Φ is a $T \times T$ isotropically distributed unitary matrix, and V is an independent $T \times M$ real, nonnegative, diagonal matrix. Furthermore, we can choose the joint density of the diagonal elements of V to be unchanged by rearrangements of its arguments.*

Proof: Proof is similar to the proof for Theorem 2 in [50]. ■

4.4 Capacity Upper and Lower Bounds

First we will state the following result which has already been established in [50], and follows in a straightforward manner from [73].

Theorem 4.3. *The expression for capacity of the isotropically random Rician fading channel when only the receiver has complete knowledge about the channel (informed receiver uninformed transmitter) is*

$$C_H = T \cdot E \log\det\left[I_N + \frac{\rho}{M}H^\dagger H\right].$$

The following general upper bound on capacity is quite intuitive [10, 50].

Proposition 4.2. *An upper bound on capacity of the isotropically random Rician fading channel when neither the transmitter nor the receiver has any knowledge about the channel (uninformed receiver uninformed transmitter) is given by*

$$C \leq C_H = T \cdot E \log \det \left[I_N + \frac{\rho}{M} H^\dagger H \right]. \tag{4.2}$$

What is lacking is a tractable lower bound on capacity of the isotropically random Rician fading channel. In this work we establish such a lower bound when no channel information is present at either the transmitter or the receiver.

Theorem 4.4. *A lower bound on capacity of the isotropically random Rician fading channel when neither the transmitter nor the receiver has any knowledge about the channel is given by*

$$\begin{aligned} C &\geq TE\left[\log_2 \det\left(I_N + \frac{\rho}{M} H^\dagger H\right)\right] - NE\left[\log_2 \det\left(I_T + \frac{\rho}{M} SS^\dagger\right)\right] \quad (4.3)\\ &\geq TE\left[\log_2 \det\left(I_N + \frac{\rho}{M} H^\dagger H\right)\right] - NM \log_2(1 + \frac{\rho}{M} T). \quad (4.4) \end{aligned}$$

Proof: First note that the capacity C is given by

$$C = \max_{p(S)} [I(X; S) = \mathcal{H}(X) - \mathcal{H}(X|S)]$$

by choosing a specific distribution on S, in this case $\mathcal{CN}(0, 1)$, we get a lower bound to C. Note that $\mathcal{H}(X|H) \leq \mathcal{H}(X)$ so that we obtain

$$C \geq \mathcal{H}(X|H) - \mathcal{H}(X|S). \tag{4.5}$$

Since $p(S) = \frac{1}{\pi^{TM}} \exp\left(-\text{tr}\{SS^\dagger\}\right)$ (the elements of S are $\mathcal{CN}(0, 1)$ random variables)

$$\mathcal{H}(X|H) = TE\left[\log_2\left((\pi e)^N \det\left(I_N + \frac{\rho}{M} H^\dagger H\right)\right)\right].$$

Now we turn to evaluating $\mathcal{H}(X|S)$. Note that since $p(\Psi^\dagger H) = p(H)$ for all unitary matrices Ψ, $E[h_{m_1n_1}h^*_{m_2n_2}] = \delta_{m_1m_2}\delta_{n_1n_2}$. Therefore given S, X has covariance given by

$$E[x_{t_1n_1}x^*_{t_2n_2}|S] = \delta_{n_1n_2} \cdot \left[\delta_{t_1t_2} + \frac{\rho}{M}\sum_{m=1}^{M} s_{t_1m}s^*_{t_2m}\right].$$

Since $\mathcal{H}(X|S)$ is bounded above by the entropy $\mathcal{H}_G(X|S)$ of a Gaussian with the same mean and covariance as X given S we have

$$\mathcal{H}_G(X|S) \leq NE\left[\log_2\left((\pi e)^T \det\left(I_T + \frac{\rho}{M}SS^\dagger\right)\right)\right]$$

where the expectation is over the distribution of S which gives us (4.3).

Next we simplify the expression above further to obtain a looser lower bound. We use the property that for any $T \times M$ matrix S, $\det(I_T + SS^\dagger) = \det(I_M + S^\dagger S)$. This, along with the fact that $\log_2 \det(K)$ is convex cap and Jensen's inequality, gives

$$\begin{aligned} NE\left[\log_2 \det\left(I_T + \frac{\rho}{M}SS^\dagger\right)\right] &\leq N\log_2 \det\left(I_M + \frac{\rho}{M}E[S^\dagger S]\right) && (4.6) \\ &= NM\log_2(1 + \frac{\rho}{M}T). && (4.7) \end{aligned}$$

Therefore, we obtain (4.4). ■

Note that the expression on the right hand side in (4.5) can be rewritten as $I(X;S|H) - I(X;H|S)$. $I(X;S|H)$ is the mutual information when H is known to the receiver and $I(X;H|S)$ is the information contained in X about H when S is known at the receiver and therefore, $I(X;H|S)$ can be viewed as the penalty for learning H in the course of decoding S from the reception of X.

A simple improved lower bound in either (4.3) or (4.4) can be obtained by optimizing over the number of transmit and receive antennas used. Note that both the expressions in the right hand side of (4.3) can be easily evaluated using Monte Carlo simulations.

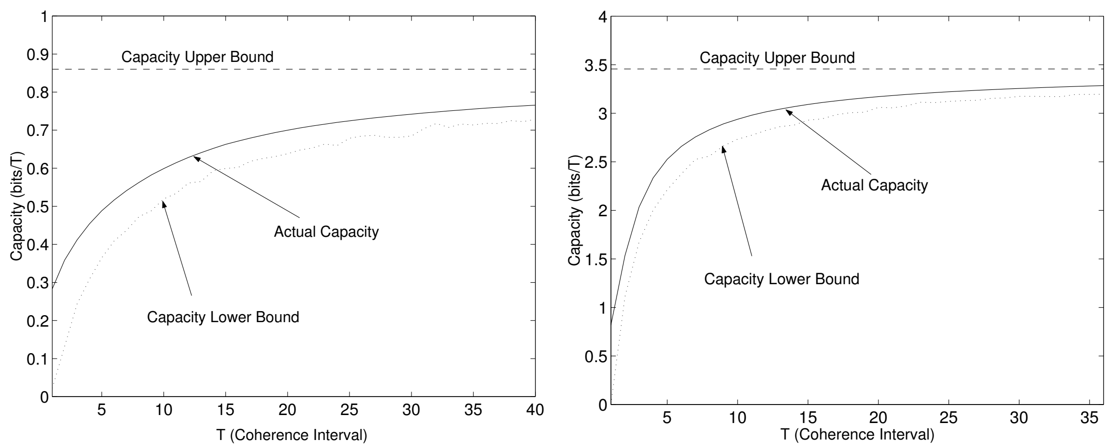
From the lower bound (4.4) it can be easily seen that as $T \to \infty$ the normalized capacity, $C_n = C/T$ in bits per channel use is given by that of the informed receiver

$$C_n = E \log\det \left[I_N + \frac{\rho}{M} H^\dagger H \right].$$

This was conjectured for the Rayleigh fading channel by Marzetta and Hochwald in [50] and discussed in [10, p. 2632].

4.5 Numerical Results

First we show the utility of the lower bound by comparing it with the actual capacity curve calculated in [50] for the case of a single transmit and receive antennas in a purely Rayleigh fading environment for increasing T. We plot the upper bound (4.2), the lower bound derived (4.3) and the actual capacity obtained in [50] from numerical integration in Figure 4.1. The utility of this bound is clearly evident from how well the lower bound follows the actual capacity.

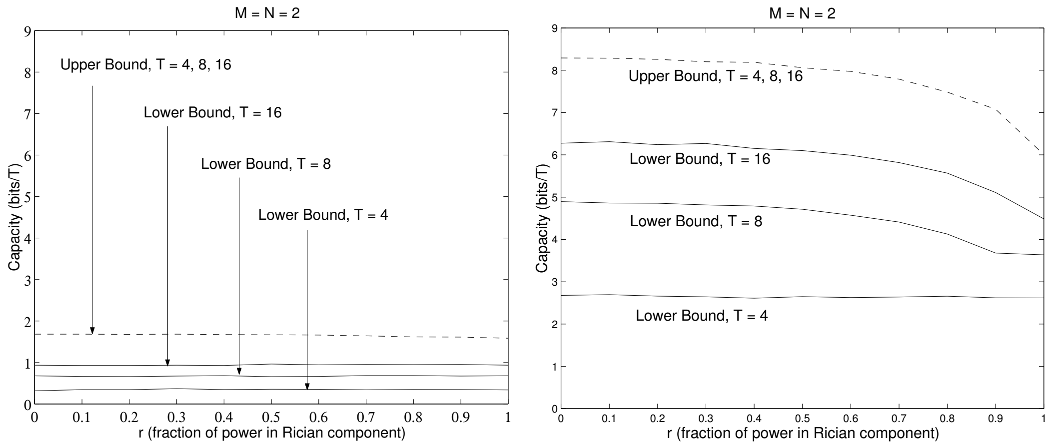


(a) $\rho = 0dB$ (b) $\rho = 12dB$

Figure 4.1: Capacity and Capacity lower bound for $M = N = 1$ as $T \to \infty$

In different simulations, using the upper and lower bounds, we found that the capacity variation as a function of r, i.e. as the channel moves from a purely Rayleigh

($r = 0$) to a purely specular channel ($r = 1$)become more pronounced as we move to higher SNR regimes and larger number of transmit and receive antennas, M and N respectively. Examples of this behavior are shown in Figures 4.2 and 4.3. Note that for the special case of $M = N = 1$ the capacity curves have an upward trend as r varies from 0 to 1. For all other cases the capacity curves have a downward trend, where the reduction in capacity becomes significant only for $r > 1/2$.



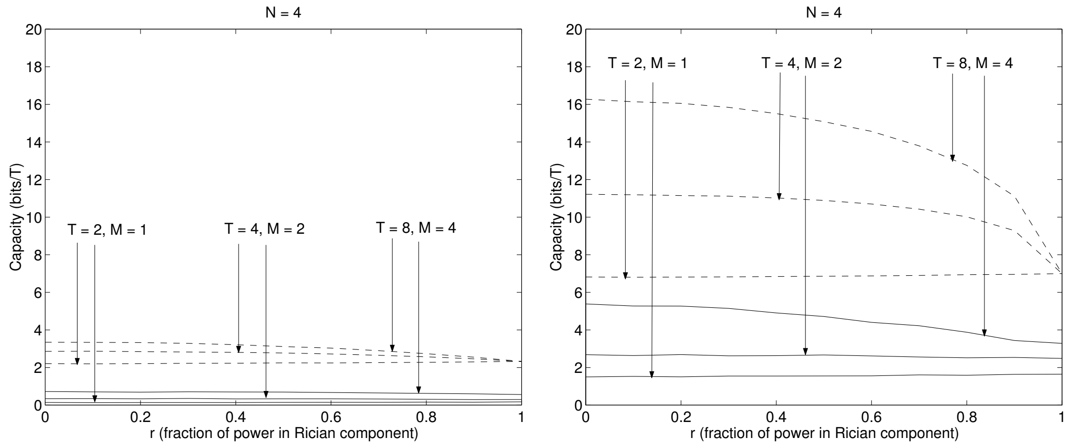
(a) $\rho = 0dB$ (b) $\rho = 15dB$

Figure 4.2: Capacity upper and lower bounds as the channel moves from purely Rayleigh to purely Rician fading

4.6 Analysis for High SNR

The techniques developed in [85] can be easily applied to the model in this section to conclude that the number of degrees of freedom is given by $M^*(T - M^*)$ where $M^* = \min\{M, N, T/2\}$. All the Theorems developed in [85] can be easily translated to the case of isotropically random Rician fading.

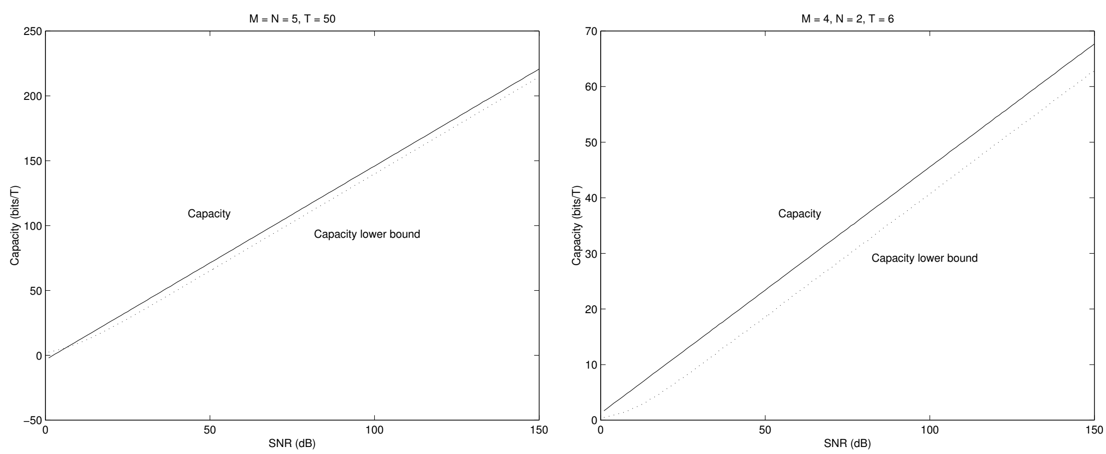
We will now investigate whether the lower bound derived in this chapter attains the required number of degrees of freedom. We will compare this lower bound with the capacity expression derived in [85] for high SNR. Figure 4.4 shows the actual



(a) $\rho = 0dB$ (b) $\rho = 15dB$

Figure 4.3: Capacity upper and lower bounds as the channel moves from purely Rayleigh to purely Rician fading

capacity and the lower bound as a function of increasing SNR for a purely Rayleigh fading channel. In both cases $T > M^* + N$ where $M^* = \min\{M, N, T/2\}$ as the actual capacity expression is not known for $T < M^* + N$ [85]. As can be seen this lower bound attains the required number of degrees of freedom.



(a) $M = N = 5, T = 50$ (b) $M = 4, N = 2, T = 6$

Figure 4.4: Capacity and capacity lower bound as a function of SNR for a purely Rayleigh fading channel

4.7 Conclusions

We have analyzed the case of isotropically random Rician fading channel. We have proposed a tractable model for Rician fading with a stochastic isotropic specular component of rank one. Using this model we were able to establish most of the results previously obtained in the case of Rayleigh fading. We were also able to derive a lower bound on capacity for any number of transmit and receive antennas. This lower bound is also applicable to the case of Rayleigh fading. For single transmit and receive antennas, the Rician channel gives superior performance with respect to the Rayleigh channel. For multiple antenna channels, Rician fading tends to degrade the performance. Our numerical results indicate that the Rayleigh model is surprisingly robust: under our Rician model, up to half of the received energy can arrive via the specular component without significant reduction in capacity compared with the purely Rayleigh case. Finally, the model considered in this chapter has a drawback as explained after Lemma 4.1. Inspite of the drawback the model is helpful in understanding the Rician channel.

CHAPTER 5

Min-Capacity: Rician Fading, Unknown Static Specular Component

5.1 Introduction

In the previous chapter we considered an isotropic model for the case of Rician fading where the fading channel consists of a Rayleigh component, modeled as in [50] and an independent rank-one isotropically distributed specular component. The fading channel was assumed to remain constant over a block of T consecutive symbol periods but take a completely independent realization over each block. We derived similar results on optimal capacity achieving signal structures as in [50]. We also established a lower bound to capacity that can be easily extended to the model considered in this chapter. The model described in the previous chapter is applicable to a mobile-wireless link where both the direct line of sight component (specular component) and the diffuse component (Rayleigh component) change with time.

In this chapter, we consider MIMO channel capacity, under the average energy constraint on the input signal, of a quasi-static Rician model where the specular component is non-changing while the Rayleigh component is varying over time. This model is similar to the traditional model where the specular component is deterministic and persists over time. The model is applicable to the case where the transmitter and receiver are fixed in space, are in motion, or the propagation medium is changing where the transmitter and the receiver are sufficiently far apart so that the specular component is practically constant while the diffuse multipath component changes rapidly. If the specular component were known to both the transmitter and the receiver then the signaling scheme as well as the capacity would depend on the specific realization of the specular component. We however deal with the case when the transmitter has no knowledge about the specular component. In this scenario the transmitter has to treat the specular component as a random variable with a prior distribution. There are two approaches the transmitter can take. It can either ignore the distribution, treat the channel as an arbitrarily varying channel and maximize the worst case rate over the ensemble of values that the specular component can take on or take into account the prior distribution of the specular component and then maximize the average rate. We address both approaches in this chapter.

Similarly to [21] the specular component is an outer product of two vectors of unit magnitude that are non-changing and unknown to the transmitter but known to the receiver. The difference between our approach and that of [21] is that in [21] the authors consider the channel to be known completely to the receiver. We assume that the receiver's extent of knowledge about the channel is limited to the specular component. That is, the receiver has no knowledge about the Rayleigh component of the model. Considering the absence of knowledge at the transmitter it is important to design a signal scheme that guarantees the largest overall rate for communication irrespective of the value of the specular component. This is formulated as the problem of determining the worst case capacity called min-capacity, in section 5.2. In section 5.5 we consider the average capacity instead of worst case capacity and show that both formulations imply the same optimal signal structure and the same maximum possible rate. Throughout this chapter we assume that the transmitter has no knowledge about the channel though it has knowledge of the statistics of the Rayleigh component of the channel. Only in 5.5 do we assume that the channel has knowledge of the statistics of the specular component as well.

5.2 Signal Model and Min-Capacity Criterion

Let there be M transmit antennas and N receive antennas. We assume that the fading coefficients remain constant over a block of T consecutive symbol periods but are independent from block to block. Keeping that in mind, we model the channel as carrying a $T \times M$ signal matrix S over a $M \times N$ MIMO channel H, producing X at the receiver according to the model:

$$X = \sqrt{\frac{\rho}{M}} SH + W \tag{5.1}$$

where the elements, w_{tn} of W are independent circular complex Gaussian random variables with mean 0 and variance 1 ($\mathcal{CN}(0,1)$).

The MIMO Rician model for the matrix H is $H = \sqrt{1-r}G + \sqrt{rNM}\alpha\beta^\dagger$ where G consists of independent $\mathcal{CN}(0,1)$ random variables and α and β are deterministic vectors of length M and N, respectively, such that $\alpha^\dagger\alpha = 1$ and $\beta^\dagger\beta = 1$. We assume α and β are known to the receiver. Since the receiver is free to apply a co-ordinate transformation by post multiplying X by a unitary matrix, without loss of generality we can take β to be identically equal to $[1\ 0\ \ldots\ 0]^\tau$. We will sometimes write H as H_α to highlight the dependence of H on α. G remains constant for T symbol periods and takes on a completely independent realization every T^{th} symbol period.

The problem we are investigating is to find the distribution $p^m(S)$ that attains the maximum in the following maximization defining the worst case channel capacity

$$C^m = \max_{p(S)} I^m(X;S) = \max_{p(S)} \inf_{\alpha \in A} I^\alpha(X;S)$$

and also to find the maximum value, C^m.

$$I^{\alpha}(X;S) = \int p(S)p(X|S,\alpha\beta^{\dagger}) \log \frac{p(X|S,\alpha\beta^{\dagger})}{\int p(S)p(X|S,\alpha\beta^{\dagger})\ dS}\ dSdX$$

is the mutual information between X and S when the specular component is given by $\alpha\beta^{\dagger}$ and $A \stackrel{\text{def}}{=} \{\alpha : \alpha \in \mathcal{C}^M \text{ and } \alpha^{\dagger}\alpha = 1\}$. Note that since $\beta = [1\ 0\ \ldots\ 0]^{\tau}$ without any loss of generality we can write $p(X|S,\alpha\beta^{\dagger})$ as simply $p(X|S,\alpha)$ and is given by

$$p(X|S,\alpha) = \frac{e^{-\mathrm{tr}\{[I_T+(1-r)\frac{\rho}{M}SS^{\dagger}]^{-1}(X-\sqrt{\rho rN}S\alpha\beta^{\dagger})(X-\sqrt{\rho rN}S\alpha\beta^{\dagger})^{\dagger}\}}}{\pi^{TN}\det^N[I_T+(1-r)\frac{\rho}{M}SS^{\dagger}]}.$$

Since A is compact the "inf" in the problem can be replaced by "min". For convenience we will refer to $I^m(X;S)$ as the *min-mutual information* and C^m as *min-capacity*.

The *min-capacity* defined above is just the capacity of a *compound channel*. For more information on the concept of compound channels, worst case capacities and the corresponding coding theorems refer to [14, chapter 5, pp. 172-178]. It is shown that [14, Prob. 13, p. 183] *min-capacity* doesn't depend on the receiver's knowledge of the channel. Hence, it is not necessary for us to assume that the specular component is known to the receiver. However, we do so because it facilitates computation of *avg-capacity* in terms of the conditional probability distribution $p(X|S)$.

5.3 Capacity Upper and Lower Bounds

Theorem 5.1. *Min-capacity, C^m_H of the quasi-static Rician fading model when the channel matrix H is known to the receiver is given by*

$$C^m_H = TE\log\det\left[I_N + \frac{\rho}{M}H^{\dagger}_{e_1}H_{e_1}\right] \tag{5.2}$$

where $e_1 = [1\ 0\ \ldots\ 0]^{\tau}$ is a unit vector in $\mathcal{C}^M$. Note that e_1 in (5.2) can be replaced by any $\alpha \in A$ without changing the answer.

Proof: First we note that for $T > 1$, given H the channel is memoryless and hence the columns of the input signal matrix S are independent of each other. That means the mutual information $I^\alpha(X;S) = \sum_{i=1}^T I^\alpha(X_i;S_i)$ where X_i and S_i denote the i^{th} row of X and S, respectively. The maximization over each term can be done separately and it is easily seen that each term will be maximized individually for the same density on S_i. That is $p(S_i) = p(S_j)$ for $i \neq j$ and $\max_{p(S)} I^\alpha(X;S) = T \max_{p(S_1)} I^\alpha(X_1;S_1)$. Therefore, WLOG we assume $T = 1$.

Given H the channel is an AWGN channel therefore, capacity is attained by Gaussian signal vectors. Let Λ_S be the input signal covariance. Since the transmitter doesn't know α, Λ_S can not depend on α and the *min-capacity* is given by

$$\max_{\Lambda_S:\mathrm{tr}\{\Lambda_S\}\leq M} \mathcal{F}(\Lambda_S) = \max_{\Lambda_S:\mathrm{tr}\{\Lambda_S\}\leq M} \min_{\alpha \in A} E \log\det\left[I_N + \frac{\rho}{M} H_\alpha^\dagger \Lambda_S H_\alpha\right] \tag{5.3}$$

where $\mathcal{F}(\Lambda_S)$ is implicitly defined in an obvious manner. First note that $\mathcal{F}(\Lambda_S)$ in (5.3) is a concave function of Λ_S (This follows from the fact that $\log\det K$ is a concave function of K). Also, $\mathcal{F}(\Psi^\dagger \Lambda_S \Psi) = \mathcal{F}(\Lambda_S)$ for any $M \times M$ $\Psi : \Psi^\dagger\Psi = I_M$ since $\Psi^\dagger \alpha \in A$ for any $\alpha \in A$ and G has i.i.d. zero mean complex Gaussian entries. Let $Q^\dagger D Q$ be the SVD of Λ_S then we have $\mathcal{F}(D) = \mathcal{F}(Q^\dagger D Q) = \mathcal{F}(\Lambda_S)$. Therefore, we can choose Λ_S to be diagonal. Moreover, $\mathcal{F}(P_k^\dagger \Lambda_S P_k) = \mathcal{F}(\Lambda_S)$ for any permutation matrix P_k, $k = 1, \ldots, M!$. Therefore, if we choose $\Lambda_S' = \frac{1}{M!}\sum_{k=1}^{M!} P_k^\dagger \Lambda_S P_k$ then by concavity and Jensen's inequality we have

$$\mathcal{F}(\Lambda_S') \geq \sum_{k=1}^{M!} \mathcal{F}(P_k^\dagger \Lambda_S P_k) = \mathcal{F}(\Lambda_S).$$

Therefore, we conclude that the maximizing input signal covariance Λ_S is a multiple of the identity matrix. To maximize the expression in (5.3) we need to choose $\mathrm{tr}\{\Lambda_S\} = M$ or $\Lambda_S = I_M$ and since $E \log\det[I_N + \frac{\rho}{M} H_{\alpha_1}^\dagger H_{\alpha_1}] = E \log\det[I_N + \frac{\rho}{M} H_{\alpha_2}^\dagger H_{\alpha_2}]$ for any $\alpha_1, \alpha_2 \in A$, (5.2) easily follows. ∎

By the data processing theorem additional information at the receiver doesn't decrease capacity. Therefore:

Proposition 5.1. *An upper bound on the channel min-capacity of the quasi-static Rician fading channel when the receiver has no knowledge of G is given by*

$$C^m \leq T \cdot E \log\det\left[I_N + \frac{\rho}{M} H_{e_1}^{\dagger} H_{e_1}\right]. \tag{5.4}$$

Now, we establish a lower bound.

Proposition 5.2. *A lower bound on min-capacity of the quasi-static Rician fading channel when the the receiver has no knowledge about G is given by*

$$\begin{aligned} C^m &\geq C_H^m - NE\left[\log_2 \det\left(I_T + (1-r)\frac{\rho}{M} SS^{\dagger}\right)\right] &(5.5)\\ &\geq C_H^m - NM\log_2(1 + (1-r)\frac{\rho}{M}T). &(5.6)\end{aligned}$$

Proof: Proof is similar to that of Theorem 4.4 and won't be repeated here. ∎

Notice that the second term in right hand side of the lower bound is

$$NE\left[\log_2 \det\left(I_T + (1-r)\frac{\rho}{M} SS^{\dagger}\right)\right]$$

instead of $NE\left[\log_2 \det\left(I_T + \frac{\rho}{M} SS^{\dagger}\right)\right]$ which occurs in the lower bound derived for the model in the previous chapter. Recall that we have seen how the second term can be viewed as a penalty term for using part of the available rate for training in order to learn the channel. When $r = 1$ or when the channel is purely specular we see that the penalty term for training goes to zero. This makes perfect sense because the specular component is known to the receiver and the penalty for learning the specular component is zero in the current model as contrasted to the model in the previous chapter.

Combining (5.4) and (4.4) gives us the following

Corollary 5.1. *The normalized min-capacity, $C_n^m = C^m/T$ in bits per channel use as $T \to \infty$ of the quasi-static Rician fading channel when the receiver has no knowledge of G is given by*

$$C_n^m = E \log \det \left[I_N + \frac{\rho}{M} H_{e_1}^\dagger H_{e_1} \right].$$

Note that this is same as the capacity when the receiver knows H, so that as $T \to \infty$ perfect channel estimation can be performed.

5.4 Properties of Capacity Achieving Signals

In this section, we derive the optimum signal structure for achieving *min-capacity*. The optimization is done under the average power constraint $E[\mathrm{tr}\{SS^\dagger\}] \leq TM$.

Lemma 5.1. *$I^m(X; S)$, for the quasi-static Rician fading channel when the receiver has no knowledge of G, as a functional of $p(S)$ is concave in $p(S)$.*

Proof: First we note that $I^\alpha(X; S)$ is a concave functional of $p(S)$ for every $\alpha \in A$. Let $I^m(X; S)_{p(S)}$ denote $I^m(X; S)$ evaluated using $p(S)$ as the signal density. Then,

$$\begin{aligned} I^m(X;S)_{\delta p_1(S)+(1-\delta)p_2(S)} &= \min_{\alpha \in A} I^\alpha(X;S)_{\delta p_1(S)+(1-\delta)p_2(S)} \\ &\geq \min_{\alpha \in A}[\delta I^\alpha(X;S)_{p_1(S)} + (1-\delta) I^\alpha(X;S)_{p_2(S)}] \\ &\geq \delta \min_{\alpha \in A} I^\alpha(X;S)_{p_1(S)} + (1-\delta) \min_{\alpha \in A} I^\alpha(X;S)_{p_2(S)} \\ &= \delta I^m(X;S)_{p_1(S)} + (1-\delta) I^m(X;S)_{p_2(S)}. \end{aligned}$$

■

Lemma 5.2. *For any $T \times T$ unitary matrix Φ and any $M \times M$ unitary matrix Ψ, if $p(S)$ generates $I^m(X; S)$ then so does $p(\Phi S \Psi^\dagger)$ when the fading is quasi-static Rician fading and the receiver has no knowledge of the Rayleigh component.*

Proof: 1) Note that $p(\Phi X|\Phi S) = p(X|S)$, therefore $I^\alpha(X;\Phi S) = I^\alpha(X;S)$ for any $T \times T$ unitary matrix Φ and all $\alpha \in A$.

2) We have, $\Psi\alpha \in A$ for any $\alpha \in A$ and any $M \times M$ unitary matrix Ψ. Therefore, if $I^m(X;S)$ achieves its minimum value at $\alpha_0 \in A$ then $I^m(X;S\Psi^\dagger)$ achieves its minimum value at $\Psi\alpha_0$ because $I^\alpha(X;S) = I^{\Psi\alpha}(X;S\Psi^\dagger)$ for $\alpha \in A$ and Ψ an $M \times M$ unitary matrix.

Combining 1) and 2) we get the lemma. ∎

Lemma 5.3. *The min-capacity achieving signal distribution, $p(S)$ when G is not known at the receiver for the quasi-static Rician fading channel is unchanged by any pre- and post- multiplication of S by unitary matrices of appropriate dimensions.*

Proof: We will show that for any signal density $p_0(S)$ generating *min-mutual information* I_0^m there exists a density $p_1(S)$ generating $I_1^m \geq I_0^m$ such that $p_1(S)$ is invariant to pre- and post- multiplication of S by unitary matrices of appropriate dimensions. By Lemma 5.2, for any pair of permutation matrices, Φ ($T \times T$) and Ψ ($M \times M$) $p_0(\Phi S \Psi^\dagger)$ generates the same *min-mutual information* as $p(S)$. Define $u_T(\Phi)$ to be the isotropically random unitary density function of a $T \times T$ unitary matrix Φ. Similarly define $u_M(\Psi)$. Let $p_1(S)$ be a mixture density given as follows

$$p_1(S) = \int\int p_0(\Phi S \Psi^\dagger) u(\Phi) u(\Psi)\ d\Phi d\Psi.$$

It is easy to see that $p_1(S)$ is invariant to any pre- and post- multiplication of S by unitary matrices and if I_1^m is the *min-mutual information* generated by $p_1(S)$ then from Jensen's inequality and concavity of $I^m(X;S)$ we have $I_1^m \geq I_0^m$. ∎

Corollary 5.2. *$p^m(S)$, the optimal min-capacity achieving signal density, for the quasi-static Rician fading channel when the receiver has no knowledge of G, lies in*

$\mathcal{P} = \cup_{I>0}\mathcal{P}_I$ *where*

$$\mathcal{P}_I = \{p(S) : I^\alpha(X;S) = I \quad \forall \alpha \in A\}. \tag{5.7}$$

Proof: Follows immediately from Lemma 5.3 because any signal density that is invariant to pre- and post- multiplication of S by unitary matrices generates the same mutual information $I^\alpha(X;S)$ irrespective of the value of α. ■

Corollary 5.3. *The min-capacity achieving signal distribution $p(S)$ for the quasi static Rician fading channel when the receiver does not know G is unchanged by rearrangements of elements in S.*

Corollary 5.4. *: The following power constraints all yield the same channel min-capacity of the quasi-static Rician fading channel when the receiver does not know G.*

- $E|s_{tm}|^2 = 1, \quad m = 1, \ldots, M, \; t = 1, \ldots, T$
- $\frac{1}{M}\sum_{m=1}^{M} E|s_{tm}|^2 = 1, \quad t = 1, \ldots, T$
- $\frac{1}{T}\sum_{t=1}^{T} E|s_{tm}|^2 = 1, \quad m = 1, \ldots, M$
- $\frac{1}{TM}\sum_{t=1}^{T}\sum_{m=1}^{M} E|s_{tm}|^2 = 1$

Theorem 5.2. *The signal matrix that achieves min-capacity, for the quasi-static Rician fading channel when the Rayleigh component is not known at the receiver, can be written as $S = \Phi V \Psi^\dagger$, where Φ and Ψ are $T \times T$ and $M \times M$ isotropically distributed matrices independent of each other, and V is a $T \times M$ real, nonnegative, diagonal matrix, independent of both Φ and Ψ.*

Proof: From the singular value decomposition (SVD) we can write $S = \Phi V \Psi^\dagger$, where Φ is a $T \times T$ unitary matrix, V is a $T \times M$ nonnegative real diagonal matrix, and Ψ is an $M \times M$ unitary matrix. In general, Φ, V and Ψ are jointly distributed. Suppose S has probability density $p_0(S)$ that generates *min-mutual information* I_0^m. Let Θ_1 and Θ_2 be isotropically distributed unitary matrices of size $T \times T$ and $M \times M$ independent of S and of each other. Define a new signal $S_1 = \Theta_1 S \Theta_2^\dagger$, generating *min-mutual information* I_1^m. Now conditioned on Θ_1 and Θ_2, the *min-mutual information* generated by S_1 equals I_0^m. From the concavity of the *min-mutual information* as a functional of $p(S)$, and Jensen's inequality we conclude that $I_1^m \geq I_0^m$.

Since Θ_1 and Θ_2 are isotropically distributed $\Theta_1\Phi$ and $\Theta_2\Psi$ are also isotropically distributed when conditioned on Φ and Ψ respectively. This means that both $\Theta_1\Phi$ and $\Theta_2\Psi$ are isotropically distributed making them independent of Φ, V and Ψ. Therefore, S_1 is equal to the product of three independent matrices, a $T \times T$ unitary matrix Φ, a $T \times M$ real nonnegative matrix V and an $M \times M$ unitary matrix Ψ.

Now, we will show that the density $p(V)$ on V is unchanged by rearrangements of diagonal entries of V. There are $\min\{M!, T!\}$ ways of arranging the diagonal entries of V. This can be accomplished by pre- and post-multiplying V by appropriate permutation matrices, P_{Tk} and P_{Mk}, $k = 1, \ldots, \min\{M!, T!\}$. The permutation does not change the *min-mutual information* because ΦP_{Tk} and ΨP_{Mk} have the same density functions as Φ and Ψ. By choosing an equally weighted mixture density for V, involving all $\min\{M!, T!\}$ arrangements we obtain a higher value of *min-mutual information* because of concavity and Jensen's inequality. This new density is invariant to the rearrangements of the diagonal elements of V. ■

5.5 Average Capacity Criterion

In this case we maximize $I_E(X; S) = E_\alpha[I^\alpha(X; S)]$, where I^α is as defined earlier and E_α denotes expectation over $\alpha \in A$ under the assumption that all α are equally likely. That is, under the assumption that α is unchanging over time, isotropically random and known to the receiver. Please note that this differs from the model considered in Chapter 4 where we consider the case of a piecewise constant, time varying, i.i.d. specular component.

Therefore, the problem can be stated as finding $p_E(S)$ the probability density function on the input signal S that achieves the following maximization

$$C_E = \max_{p(S)} E_\alpha[I^\alpha(X;S)] \tag{5.8}$$

and also to find the value C_E. We will refer to $I_E(X;S)$ as *avg-mutual information* and C_E as *avg-capacity.*

We will show that the signal density $p^m(S)$ that attains C^m also attains C_E. For this we need to establish the following Lemmas.

Lemma 5.4. *$I_E(X;S)$ for the quasi-static Rician fading channel is a concave function of the signal density $p(S)$ when the receiver does not know G and the transmitter takes into account the statistics of the specular component.*

Proof: First we note that $I^\alpha(X;S)$ is a concave functional of $p(S)$ for every $\alpha \in A$. Then,

$$\begin{aligned}
I_E(X;S)_{\delta p_1(S)+(1-\delta)p_2(S)} &= E_\alpha[I^\alpha(X;S)_{\delta p_1(S)+(1-\delta)p_2(S)}] \\
&\geq E_\alpha[\delta I^\alpha(X;S)_{p_1(S)} + (1-\delta)I^\alpha(X;S)_{p_2(S)}] \\
&= \delta E_\alpha[I^\alpha(X;S)_{p_1(S)}] + (1-\delta)E_\alpha[I^\alpha(X;S)_{p_2(S)}] \\
&= \delta I_E(X;S)_{p_1(S)} + (1-\delta)I_E(X;S)_{p_2(S)}.
\end{aligned}$$

■

Lemma 5.5. *For any $T \times T$ unitary matrix Φ and any $M \times M$ unitary matrix Ψ, if $p(S)$ generates $I_E(X;S)$ then so does $p(\Phi S \Psi^\dagger)$ when the fading is quasi-static, the* *receiver does not know G and the transmitter takes into account the statistics of the specular component.*

Proof: We want to show if $p(S)$ generates $I_E(X;S)$ then so does $p(\Phi S\Psi^\dagger)$. Now since the density function of α, $p(\alpha) = \frac{\Gamma(M)}{\pi^M}\delta(\alpha^\dagger\alpha - 1)$ we have

$$I_E(X;S) = \frac{\pi^M}{\Gamma(M)}\int I^\alpha(X;S)\, d\alpha.$$

Note that $I^\alpha(X;\Phi S) = I^\alpha(X;S)$ and $I^{\Psi\alpha}(X;S\Psi^\dagger) = I^\alpha(X;S) \Rightarrow I^{\Psi^\dagger\alpha}(X;S) = I^\alpha(X;S\Psi^\dagger)$. Therefore,

$$\begin{aligned} I'_E(X;S) &= \frac{\pi^M}{\Gamma(M)}\int I^\alpha(X;\Phi S\Psi^\dagger)\, d\alpha \\ &= \frac{\pi^M}{\Gamma(M)}\int I^\alpha(X;S\Psi^\dagger)\, d\alpha \\ &= \frac{\pi^M}{\Gamma(M)}\int I^{\Psi^\dagger\alpha}(X;S)\, d\alpha \\ &= \frac{\pi^M}{\Gamma(M)}\int I^\omega(X;S)\, d\omega \\ &= I_E(X;S) \end{aligned}$$

where the last two equalities follow from the transformation $\omega = \Psi^\dagger\alpha$ and the fact the Jacobian of the transformation is equal to 1. ■

Lemma 5.6. *The avg-capacity achieving signal distribution, $p(S)$ is unchanged by any pre- and post- multiplication of S by unitary matrices of appropriate dimensions when the fading is quasi-static, the receiver does not know G and the transmitter takes into account the statistics of the specular component.*

Corollary 5.5. *$p_E(S)$, the optimal avg-capacity achieving signal density lies in $\mathcal{P} = \cup_{I>0}\mathcal{P}_I$ where $\mathcal{P}_I$ is as defined in (5.7) when the fading is quasi-static Rician fading, the receiver has no knowledge of G and the transmitter takes into account the statistics of the specular component.*

Based on the last corollary we conclude that for a given $p(S)$ in $\mathcal{P}$ we have $I^m(X;S) = \min_{\alpha\in A} I^\alpha(X;S) = E_\alpha[I^\alpha(X;S)] = I_E(X;S)$. Therefore, the maximizing densities for C_E and C^m are the same and also $C_E = C^m$. Therefore, designing the signal constellation with the objective of maximizing the worst case performance is not more pessimistic than maximizing the average performance.

5.6 Coding Theorem for *Min-capacity*

In this section, we will prove the following theorem.

Theorem 5.3. *For the quasi-static Rician fading model, for every $R < C^m$ there exists a sequence of $(2^{nR}, n)$ codes with codewords, m_i^n, $i = 1, \ldots, 2^{nR}$, satisfying the power constraint such that*

$$\lim_{n\to\infty} \sup_\alpha P_{e\alpha,n} = 0$$

where $P_{e\alpha,n} = \max_{i=1}^{2^{nR}} P_e(m_i^n, \alpha)$ and $P_e(m_i)$ is the probability of incorrectly decoding the messages m_i when the channel is given by H_α.

Proof: Proof follows if we can show that $P_{e\alpha,n}$ is bounded above by the same Gallager error exponent [26, 27] irrespective of the value of α. That follows from the following lemma (Lemma 5.7). ■

The intuition behind the existence of a coding theorem is that the *min-capacity* C^m, is the mutual information between the output and the input for some particular channel H_α (Since $A = \{\alpha : \alpha \in \mathbb{C}^M \text{ and } \alpha^\dagger\alpha = 1\}$ is compact). Therefore, any codes generated from the random coding argument designed to achieve rates up to C^m for that particular H_α achieve rates up to C^m for all H_α.

For Lemma 5.7, we first need to briefly describe the Gallager error exponents [26, 27] for the quasi-static Rician fading channel. For a system communicating at a rate R the upper bound on the maximum probability of error is given as follows

$$P_{e\alpha,n} \leq \exp\left(-n \max_{p(S)} \max_{0\leq\gamma\leq 1} \left[E_0(\gamma, p(S), \alpha) - \gamma R \log 2\right]\right)$$

where n is the length of the codewords in the codebook used and $E_0(\gamma, p(S), \alpha)$ is as follows

$$E_0(\gamma, p(S), \alpha) = -\log \int \left[\int p(S) p(X|S,\alpha)^{\frac{1}{1+\gamma}} dS\right]^{\gamma} dX$$

where S is the input to the channel and X is the observed output and

$$p(X|S,\alpha) = \frac{e^{-\mathrm{tr}\{[I_T+(1-r)\frac{\rho}{M}SS^\dagger]^{-1}(X-\sqrt{\rho r N}S\alpha\beta^\dagger)(X-\sqrt{\rho r N}S\alpha\beta^\dagger)^\dagger\}}}{\pi^{TN}\det^N[I_T+(1-r)\frac{\rho}{M}SS^\dagger]}.$$

where β is simply $[1\ 0\ \ldots\ 0]^\tau$. Maximization over γ in the error exponent yields a value of γ such that $\frac{\partial E_0(\gamma,p(S),\alpha)}{\partial\gamma} = R$. Note that for $\gamma = 0$, $\frac{\partial E_0(\gamma,p(S),\alpha)}{\partial\gamma} = I^\alpha(X;S)$ [26, 27] where the mutual information has been evaluated when the input is $p(S)$. If $p(S)$ is the *min-capacity* achieving density, $p^m(S)$ then $\frac{\partial E_0(\gamma,p^m(S),\alpha)}{\partial\gamma} = C^m$. For more information refer to [26, 27].

Lemma 5.7. *The $E_0(\gamma, p^m(S), \alpha)$ for the quasi-static Rician fading model is independent of α.*

Proof: First, note that

$$p^m(S) = p^m(S\Psi^\dagger)$$

for any $M \times M$ unitary matrix Ψ. Second,

$$\begin{aligned}
E_0(\gamma, p^m(S), \alpha) &= -\log \int \left[\int p^m(S) p(X|S,\alpha)^{\frac{1}{1+\gamma}} dS\right]^{\gamma} dX \\
&= -\log \int \left[\int p^m(S\Psi^\dagger) p(X|S\Psi^\dagger,\alpha)^{\frac{1}{1+\gamma}} dS\right]^{\gamma} dX \\
&= -\log \int \left[\int p^m(S) p(X|S,\Psi^\dagger\alpha)^{\frac{1}{1+\gamma}} dS\right]^{\gamma} dX \\
&= E_0(\gamma, p^m(S), \Psi^\dagger\alpha)
\end{aligned}$$

where the second equation follows from the fact that Ψ is a unitary matrix and its Jacobian is equal to 1 and the third equation follows from the fact that $p(X|S\Psi^\dagger, \alpha)^{\frac{1}{1+\gamma}} = p(X|S, \Psi^\dagger \alpha)^{\frac{1}{1+\gamma}}$. Since Ψ is arbitrary we obtain that $E_0(\gamma, p^m(S), \alpha)$ is independent of α. ■

5.7 Numerical Results

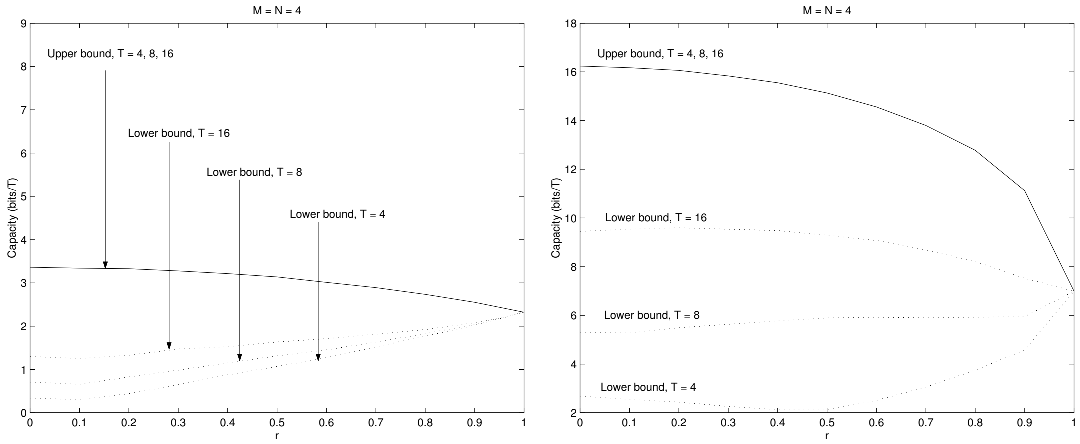
Plotting the capacity upper and lower bounds leads to similar conclusions as in the previous chapter except for the fact when $r = 1$ the upper and lower bounds on capacity coincide.

In Figure 5.1 we plot the capacity and upper lower bounds as a function of the Rician parameter r. We see that the change in capacity is not drastic for low SNR as compared to larger SNR values. Also, from Figure 5.2 we conclude that this change in capacity is more prominent for larger number of antennas. We also conclude that for a purely specular channel increasing the number of transmit antennas has no effect on the capacity. This is due to the fact that with a rank-one specular component, only beamforming SNR gains can be exploited, no diversity gains are possible.

5.8 Conclusions

We have proposed another tractable model for Rician fading channel different from the one in Chapter 4 but, along the lines of [21]. We were able to analyze this channel and derive some interesting results on optimal signal structure. We were also able to show that the optimization effort is over a much smaller set of parameters than the set originally started with. We were also able to derive a lower bound that is useful since the capacity is not in closed form.

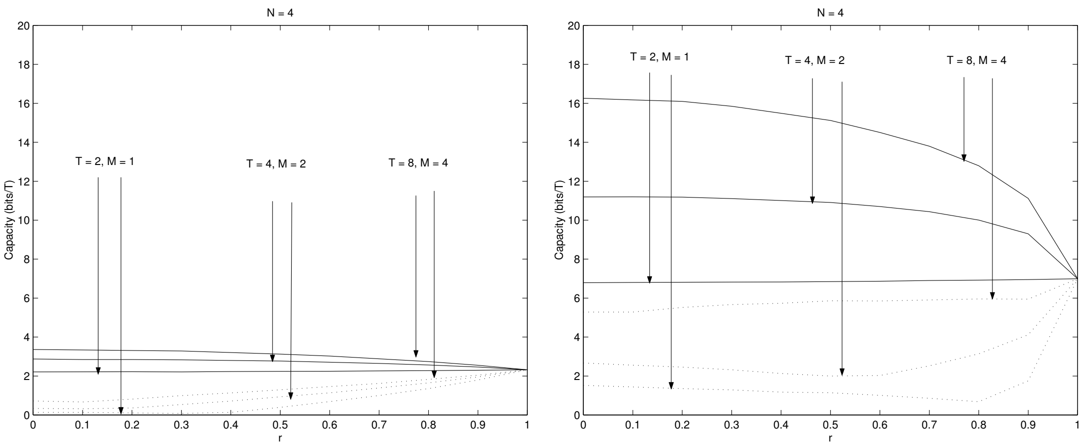
Finally, we were able to show that the approach of maximizing the worst case scenario is not overly pessimistic in the sense that the signal density maximizing the



(a) $\rho = 0dB$

(b) $\rho = 15dB$

Figure 5.1: Capacity upper and lower bounds as the channel moves from purely Rayleigh to purely Rician fading



(a) $\rho = 0dB$

(b) $\rho = 15dB$

Figure 5.2: Capacity upper and lower bounds as the channel moves from purely Rayleigh to purely Rician fading

worst case performance also maximizes the average performance and the capacity value in both formulations turns out to be the same. The average capacity being equal to the worst case capacity can also be interpreted in a different manner: we have shown that the average capacity criterion is a quality of service guaranteeing capacity.

CHAPTER 6

Capacity: Rician Fading, Known Static Specular Component

6.1 Introduction

In Chapter 4, we considered a model where the specular component is rank-one and has an isotropic distribution. In Chapter 5, we considered a model where the specular component is deterministic and static but unknown to the transmitter. Both these models led to a tractable analysis of the capacity or min-capacity of the Rician fading model. Also, the specular component considered in both the chapters was of rank one. Here we consider the Rician fading model with a fixed deterministic specular fading component.

In this chapter, we analyze the standard Rician fading model for capacity under the average energy constraint on the input signal. Throughout the chapter, we assume that the specular component is deterministic and is known to both the transmitter and the receiver. The specular component in this chapter is of general rank except in section 6.2 where it is restricted to be of rank one. Throughout this chapter the Rayleigh component is never known to the transmitter. There are some cases we consider where the receiver has complete knowledge of the channel. In such cases, the receiver has knowledge about the Rayleigh as well as the specular component whereas the transmitter has knowledge only about the specular component. The capacity when the receiver has complete knowledge about the channel will be referred to as *coherent capacity* and the capacity when the receiver has no knowledge about the Rayleigh component will be referred to as *non-coherent capacity.* This chapter is organized as follows. In section 6.2 we deal with the special case of a rank-one specular component with the characterization of coherent capacity in section 6.2.1. The case of rank greater than one is dealt with in section 6.3. The coherent capacity for this case is considered in section 6.3.1, the non-coherent capacity for low SNR in section 6.3.3 and the non-coherent capacity for high SNR in section 6.3.4. Finally, in section 6.4 we consider the performance of a Rician channel in terms of capacity when pilot symbol based training is used in the communication system.

6.2 Rank-one Specular Component

We adopt the following model for the Rician fading channel

$$X = \sqrt{\frac{\rho}{M}} SH + W \tag{6.1}$$

where X is the $T \times N$ matrix of received signals, H is the $M \times N$ matrix of propagation coefficients, S is the $T \times M$ matrix of transmitted signals, W is the $T \times N$ matrix of additive noise components and ρ is the expected signal to noise ratio at the receivers.

For a deterministic rank one Rician channel H is defined as

$$H = \sqrt{1-r}G + \sqrt{rNM}H_m \tag{6.2}$$

where G is a matrix of independent $\mathcal{CN}(0,1)$ random variables, H_m is an $M \times N$ deterministic matrix of rank one such that $\text{tr}\{H_m^\dagger H_m\} = 1$ and r is a non-random constant lying between 0 and 1. Without loss of generality we can assume that

$H_m = \alpha\beta^\dagger$ where α is a length M vector and β is a length N vector such that

$$H_m = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} [1\ 0 \ldots\ 0] \tag{6.3}$$

where the column and row vectors are of appropriate lengths.

In this case, the conditional probability density function of X given S is given by,

$$p(X|S) = \frac{e^{-\mathrm{tr}\{[I_T+(1-r)(\rho/M)SS^\dagger]^{-1}(X-\sqrt{rNM}SH_m)(X-\sqrt{rNM}SH_m)^\dagger\}}}{\pi^{TN}\det^N[I_T+(1-r)(\rho/M)SS^\dagger]}.$$

The conditional probability density enjoys the following properties

1. For any $T \times T$ unitary matrix ϕ

$$p(\phi X|\phi S) = p(X|S)$$

2. For any $(M-1) \times (M-1)$ unitary matrix ψ

$$p(X|S\Psi) = p(X|S)$$

where

$$\Psi = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0}^\dagger & \psi \end{bmatrix}. \tag{6.4}$$

6.2.1 Coherent Capacity

The mutual information (MI) expression for the case where H is known by the receiver has already been derived in [23]. The informed receiver capacity achieving signal S is zero mean Gaussian independent from time instant to time instant. For such a signal the MI is

$$I(X;S|H) = T \cdot E \log\det\left[I_N + \frac{\rho}{M}H^\dagger \Lambda H\right]$$

where $\Lambda = E[S_t^\tau S_t^*]$ for $t = 1, \ldots, T$, S_t is the t^{th} row of the $T \times M$ matrix S. S_t^τ denotes the transpose of S_t and $S_t^* \stackrel{\text{def}}{=} (S_t^\tau)^\dagger$.

Theorem 6.1. *Let the channel H be Rician (6.2) and be known to the receiver. Then the capacity is*

$$C_H = \max_{l,d} TE \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda^{(l,d)} H] \tag{6.5}$$

where the signal covariance matrix $\Lambda^{(l,d)}$ is of the form

$$\Lambda^{(l,d)} = \begin{bmatrix} M-(M-1)d & l\underline{1}_{M-1} \\ l\underline{1}_{M-1}^\tau & dI_{M-1} \end{bmatrix}$$

where d is a positive real number such that $0 \le d \le M/(M-1)$ and l is such that $|l| \le \sqrt{(\frac{M}{M-1} - d)d}$. I_{M-1} is the identity matrix of dimension $M-1$ and $\underline{1}_{M-1}$ is the all ones column vector of length $M-1$.

Proof: This proof is a modification of the proof in [73]. Using the property that $\Psi^\dagger H$ has the same distribution as H where Ψ is of the form given in (6.4) we conclude that

$$T \cdot E \log\det\left[I_N + \frac{\rho}{M} H^\dagger \Lambda H\right] = T \cdot E \log\det\left[I_N + \frac{\rho}{M} H^\dagger \Psi\Lambda\Psi^\dagger H\right].$$

If Λ is written as

$$\Lambda = \begin{bmatrix} c & A \\ A^\dagger & B \end{bmatrix}$$

where c is a positive number such that $c \ge A^\dagger B^{-1} A$ (to ensure positive semi-definiteness of the covariance matrix Λ), A is a row vector of length $M-1$ and B is a positive definite matrix of size $(M-1) \times (M-1)$. Then

$$\Psi\Lambda\Psi^\dagger = \begin{bmatrix} c & A\psi^\dagger \\ \psi A^\dagger & \psi B \psi^\dagger \end{bmatrix}.$$

Since $B = UDU^\dagger$ where D is a diagonal matrix and U is a unitary matrix of size $(M-1)\times(M-1)$, choosing $\psi = \Pi U$ where Π is a $(M-1)\times(M-1)$ permutation matrix, we obtain that

$$T \cdot E \log\det\left[I_N + \frac{\rho}{M}H^\dagger \Lambda H\right] = T \cdot E \log\det\left[I_N + \frac{\rho}{M}H^\dagger \Lambda^\Pi H\right]$$

where

$$\Lambda^\Pi = \begin{bmatrix} c & AU^\dagger\Pi^\dagger \\ \Pi U A^\dagger & \Pi D \Pi^\dagger \end{bmatrix}.$$

Since log det is a concave (convex cap) function we have

$$\begin{aligned} T \cdot E \log\det\left[I_N + \frac{\rho}{M}H^\dagger\ \overline{\Lambda^\Pi}\ H\right] &\geq T \cdot \frac{1}{(M-1)!}\sum_\Pi E\log\det\left[I_N + \frac{\rho}{M}H^\dagger\Lambda^\Pi H\right] \\ &= I(X;S) \end{aligned}$$

where $\overline{\Lambda^\Pi} = \frac{1}{(M-1)!}\sum_\Pi \Lambda^\Pi$ and the summation is over all $(M-1)!$ possible permutation matrices Π. Therefore, the capacity achieving Λ is given by $\overline{\Lambda^\Pi}$ and is of the form

$$\Lambda = \begin{bmatrix} c & b\underline{1}_{M-1} \\ b\underline{1}^\tau_{M-1} & dI_{M-1} \end{bmatrix}$$

where $d = \text{tr}\{B\}/(M-1)$. Now, the capacity achieving signal matrix has to satisfy $\text{tr}\{\Lambda\} = M$ since MI is monotonically increasing in $\text{tr}\{\Lambda\}$. Therefore, $c = M-(M-1)d$. And since $c \geq L^\dagger D^{-1}L$ this implies $M-(M-1)d \geq \frac{(M-1)|l|^2}{d}$ and we obtain the desired signal covariance structure. ■

The problem remains to find the l and d that achieve the maximum in (6.5). This problem has an analytical solution for the special cases of: 1) $r=0$ for which $d=1$ and $l=0$ (rank M signal S); and 2) $r=1$ for which $d=l=0$ (rank 1 signal S). In general, the optimization problem (6.5) can be solved by using the method of steepest descent over the space of parameters that satisfy the average power constraint (See Appendix B.1). Results for $\rho = 100, 10, 1, 0.1$ are shown in Figure 6.1. The optimum values of l for different values of ρ turned out to be zero, i.e. the signal energy transmitted is uncorrelated over different antenna elements and over time. As can be seen from the plot the optimum value of d stays close to 1 for high SNR and close to 0 for low SNR. That is, the optimum covariance matrix is close to an identity matrix for high SNR. For low SNR, all the energy is concentrated in the direction of the specular component or in other words the optimal signaling strategy is beamforming. These observations are proven in section 6.3.1.

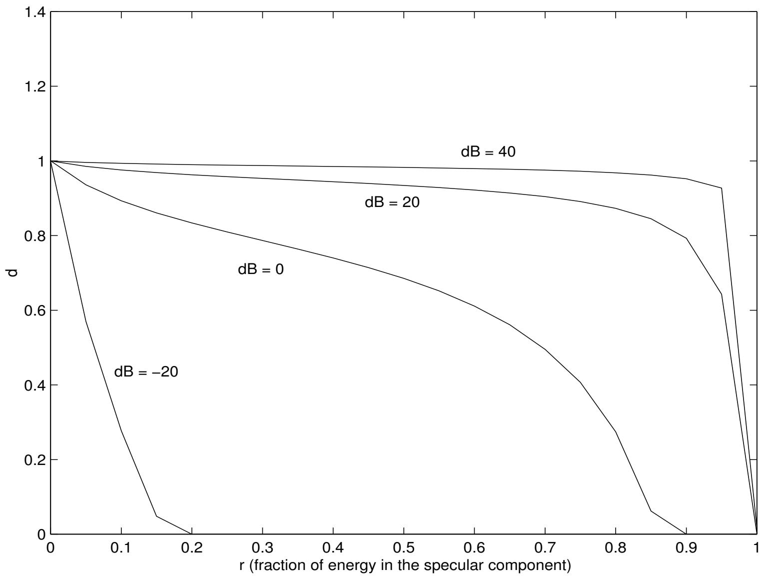


Figure 6.1: Optimum value of d as a function of r for different values of ρ

6.3 General Rank Specular Component

In this case the channel matrix can be written as

$$H = \sqrt{1-r}G + \sqrt{r}H_m \tag{6.6}$$

where G is the Rayleigh Fading component and H_m is a deterministic matrix such that $\text{tr}\{H_m H_m^\dagger\} = MN$ with no restriction on its rank. Without loss of generality, we can assume H_m to be an $M \times N$ diagonal matrix with positive real entries.

6.3.1 Coherent Capacity

For high SNR, we show that the capacity achieving signal structure basically ignores the specular component. There is no preference given to the channel directions in the specular component.

Proposition 6.1. *Let H be Rician (6.6). Let C_H be the capacity for H known at the receiver. For high SNR ρ, C_H is attained by an identity signal covariance matrix when $M \leq N$ and*

$$C_H = T \cdot E \log\det[\frac{\rho}{M} H H^\dagger] + O(\frac{\log(\sqrt{\rho})}{\sqrt{\rho}}).$$

Proof: The expression for capacity, C_H is

$$C_H = T \cdot E \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H].$$

Let H have SVD $H = \Phi\Sigma\Psi^\dagger$ then

$$\log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] = \log\det[I_N + \frac{\rho}{M} \Sigma^\dagger \Phi^\dagger \Lambda \Phi \Sigma].$$

Let $\Phi^\dagger \Lambda \Phi = D$. Then

$$\log\det[I_N + \frac{\rho}{M} \Sigma^\dagger D \Sigma] = \log\det[I_M + \frac{\rho}{M} D \Sigma\Sigma^\dagger].$$

The right hand side expression is maximized by choosing Λ such that D is diagonal [13, page 255] (We will show finally that the optimum D does not depend on the specific realization of H). Let $D = \mathrm{diag}\{d_1,\ d_2,\ \ldots, d_M\}$ and σ_i be the eigenvalues of $\Sigma\Sigma^\dagger$ and

$$E_A[f(x)] \stackrel{\text{def}}{=} E[f(x)\chi_A(x)] \tag{6.7}$$

where $\chi_A(x)$ is the characteristic function for the set A so that $\chi_A(x) = 0$ if $x \notin A$ and $\chi_A(x) = 1$ otherwise, then for large ρ

$$\begin{aligned} E \log\det[I_M + \frac{\rho}{M} D\Sigma\Sigma^\dagger] &= \sum_{i=1}^{M} E_{\sigma_i < 1/\sqrt{\rho}} \log[1 + \frac{\rho}{M} d_i \sigma_i] + \\ & \sum_{i=1}^{M} E_{\sigma_i \geq 1/\sqrt{\rho}} \log[1 + \frac{\rho}{M} d_i \sigma_i]. \end{aligned}$$

Let K denote the first term in the right hand side of the expression above and L denote the second term. We can show that

$$\begin{aligned} E \log\det[I_M + \frac{\rho}{M} D\Sigma\Sigma^\dagger] &= \log \frac{\rho}{M} + \sum_{i=1}^{M} \log(d_i) + \sum_{i=1}^{M} E_{\sigma_i > 1/\sqrt{\rho}} [\log(\sigma_i)] + \\ & O(\log(\sqrt{\rho})/\sqrt{\rho}) \end{aligned}$$

since

$$K \leq \log[1 + \sqrt{\rho}] \sum_{i=1}^{M} P(\sigma_i < 1/\sqrt{\rho}) = O(\log(\sqrt{\rho})/\sqrt{\rho})$$

and

$$L = \log \frac{\rho}{M} + \sum_{i=1}^{M} \log(d_i) + \sum_{i=1}^{M} E_{\sigma_i > 1/\sqrt{\rho}} [\log(\sigma_i)] + O(1/\sqrt{\rho}).$$

On account of log being a convex cap function the first term in the expression on the last line above is maximized by choosing $d_i = d$ for $i = 1, \ldots, M$ such that $M \cdot d = M$. ∎

For $M > N$, optimization using the steepest descent algorithm similar to the one described in Appendix B.1 shows that for high SNR the capacity achieving signal matrix is an identity matrix as well and the capacity is given by

$$C_H \approx T \cdot E \log\det[I_N + \frac{\rho}{M} H^\dagger H].$$

For low SNR, we next show that the Rician fading channel essentially behaves like an AWGN channel in the sense that the Rayleigh fading component has no effect on the structure of the optimum covariance structure.

Proposition 6.2. *Let H be Rician (6.6) and let the receiver have complete knowledge of the Rayleigh component G. For low SNR, C_H is attained by the same signal covariance matrix that attains capacity when $r = 1$, irrespective of the value of M and N, and*

$$C_H = T\rho[r\lambda_{max}(H_m H_m^\dagger) + (1-r)N] + O(\rho^2).$$

Proof: Let $\|H\|$ denote the matrix 2-norm of H, γ be a positive number such that $\gamma \in (0, 1)$ then

$$\begin{aligned} C_H &= T \cdot E \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] \\ &= T \cdot E_{\|H\| \geq 1/\rho^\gamma} \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] + E_{\|H\| < 1/\rho^\gamma} \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] \\ &= TE\mathrm{tr}\{\frac{\rho}{M} H^\dagger \Lambda H\} + O(\rho^{2-2\gamma}) \end{aligned}$$

where $E_{\|H\| \geq 1/\rho^\gamma}[\cdot]$ is as defined in (6.7). This follows from the fact that $P(\|H\| \geq 1/\rho^\gamma) \leq O(e^{-\frac{1}{TM\rho^\gamma}})$ and for $\|H\| < 1/\rho^\gamma$ $\log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] = \mathrm{tr}[\frac{\rho}{M} H^\dagger \Lambda H] + O(\rho^{2-2\gamma})$. Since γ is arbitrary

$$E \log\det[I_N + \frac{\rho}{M} H^\dagger \Lambda H] = E\mathrm{tr}[\frac{\rho}{M} H^\dagger \Lambda H] + O(\rho^2).$$

Now

$$\begin{aligned} E\mathrm{tr}[H^\dagger \Lambda H] &= \mathrm{tr}\{(1-r)E[G^\dagger \Lambda G] + rH_m^\dagger \Lambda H_m]\} \\ &= \mathrm{tr}\{(1-r)\Lambda E[GG^\dagger] + r\Lambda H_m H_m^\dagger\}. \end{aligned}$$

Therefore, we have to choose Λ to maximize $\mathrm{tr}\{(1-r)N\Lambda + r\Lambda H_m H_m^\dagger\}$. Since H_m is diagonal the trace depends only on the diagonal elements of Λ. Therefore, Λ can be chosen to be diagonal. Also, because of the power constraint, $\mathrm{tr}\{\Lambda\} \leq M$, to maximize the expression we choose $\mathrm{tr}\{\Lambda\} = M$. The maximizing Λ has as many non-zero elements as the multiplicity of the maximum eigenvalue of $(1-r)NI_M + rH_m H_m^\dagger$.

The non-zero elements of Λ multiply the maximum eigenvalues of $(1-r)NI_M + rH_mH_m^\dagger$ and can be chosen to be of equal magnitude summing up to M. This is the same Λ maximizing the capacity for additive white Gaussian noise channel with channel H_m. ■

Note that if we choose $\Lambda = I_M$ then varying r has no effect on the value of the capacity. This explains the trend seen in Figures 4.2 and 4.3 in Chapter 4 and Figures 5.1 and 5.2 in Chapter 5.

6.3.2 Non-Coherent Capacity Upper and Lower Bounds

It follows from the data processing theorem that the non-coherent capacity, C can never be greater than the coherent capacity C_H, that is, the uninformed capacity is never decreased when the channel is known to the receiver.

Proposition 6.3. *Let H be Rician (6.6) and the receiver have no knowledge of the Rayleigh component then*

$$C \leq C_H.$$

Now, we establish a lower bound which is similar in flavor to those derived in chapters 4 and 5.

Proposition 6.4. *Let H be Rician (6.6). A lower bound on capacity when the receiver has no knowledge of G is*

$$\begin{aligned} C &\geq C_H - NE\left[\log_2 \det\left(I_T + (1-r)\frac{\rho}{M}SS^\dagger\right)\right] && (6.8) \\ &\geq C_H - NM\log_2(1+(1-r)\frac{\rho}{M}T). && (6.9) \end{aligned}$$

Proof: Proof is similar to that of Proposition 5.2 in Chapter 5 and won't be repeated here. ■

We notice that the second term in the lower bound goes to zero when $r = 1$: as the channel becomes purely Gaussian the capacity of the channel is completely determined.

6.3.3 Non-Coherent Capacity: Expressions for Low SNR

In this section, we introduce some new notation for ease of description. if X is a $T \times N$ matrix then let $\tilde{X}$ denote the "unwrapped" $NT \times 1$ vector formed by placing the transposed rows of X in a single column in an increasing manner. That is, if $X_{i,j}$ denotes the element of X in the i^{th} row and j^{th} column then $\tilde{X}_{i,1} = X_{\lfloor i/N \rfloor, i\%N}$. The channel model $X = \sqrt{\frac{\rho}{M}} SH + W$ can now be written as $\tilde{X} = \sqrt{\frac{\rho}{M}} \hat{H}\tilde{S} + \tilde{W}$. $\hat{H}$ is given by $\hat{H} = I_T \otimes H^\tau$ where H^τ denotes the transpose of H. The notation $A \otimes B$ denotes the Kronecker product of the matrices A and B and is defined as follows. If A is a $I \times J$ matrix and B a $K \times L$ matrix then $A \otimes B$ is a $IK \times JL$ matrix

$$A \otimes B = \begin{bmatrix} (A)_{11}B & (A)_{12}B & \dots & (A)_{1J}B \\ (A)_{21}B & (A)_{22}B & \dots & (A)_{2J}B \\ \vdots & \vdots & \ddots & \vdots \\ (A)_{I1}B & (A)_{I2}B & \dots & (A)_{IJ}B \end{bmatrix}.$$

This way, we can describe the conditional probability density function $p(X|S)$ as follows

$$p(X|S) = \frac{1}{\pi^{TN}|\Lambda_{\tilde{X}|\tilde{S}}|} e^{-(\tilde{X} - \sqrt{r\frac{\rho}{M}}\hat{H}_m\tilde{S})^\dagger \Lambda_{\tilde{X}|\tilde{S}}^{-1} (\tilde{X} - \sqrt{r\frac{\rho}{M}}\hat{H}_m\tilde{S})}$$

where $|\Lambda_{\tilde{X}|\tilde{S}}| = \det(I_{TN} + (1-r)SS^\dagger \otimes I_N)$.

For low SNR, it will be shown that the channel behaves as an AWGN channel. Calculation of capacity for the special case of peak power constraint has been shown in Appendix B.2.

Theorem 6.2. *Let the channel H be Rician (6.6) and the receiver have no knowledge* *of G. For fixed M, N and T if S is a Gaussian distributed source then as $\rho \to 0$*

$$I(X;S_G) = rT\rho\lambda_{max}(H_m H_m^\dagger) + O(\rho^2)$$

where $I(X;S_G)$ is the mutual information between the output and the Gaussian source.

Proof: First, $I(X;S) = \mathcal{H}(X) - \mathcal{H}(X|S)$. Since S is Gaussian distributed, $E[\log\det(I_N + \frac{\rho}{M}\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger)] \leq \mathcal{H}(X) \leq \log\det(I_N + \frac{\rho}{M}\Lambda_{\tilde{X}})$ where the expectation is taken over the distribution of H and $\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger = \Lambda_{\tilde{X}|H}$ is the covariance of $\tilde{X}$ for a particular H. Next, we show that $\mathcal{H}(X) = \frac{\rho}{M}\mathrm{tr}\{\Lambda_{\tilde{X}}\} + O(\rho^2)$. First, the upper bound to $\mathcal{H}(X)$ can be written as $\frac{\rho}{M}\mathrm{tr}\{\Lambda_{\tilde{X}}\} + O(\rho^2)$ because H is Gaussian distributed and the probability that $\|H\| > R$ is of the order e^{-R^2}. Second, $E[\log\det(I_{TN} + \frac{\rho}{M}\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger)] = E_{\|H\|<(\frac{M}{\rho})^\gamma}[\cdot] + E_{\|H\|\geq(\frac{M}{\rho})^\gamma}[\cdot]$ where γ is a number such that $2-\gamma > 1$ or $\gamma < 1$. Then

$$\begin{aligned} E[\log\det(I_{TN} + \frac{\rho}{M}\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger)] &= \frac{\rho}{M}E_{\|H\|<(\frac{M}{\rho})^\gamma}[\mathrm{tr}\{\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger\}] + \\ &\quad O(\rho^{2-\gamma}) + O(\log((\frac{M}{\rho})^\gamma)e^{-(\frac{M}{\rho})^\gamma}) \\ &= \frac{\rho}{M}E[\mathrm{tr}\{\hat{H}\Lambda\hat{H}^\dagger\}] + O(\rho^{2-\gamma}). \end{aligned}$$

Since γ is arbitrary, we have $\mathcal{H}(X) = \frac{\rho}{M}E[\mathrm{tr}\{\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger\}] + O(\rho^2)$. Note that $\Lambda_{\tilde{X}} = E[\Lambda_{\tilde{X}|H}]$ and since $\mathcal{H}(X)$ is sandwiched between two expressions of the form $\frac{\rho}{M}\mathrm{tr}\{\Lambda_{\tilde{X}}\} + O(\rho^2)$ the assertion follows.

Calculate $\mathcal{H}(X|S) = E[\log\det(I_{TN} + (1-r)\frac{\rho}{M}SS^\dagger \otimes I_N)]$. We have S to be Gaussian distributed therefore in a similar manner it can be shown that $\mathcal{H}(X|S) = (1-r)\frac{\rho}{M}\mathrm{tr}\{E[SS^\dagger \otimes I_N]\} + O(\rho^2)$.

$H = \sqrt{r}H_m + \sqrt{1-r}G$ therefore, $\Lambda_{\tilde{X}} = E[\hat{H}\Lambda_{\tilde{S}}\hat{H}^\dagger] = r\hat{H}_m\Lambda_{\tilde{S}}\hat{H}_m^\dagger + (1-r)E[SS^\dagger] \otimes I_N$. Therefore, we have for a Gaussian distributed input $I(X;S_G) = r\frac{\rho}{M}\mathrm{tr}\{\hat{H}_m\Lambda_{\tilde{S}}\hat{H}_m^\dagger\} +$

$O(\rho^2)$. Since H_m is a diagonal matrix only the diagonal elements of $\Lambda_{\tilde{S}}$ matter. Therefore, we can choose the signals to be independent from time instant to time instant. Also, to maximize $\text{tr}\{\hat{H}_m \Lambda_{\tilde{S}} \hat{H}_m^\dagger\}$ under the condition $\text{tr}\{\Lambda\} \leq TM$ it is best to concentrate all the available energy on the largest eigenvalues of H_m. Therefore, we obtain

$$I(X; S_G) = r\frac{\rho}{M} TM\lambda_{max}(H_m H_m^\dagger) + O(\rho^2).$$

■

Corollary 6.1. *For purely Rayleigh fading channels when the receiver has no knowledge of G a Gaussian transmitted signal satisfies* $\lim_{\rho\to 0} I(X; S_G)/\rho = 0$.

The peak constraint results in Appendix B.2 and the Gaussian input results imply that for low SNR Rayleigh fading channels are at a disadvantage compared to Rician fading channels. But, it has been shown in [2, 75] for single antenna transmit and receive channel Rayleigh fading provides as much capacity as a Gaussian channel with the same energy for low SNR. We will extend that result to multiple transmit and receive antenna channel for the general case of Rician fading. The result for Rayleigh fading will follow as a special case.

Theorem 6.3. *Let H be Rician (6.6) and the receiver have no knowledge of G. For fixed M, N and T*

$$\lim_{\rho\to 0} \frac{C}{\rho} = T\left[r\lambda_{max}(H_m H_m^\dagger) + N(1-r)\right].$$

Proof: First, absorb $\sqrt{\frac{\rho}{M}}$ into $\tilde{S}$ and rewrite the channel as

$$\tilde{X} = \hat{H}\tilde{S} + W$$

with the average power constraint on the signal $\tilde{S}$ $E[\text{tr}\{\tilde{S}\tilde{S}^\dagger\}] \leq \frac{\rho}{M} TM = \rho T$.

It has been shown [75] that if the input alphabet includes the value "0" (symbol with 0 power) for a channel with output X, and condition probability given by $p(X|S)$, then

$$\lim_{P_C \to 0} \frac{C}{P_C} = \sup_{s \in \mathcal{S}} \frac{D(\ p(X|S=s) \parallel p(X|S=0)\ )}{P_s}$$

where $\mathcal{S}$ is the set of values that the input can take, P_C is the average power constraint on the input (in our case, $E[\text{tr}\{SS^\dagger\}] \leq P_C = \rho T$) and $P_s = \text{tr}\{ss^\dagger\}$ is the energy in the specific realization of the input $S = s$ and $D(p_A \| p_B)$ is the Kullback-Leibler distance for continuous density functions with argument x defined as

$$D(p_A \| p_B) = \int p_A(x) \log \frac{p_A(x)}{p_B(x)} dx.$$

Applying the above result to the case of Rician fading channels, we obtain

$$\lim_{\rho \to 0} \frac{C}{\rho T} = \sup_{\tilde{S}} \frac{D(\ p(\tilde{X}|\tilde{S}) \parallel p(\tilde{X}|0)\ )}{\text{tr}\{\tilde{S}\tilde{S}^\dagger\}}.$$

First, we have

$$p(\tilde{X}|\tilde{S}) = \frac{1}{\pi^{TN}|\Lambda_{\tilde{X}|\tilde{S}}|} e^{-(\tilde{X}-\sqrt{r}\hat{H}_m\tilde{S})^\dagger\ \Lambda^{-1}_{\tilde{X}|\tilde{S}}\ (\tilde{X}-\sqrt{r}\hat{H}_m\tilde{S})}$$

and

$$p(\tilde{X}|0) = \frac{1}{\pi^{TN}} e^{-\tilde{X}^\dagger\tilde{X}}.$$

Therefore,

$$\begin{aligned}
D(\ p(\tilde{X}|\tilde{S}) \parallel p(\tilde{X}|0)\ ) &= \int p(\tilde{X}|\tilde{S}) \Bigg[\log \frac{1}{|\Lambda_{\tilde{X}|\tilde{S}}|} \ + \ \tilde{X}^\dagger\tilde{X} \ - \\
&\quad \left(\tilde{X} - \sqrt{r}\hat{H}_m\tilde{S}\right)^\dagger \ \Lambda^{-1}_{\tilde{X}|\tilde{S}} \ \left(\tilde{X} - \sqrt{r}\hat{H}_m\tilde{S}\right)\Bigg] d\tilde{X} \\
&= \log \frac{1}{|\Lambda_{\tilde{X}|\tilde{S}}|} \ + \ \text{tr}\left\{r\hat{H}_m\tilde{S}\tilde{S}^\dagger\hat{H}_m^\dagger + \Lambda_{\tilde{X}|\tilde{S}}\right\} - TN \\
&= \log \frac{1}{\det(I_{TN} + (1-r)SS^\dagger \otimes I_N)} \ + \\
&\quad \text{tr}\left\{r\hat{H}_m\tilde{S}\tilde{S}^\dagger\hat{H}_m^\dagger + (1-r)SS^\dagger \otimes I_N\right\}.
\end{aligned}$$

This gives,

$$\frac{D(\ p(\tilde{X}|\tilde{S}) \parallel p(\tilde{X}|0)\ )}{\mathrm{tr}\{\tilde{S}\tilde{S}^\dagger\}} = -N\frac{\sum_{i=1}^{T}\log(1+\lambda_i(SS^\dagger))}{\sum_{i=1}^{T}\lambda_i(SS^\dagger)} + \frac{\mathrm{tr}\{r\hat{H}_m\tilde{S}\tilde{S}^\dagger\hat{H}_m^\dagger\}}{\sum_{i=1}^{T}\mathrm{tr}\{S_iS_i^\dagger\}} + N(1-r)$$

where we have used the facts that $\det(I_{TN}+(1-r)SS^\dagger\otimes I_N) = \det(I_T+(1-r)SS^\dagger)^N$, $\mathrm{tr}\{\tilde{S}\tilde{S}^\dagger\} = \mathrm{tr}\{SS^\dagger\} = \sum_{i=1}^{T}\mathrm{tr}\{S_i^\tau S_i^*\}$ where S_i is the i^{th} row in the matrix S.

Since,

$$\hat{H}_m\tilde{S}\tilde{S}^\dagger\hat{H}_m^\dagger = \begin{bmatrix} H_m^\tau S_1^\tau S_1^* H_m^* & H_m^\tau S_1^\tau S_2^* H_m^* & \dots & H_m^\tau S_1^\tau S_T^* H_m^* \\ H_m^\tau S_2^\tau S_1^* H_m^* & H_m^\tau S_2^\tau S_2^* H_m^* & \dots & H_m^\tau S_2^\tau S_T^* H_m^* \\ \vdots & \vdots & \ddots & \vdots \\ H_m^\tau S_T^\tau S_1^* H_m^* & H_m^\tau S_T^\tau S_2^* H_m^* & \dots & H_m^\tau S_T^\tau S_T^* H_m^* \end{bmatrix}$$

we have $\mathrm{tr}\{\hat{H}_m\tilde{S}\tilde{S}^\dagger\hat{H}_m^\dagger\} = \sum_{i=1}^{T}\mathrm{tr}\{H_m^\tau S_i^\tau S_i^* H_m^*\} = \mathrm{tr}\{H_m^* H_m^\tau \sum_{i=1}^{T} S_i^\tau S_i^*\}$. Therefore,

$$\frac{D(\ p(\tilde{X}|\tilde{S}) \parallel p(\tilde{X}|0)\ )}{\mathrm{tr}\{\tilde{S}\tilde{S}^\dagger\}} = -N\frac{\sum_{i=1}^{N}\log(1+\lambda_i(S^\dagger S))}{\sum_{i=1}^{N}\lambda_i(S^\dagger S)} + r\frac{\mathrm{tr}\{H_m^* H_m^\tau \sum_{i=1}^{T} S_i^\tau S_i^*\}}{\sum_{i=1}^{T}\mathrm{tr}\{S_i^\tau S_i^*\}} + N(1-r).$$

Note that since H_m is a diagonal matrix only the diagonal elements of $S_iS_i^\dagger$ affect the second term. Therefore, for a given $\sum_{i=1}^{T} S_i^\tau S_i^*$ the second term in right hand side of the expression above can be maximized by choosing S_i such that $S_i^\tau S_i^*$ is diagonal. In addition the non-zero values of $S_iS_i^\dagger$ should be located at the same diagonal positions as the maximum entries of $H_m^* H_m^\tau$. In such a case the expression above evaluates to

$$\frac{D(\ p(\tilde{X}|\tilde{S}) \parallel p(\tilde{X}|0)\ )}{\mathrm{tr}\{\tilde{S}\tilde{S}^\dagger\}} = -N\frac{\log(1+\mathrm{tr}\{S^\dagger S\})}{\mathrm{tr}\{S^\dagger S\}} + r\lambda_{max}(H_m^* H_m^\tau) + N(1-r).$$

The first term can be made arbitrarily small by letting $\mathrm{tr}\{S^\dagger S\} \to \infty$. Therefore, we have $\lim_{\rho\to 0}\frac{C}{\rho T} = r\lambda_{max}(H_m H_m^\dagger) + N(1-r)$. ■

Theorem 6.3 suggests that at low SNR all the energy has to be concentrated in the strongest directions of the specular component. In [2] it is shown that the optimum signaling scheme for Rayleigh fading channels is an "on-off" signaling scheme. We conjecture that the capacity achieving signaling scheme for low SNR in the case of the Rician fading is also a similar "on-off" signaling scheme.

6.3.4 Non-Coherent Capacity: Expressions for High SNR

In this section we apply the method developed in [85] for the analysis of Rayleigh fading channels. The only difference in the models considered in [85] and here is that we assume H has a deterministic non-zero mean. For convenience, we use a different notation for the channel model. We rewrite the channel model as

$$X = SH + W$$

with $H = \sqrt{r}H_m + \sqrt{1-r}G$ where H_m is the specular component of H and G denotes the Rayleigh component. G and W consist of Gaussian circular independent random variables and the covariance matrices of G and W are given by $(1-r)I_{MN}$ and $\sigma^2 I_{TN}$, respectively. H_m is deterministic such that $E[\mathrm{tr}\{HH^\dagger\}] = MN$ and r is a number between 0 and 1.

Lemma 6.1. *Let the channel be Rician (6.6) and the receiver have no knowledge of G. Then the capacity achieving signal, S can be written as $S = \Phi V \Psi^\dagger$ where Φ is a $T \times M$ unitary matrix independent of V and Ψ. V and Ψ are $M \times M$.*

Proof: Follows from the fact that $p(\Phi X|\Phi S) = p(X|S)$. ∎

In [85] the requirement for $X = SH + W$ was that X had to satisfy the property that in the singular value decomposition of X, $X = \Phi V \Psi^\dagger$ Φ be independent of V and Ψ. This property holds for the case of Rician fading too because the density functions of X, SH and S are invariant to pre-multiplication by a unitary matrix.

Therefore, the leading unitary matrix in the SVD decomposition of any of X, SH and S is independent of the other two components in the SVD and isotropically distributed.

Therefore, Lemma 6 in [85] holds

Lemma 6.2. *Let $R = \Phi_R \Sigma_R \Psi_R^\dagger$ be such that Φ_R is independent of Σ_R and Ψ_R. Then*

$$\mathcal{H}(R) = \mathcal{H}(Q\Sigma_R\Psi_R^\dagger) + \log|G(T,M)| + (T-M)E[\log\det\Sigma_R^2]$$

where Q is an $M \times M$ unitary matrix independent of V and Ψ and $|G(T,M)|$ is the volume of the Grassmann Manifold and is equal to

$$\frac{\prod_{i=T-M+1}^{T} \frac{2\pi^i}{(i-1)!}}{\prod_{i=1}^{M} \frac{2\pi^i}{(i-1)!}}.$$

The Grassmann Manifold $G(T,M)$ [85] is the set of equivalence classes of all $T \times M$ unitary matrices such that if P, Q belong to an equivalence class then $P = QU$ for some $M \times M$ unitary matrix U.

$M = N$, $T \geq 2M$

To calculate $I(X;S)$ we need to compute $\mathcal{H}(X)$ and $\mathcal{H}(X|S)$. To compute $\mathcal{H}(X|S)$ we note that given S, X is a Gaussian random vector with columns of X independent of each other. Each row has the common covariance matrix given by $(1-r)SS^\dagger + \sigma^2 I_T = \Phi V^2 \Phi^\dagger + \sigma^2 I_T$. Therefore

$$\mathcal{H}(X|S) = ME[\sum_{i=1}^{M} \log(\pi e((1-r)\|s_i\|^2 + \sigma^2)] + M(T-M)\log(\pi e \sigma^2).$$

To compute $\mathcal{H}(X)$, we write the SVD: $X = \Phi_X \Sigma_X \Psi_X^\dagger$. Note that Φ_X is isotropically distributed and independent of $\Sigma_X \Psi_X^\dagger$, therefore from Lemma 6.2 we have

$$\mathcal{H}(X) = \mathcal{H}(Q\Sigma_X\Psi_X^\dagger) + \log|G(T,M)| + (T-M)E[\log\det\Sigma_X^2].$$

We first characterize the optimal input distribution in the following lemma.

Lemma 6.3. *Let H be Rician (6.6) and the receiver have no knowledge of G. Let $(s_i^\sigma, i = 1, \ldots, M)$ be the optimal input signal of each antenna at when the noise power at the receive antennas is given by σ^2. If $T \geq 2M$,*

$$\frac{\sigma}{\|s_i^\sigma\|} \xrightarrow{P} 0, \text{ for } i = 1, \ldots, M \tag{6.10}$$

where $\xrightarrow{P}$ denotes convergence in probability.

Proof: See Appendix B.3. ■

Lemma 6.4. *Let H be Rician (6.6) and the receiver have no knowledge of G. The maximal rate of increase of capacity, $\max_{p(S):E[\mathrm{tr}\{SS^\dagger\}]\leq TM} I(X;S)$ with SNR is $M(T-M)\log\rho$ and the constant norm source $\|s_i\|^2 = T$ for $i = 1, \ldots, M$ attains this rate.*

Proof: See Appendix B.3. ■

Lemma 6.5. *Let H be Rician (6.6) and the receiver have no knowledge of G. As $T \to \infty$ the optimal source in Lemma 6.4 is the constant norm input*

Proof: See Appendix B.3. ■

From now on, we assume that the optimal input signal is the constant norm input. For the constant norm input $\Phi V \Psi^\dagger = \Phi V$ since Φ is isotropically distributed.

Theorem 6.4. *Let the channel be Rician (6.6) and the receiver have no knowledge of G. For the constant norm input, as $\sigma^2 \to 0$ the capacity is given by*

$$\begin{aligned} C &= \log|G(T,M)| + (T-M)E[\log\det H^\dagger H] - M(T-M)\log \pi e \sigma^2 - \\ &\quad M^2 \log \pi e + \mathcal{H}(QVH) + (T-2M)M\log T - M^2\log(1-r) \end{aligned}$$

where Q, V and $|G(T,M)|$ are as defined in Lemma 6.2.

Proof: Since $\|s_i^2\| \gg \sigma^2$ for all $i = 1, \ldots, M$

$$\begin{aligned}
\mathcal{H}(X|S) &= ME[\sum_{i=1}^{M} \log \pi e((1-r)\|s_i\|^2 + \sigma^2)] + M(T-M)\log(\pi e \sigma^2) \\
&\approx ME[\sum_{i=1}^{M} \log \pi e(1-r)\|s_i\|^2] + M(T-M)\log \pi e \sigma^2 \\
&= ME[\log\det(1-r)V^2] + M^2 \log \pi e + M(T-M)\log \pi e \sigma^2
\end{aligned}$$

and from Appendix B.5

$$\begin{aligned}
\mathcal{H}(X) &\approx \mathcal{H}(SH) \\
&= \mathcal{H}(QVH) + \log|G(T,M)| + (T-M)E[\log\det(H^\dagger V^2 H)] \\
&= \mathcal{H}(QVH) + \log|G(T,M)| + (T-M)E[\log\det V^2] + \\
&\quad (T-M)E[\log\det HH^\dagger].
\end{aligned}$$

Combining the two equations

$$\begin{aligned}
I(X;S) \approx\ & \log|G(T,M)| + (T-M)E[\log\det H^\dagger H] - M(T-M)\log \pi e \sigma^2 + \\
& \mathcal{H}(QVH) - M^2 \log \pi e + (T-2M)E[\log\det V^2] - M^2 \log(1-r).
\end{aligned}$$

Now, since the optimal input signal is $\|s_i\|^2 = T$ for $i = 1, \ldots, M$, we have

$$\begin{aligned}
C &= I(X;S) \\
&\approx \log|G(T,M)| + (T-M)E[\log\det H^\dagger H] - M(T-M)\log \pi e \sigma^2 - \\
&\quad M^2 \log \pi e + \mathcal{H}(QVH) + (T-2M)M\log T - M^2 \log(1-r).
\end{aligned}$$

■

Theorem 6.5. *Let H be Rician (6.6) and the receiver have no knowledge of G. As $T \to \infty$ the normalized capacity $C/T \to E[\log\det \frac{\rho}{M} H^\dagger H]$ where $\rho = M/\sigma^2$.*

Proof: First, a lower bound to capacity as $\sigma^2 \to 0$ is given by

$$\begin{aligned} C &\geq \log|G(T,M)| + (T-M)E[\log\det H^\dagger H] + M(T-M)\log\frac{T\rho}{M\pi e} - \\ &\quad M^2\log T - M^2\log(1-r) - M^2\log \pi e. \end{aligned}$$

In [85] it's already been shown that $\lim_{T\to\infty}(\frac{1}{T}\log|G(T,M)| + M(1-\frac{M}{T})\log\frac{T}{\pi e}) = 0$. Therefore we have as $T \to \infty$

$$C/T \geq ME[\log\det\frac{\rho}{M}H^\dagger H].$$

Second, since $\mathcal{H}(QVH) \leq M^2\log(\pi e T)$ an asymptotic upper bound on capacity is given by

$$\begin{aligned} C &\leq \log|G(T,M)| + (T-M)E[\log\det H^\dagger H] + M(T-M)\log\frac{T\rho}{M\pi e} - \\ &\quad M^2\log(1-r). \end{aligned}$$

Therefore, we have as $T \to \infty$

$$C/T \leq E[\log\det\frac{\rho}{M}H^\dagger H].$$

■

$M < N$ $T \geq M + N$

In this case we show that the optimal rate of increase is given by $M(T-M)\log\rho$. The higher number of receive antennas can provide only a finite increase in capacity for all SNRs.

Theorem 6.6. *Let the channel be Rician (6.6) and the receiver have no knowledge of G. Then the maximum rate of increase of capacity with respect to $\log\rho$ is given by $M(T-M)$.*

Proof: See Appendix B.3. ■

6.4 Training in Non-Coherent Communications

It is important to know whether training based signal schemes are practical and if they are how much time can be spent in learning the channel and what the optimal training signal is like. Hassibi and Hochwald [39] have addressed these issues for the case of Rayleigh fading channels. They showed that 1) pilot symbol training based communication schemes are highly suboptimal for low SNR and 2) when practical the optimal amount of time devoted to training is equal to the number of transmitters, M when the fraction of power devoted to training is allowed to vary and 3) the orthonormal signal is the optimal signal for training.

In [85] the authors demonstrate a very simple training method that achieves the optimal rate of increase with SNR. The same training method can also be easily applied to the Rician fading model with deterministic specular component. The training signal is the $M \times M$ diagonal matrix dI_M. d is chosen such that the same power is used in the training and the communication phase. Therefore, $d = \sqrt{M}$. Using $S = dI_M$, the output of the MIMO channel in the training phase is given by

$$X = \sqrt{M}\sqrt{r}H_m + \sqrt{M}\sqrt{1-r}G + W$$

The Rayleigh channel coefficients G can be estimated independently using scalar minimum mean squared error (MMSE) estimates since the elements of W and G are i.i.d. Gaussian random variables

$$\hat{G} = \frac{\sqrt{1-r}\sqrt{M}}{(1-r)M + \sigma^2}[X - \sqrt{M}\sqrt{r}H_m]$$

where recall that σ^2 is the variance of the components of W. The elements of the estimate $\hat{G}$ are i.i.d. Gaussian with variance $\frac{(1-r)M}{(1-r)M+\sigma^2}$. Similarly, the estimation error matrix $G - \hat{G}$ has i.i.d Gaussian distributed elements with zero mean and variance $\frac{\sigma^2}{(1-r)M+\sigma^2}$.

The output of the channel in the communication phase is given by

$$\begin{aligned} X &= SH + W \\ &= \sqrt{r}SH_m + \sqrt{1-r}S\hat{G} + \sqrt{1-r}S(G-\hat{G}) + W \end{aligned}$$

where S consists of zero mean i.i.d circular Gaussian random variables with zero mean and unit variance. This choice of S is sub-optimal as this might not be the capacity achieving signal, but this choice gives us a lower bound on capacity. Let $\hat{W} = \sqrt{1-r}S(G-\hat{G}) + W$. For the choice of S given above the entries of $\hat{W}$ are uncorrelated with each other and also with $S(\sqrt{r}H_m + \sqrt{1-r}\hat{G})$. The variance of each of the entries of $\hat{W}$ is given by $\sigma^2 + (1-r)M\frac{\sigma^2}{(1-r)M+\sigma^2}$. If $\hat{W}$ is replaced with a white Gaussian noise with the same covariance matrix then the resulting mutual information is a lower bound on the actual mutual information [13, p. 263]. This result is formally stated in Proposition 6.5. In this section we deal with normalized capacity C/T instead of capacity C. The lower bound on the normalized capacity is given by

$$C/T \geq \frac{T-T_t}{T} E \log\det\left(I_M + \frac{\rho_{eff}}{M} H_1 H_1^\dagger\right)$$

where ρ_{eff} in the expression above is the effective SNR at the output (explained at the end of this paragraph), and H_1 is a Rician channel with a new Rician parameter r_{new} where $r_{new} = \frac{r}{r+(1-r)\frac{(1-r)M}{(1-r)M+\sigma^2}}$. This lower bound can be easily calculated because the lower bound is essentially the coherent capacity with H replaced by $\sqrt{r_{new}}H_m + \sqrt{1-r_{new}}\hat{G}$. The signal covariance structure was chosen to be an identity matrix as this is the optimum covariance matrix for high SNR. The effective SNR is now given by the ratio of the energy of the elements of $S(\sqrt{r}H_m + \sqrt{1-r}\hat{G})$ to the energy of the elements of $\hat{W}$. The energy in the elements of $S(\sqrt{r}H_m + \sqrt{1-r}\hat{G})$ is given by $M(r + (1-r)^2\frac{M}{(1-r)M+\sigma^2})$ and the energy in the elements of $\hat{W}$ are given by $\sigma^2 + \frac{(1-r)M\sigma^2}{(1-r)M+\sigma^2}$. Therefore, the effective SNR, ρ_{eff} is given by $\frac{\rho[r+r(1-r)\rho+(1-r)^2\rho]}{[1+2(1-r)\rho]}$ where $\rho = \frac{M}{\sigma^2}$ is the actual SNR. Note, for $r = 1$ no training is required since the channel is completely known.

This simple scheme achieves the optimum increase of capacity with SNR and uses only M of the T symbols for training.

We will compare the LMS algorithm for estimation of the channel coefficients to this simple scheme. LMS is obviously at a disadvantage because when $\sigma^2 = 0$ the simple scheme outlined above generates a perfect channel estimate after M training symbols whereas the LMS algorithm requires much more than M training symbols to obtain an accurate estimate. The performance of the simple training scheme is plotted with respect to different SNR values for comparison with the asymptotic upper bound to capacity in the proof of Theorem 6.5. The plot also verifies the result of Theorem 6.5. The plots are for $M = N = 5$, $r = 0.9$ and $T = 50$ in Figure 6.2 the specular component is a rank-one specular component given by (6.3).

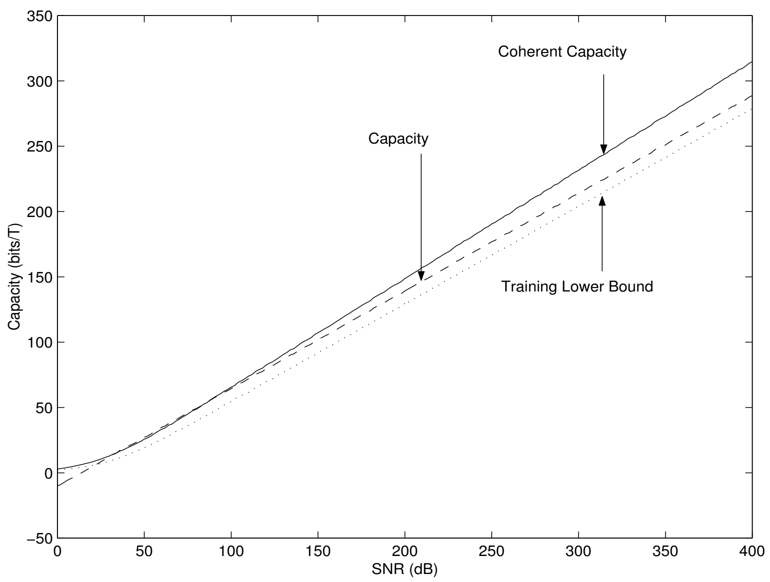


Figure 6.2: Asymptotic capacity upper bound, Capacity Upper and Lower bounds for different values of SNR

We can quantify the amount of training required using the techniques in [39]. In [39], the authors use the optimization of the lower bound on capacity to find the optimal allocation of training as compared to communication. Let T_t denote the amount of time devoted to training and T_c the amount of time devoted to actual communication. Let S_t be the $T_t \times M$ signal used for training and S_c the $T_c \times M$ signal used for communication.

Let κ denote the fraction of the energy used for communication. Then $T = T_t + T_c$ and $\mathrm{tr}\{S_t S_t^\dagger\} = (1-\kappa)TM$ and $\mathrm{tr}\{S_c S_c^\dagger\} = \kappa TM$.

$$\begin{aligned} X_t &= S_t(\sqrt{r}H_m + \sqrt{1-r}G) + W_t \\ X_c &= S_c(\sqrt{r}H_m + \sqrt{1-r}G) + W_c \end{aligned}$$

where X_t is $T_t \times N$ and X_c is $T_c \times N$. G is estimated from the training phase. For that we need $T_t \geq M$. Since G and W_t are Gaussian the MMSE estimate of G is also the linear MMSE estimate conditioned on S. The optimal estimate is given by

$$\hat{G} = \sqrt{1-r}(\sigma^2 I_M + (1-r)S_t^\dagger S_t)^{-1} S_t^\dagger (X_t - \sqrt{r} S_t H_m).$$

Let $\bar{G} = G - \hat{G}$ then

$$X_c = S_c(\sqrt{r}H_m + \sqrt{1-r}\hat{G}) + \sqrt{1-r}S_c\bar{G} + W_c.$$

Let $\hat{W}_c = \sqrt{1-r}S_t\bar{G} + W$. Note that elements of $\hat{W}_c$ are uncorrelated with each other and have the same marginal densities when the elements of S_c are chosen to be i.i.d Gaussian. If we replace $\hat{W}_c$ with Gaussian noise that is zero-mean and spatially and temporally independent the elements of which have the same variance as the elements of $\hat{W}_c$ then the resulting mutual information is a lower bound to the actual mutual information in the above channel. This is stated formally in the following proposition.

Proposition 6.5 (Theorem 1 in [39]). *Let*

$$X = SH + W$$

be a Rician fading channel with H known to the receiver. Let S and W satisfy $\frac{1}{M}E[SS^\dagger] = 1$ and $\frac{1}{M}E[WW^\dagger] = \sigma^2$ and be uncorrelated with each other. Then the worst case noise has i.i.d. zero mean Gaussian distribution, i.e. $W \sim \mathcal{CN}(0, I_N)$. Moreover, this distribution has the following minimax property

$$I_{W\sim\mathcal{CN}(0,\sigma^2 I_N),S}(X;S) \leq I_{W\sim\mathcal{CN}(0,\sigma^2 I_N),S\sim\mathcal{CN}(0,I_M)}(X;S) \leq I_{W,S\sim\mathcal{CN}(0,I_M)}(X;S)$$

where $I_{W\sim\mathcal{CN}(0,\sigma^2 I_N),S}(X;S)$ denotes the mutual information between X and S when W has a zero mean complex circular Gaussian distribution and S has any arbitrary distribution.

The variance of the elements of $\hat{W}_c$ is given by

$$\begin{aligned}
\sigma^2_{w_c} &= \sigma^2 + \frac{1-r}{NT_c}\mathrm{tr}\{E[\bar{G}\bar{G}^\dagger]\kappa T I_M\} \\
&= \sigma^2 + \frac{(1-r)\kappa TM}{T_c}\frac{1}{NM}\mathrm{tr}\{E[\bar{G}\bar{G}^\dagger]\} \\
&= \sigma^2 + \frac{(1-r)\kappa TM}{T_c}\sigma^2_{\bar{G}}
\end{aligned}$$

and the lower bound is

$$C_t/T \geq \frac{T-T_t}{T} E \log\det\left(I_M + \frac{\rho_{eff}}{M} H_1 \Lambda H_1^\dagger\right) \tag{6.11}$$

where ρ_{eff} is the ratio of the energies in the elements of $S_c\hat{H}$ and energies in the elements of $\hat{W}_c$ and $H_1 = \sqrt{r_{new}}H_m + \sqrt{1-r_{new}}\hat{G}$ where $r_{new} = \frac{r}{r+(1-r)\sigma^2_{\hat{G}}}$. Λ is the optimum signal correlation matrix the form of which depends on the distribution of H_1 according to Proposition 6.2 for low SNR and Proposition 6.1 for high SNR and $M \leq N$ as given in section 6.3.1.

To calculate ρ_{eff}, the energy in the elements of $S\hat{H}$ is given by

$$\begin{aligned}\sigma_{SH}^2 &= \frac{1}{NT_c}[r\text{tr}\{H_m H_m^\dagger \kappa T I_M\} + (1-r)\text{tr}\{\hat{G}\hat{G}^\dagger \kappa T I_M\}] \\ &= \frac{\kappa T M}{T_c}\frac{1}{NM}[rNM + (1-r)\text{tr}\{\hat{G}\hat{G}^\dagger\}] \\ &= \frac{\kappa T M}{T_c}[r + (1-r)\sigma_{\hat{G}}^2]\end{aligned}$$

which gives us

$$\rho_{eff} = \frac{\kappa T\rho[r + (1-r)\sigma_{\hat{G}}^2]}{T_c + (1-r)\kappa T\rho\sigma_{\hat{G}}^2}.$$

6.4.1 Optimization of S_t, κ and T_t

We will optimize S_t, κ and T_t to maximize the lower bound (6.11). In this section we merely state the main results and their interpretations. Derivations and details are given in the Appendices.

Optimization of the lower bound over S_t is difficult as S_t effects the distribution of $\hat{H}$, the form of Λ as well as ρ_{eff}. To make the problem simpler we will just find the value of S_t that maximizes ρ_{eff}.

Theorem 6.7. *The signal S_t that maximizes ρ_{eff} satisfies the following condition*

$$S_t^\dagger S_t = (1-\kappa)TI_M$$

and the corresponding ρ_{eff} is

$$\rho_{eff}^* = \frac{\kappa T\rho[Mr + \rho(1-\kappa)T]}{T_c(M + \rho(1-\kappa)T) + (1-r)\kappa T\rho M}.$$

Proof: See Appendix B.6. ■

The optimum signal derived above is the same as the optimum signal derived in [39].

The corresponding capacity lower bound using the S_t obtained above is

$$C_t/T \geq \frac{T-T_t}{T} E \log\det\left(I_M + \frac{\rho_{eff}}{M} H_1 \Lambda H_1^\dagger\right)$$

where ρ_{eff} is as given above and $H_1 = \sqrt{r_{new}} H_m + \sqrt{1-r_{new}} G$ where $r_{new} = r\frac{1+(1-r)(1-\kappa)\frac{\rho}{M}T}{r+(1-r)(1-\kappa)\frac{\rho}{M}T}$ and as before G is a matrix consisting of i.i.d. Gaussian circular random variables with mean zero and unit variance. Now, Λ is the covariance matrix of the source S_c when the channel is Rician and known to the receiver. The form of Λ was derived for $\rho_{eff} \to 0$ and $\rho_{eff} \to \infty$ in section 6.3.1.

Optimization of (6.11) over κ is straightforward as κ affects the lower bound only through ρ_{eff} and can be stated in the following proposition.

Theorem 6.8. *For fixed T_t and T_c the optimal power allocation κ in a training based scheme is given by*

$$\kappa = \begin{cases} \min\{\gamma - \sqrt{\gamma(\gamma-1-\eta)}, 1\} & \text{for } T_c > (1-r)M \\ \min\{\frac{1}{2} + \frac{rM}{2T\rho}, 1\} & \text{for } T_c = (1-r)M \\ \min\{\gamma + \sqrt{\gamma(\gamma-1-\eta)}, 1\} & \text{for } T_c < (1-r)M \end{cases}$$

where $\gamma = \frac{MT_c + T\rho T_c}{T\rho[T_c - (1-r)M]}$ and $\eta = \frac{rM}{T\rho}$. The corresponding lower bound is given by

$$C_t/T \geq \frac{T-T_t}{T} E \log\det\left(I_M + \frac{\rho_{eff}}{M} H_1 \Lambda H_1^\dagger\right)$$

where for $T_c > (1-r)M$

$$\rho_{eff} = \begin{cases} \frac{T\rho}{T_c-(1-r)M}(\sqrt{\gamma} - \sqrt{\gamma-1-\eta})^2 & \text{when } \kappa = \gamma - \sqrt{\gamma(\gamma-1-\eta)} \\ \frac{r\rho}{1+(1-r)\rho} & \text{when } \kappa = 1 \end{cases}$$

for $T_c = (1-r)M$

$$\rho_{eff} = \begin{cases} \frac{T^2\rho^2}{4(1-r)M(M+T\rho)}(1+\frac{rM}{T\rho})^2 & \text{when } \kappa = \frac{1}{2} + \frac{rM}{2T\rho} \\ \frac{rT\rho}{(1-r)(M+T\rho)} & \text{when } \kappa = 1 \end{cases}$$

and for $T_c < (1-r)M$

$$\rho_{eff} = \begin{cases} \frac{T\rho}{(1-r)M-T_c}(\sqrt{-\gamma} - \sqrt{-\gamma+1+\eta})^2 & when\ \kappa = \gamma + \sqrt{\gamma(\gamma-1-\eta)} \\ \frac{r\rho}{1+(1-r)\rho} & when\ \kappa = 1 \end{cases}$$

and r_{new} *is given by substituting the appropriate value of* κ *in the expression*

$$r\frac{1+(1-r)(1-\kappa)\frac{\rho}{M}T}{r+(1-r)(1-\kappa)\frac{\rho}{M}T}.$$

Proof: See Appendix B.7. ■

For optimization over T_t we draw somewhat similar conclusions as in [39]. In [39] the optimal setting for T_t was shown to be $T_t = M$ for all values of SNR. We however show that for small SNR the optimal setting is $T_t = 0$ or that no training is required. When training is required, the intuition is that increasing T_t linearly decreases the capacity through the term $(T - T_t)/T$, but only logarithmically increases the capacity through the higher effective SNR ρ_{eff} [39]. Therefore, it makes sense to make T_t as small as possible. For small SNR we show that $\kappa = 1$. It is clear that optimization of T_t makes sense only when κ is strictly less than 1. When $\kappa = 1$ no power is devoted to training and T_t can be made as small as possible which is zero. When $\kappa < 1$ the smallest value T_t can be is M since it takes atleast that many intervals to completely determine the unknowns.

Theorem 6.9. *The optimal length of the training interval is* $T_t = M$ *whenever* $\kappa < 1$ *for all values of* ρ *and* $T > M$*, and the capacity lower bound is*

$$C_t/T \geq \frac{T-M}{T} E \log\det\left(I_M + \frac{\rho_{eff}}{M} H_1 \Lambda H_1^\dagger\right) \tag{6.12}$$

where

$$\rho_{eff} = \begin{cases} \frac{T\rho}{T-(2-r)M}(\sqrt{\gamma} - \sqrt{\gamma-1-\eta})^2 & for\ T > (2-r)M \\ \frac{T^2\rho^2}{4(1-r)M(M+T\rho)}(1+\frac{rM}{T\rho})^2 & for\ T = (2-r)M \\ \frac{T\rho}{T-(2-r)M}(\sqrt{-\gamma} - \sqrt{-\gamma+1+\eta})^2 & for\ T < (2-r)M \end{cases}$$

The optimal power allocations are easily obtained from Theorem 6.8 by simply setting $T_c = T - M$.

Proof: See Appendix B.8 ■

6.4.2 Equal training and data power

As stated in [39], sometimes it is difficult for the transmitter to assign different powers for training and communication phases. In this section, we will concentrate on setting the training and communication powers equal to each other in the following sense

$$\frac{(1-\kappa)T}{T_t} = \frac{\kappa T}{T_c} = \frac{\kappa T}{T - T_t} = 1$$

this means $\kappa = 1 - T_t/T$ and that the power transmitted in T_t and T_c are equal.

In this case,

$$\rho_{eff} = \frac{\rho[r + \rho\frac{T_t}{M}]}{1 + \rho[\frac{T_t}{M} + (1-r)]}$$

and the capacity lower bound is

$$C_t/T \geq \frac{T - T_t}{T} E \log\det(I_M + \frac{\rho_{eff}}{M} H_1 \Lambda H_1^\dagger)$$

where ρ_{eff} is as given above and $H_1 = \sqrt{r_{new}} H_m + \sqrt{1 - r_{new}} G$ where $r_{new} = r\frac{1+(1-r)\frac{\rho}{M}T_t}{r+(1-r)\frac{\rho}{M}T_t}$.

6.4.3 Numerical Comparisons

Throughout the section we have chosen the number of transmit antennas M, and receive antennas N, to be equal and $H_m = I_M$.

The Figures 6.3 and 6.4 show r_{new} and κ respectively as a function of r for different values of SNR. The plots have been calculated for a block length given by $T = 40$ and the number of transmit and receive antennas given by $M = N = 5$. Figure 6.3 shows that for low SNR values the channel behaves like a purely AWGN channel given by

$\sqrt{r}H_m$ and for high SNR values the channel behaves exactly like the original Rician fading channel. Figure 6.4 shows that as the SNR goes to zero less and less power is allocated for training. This agrees with the plot in Figure 6.3.

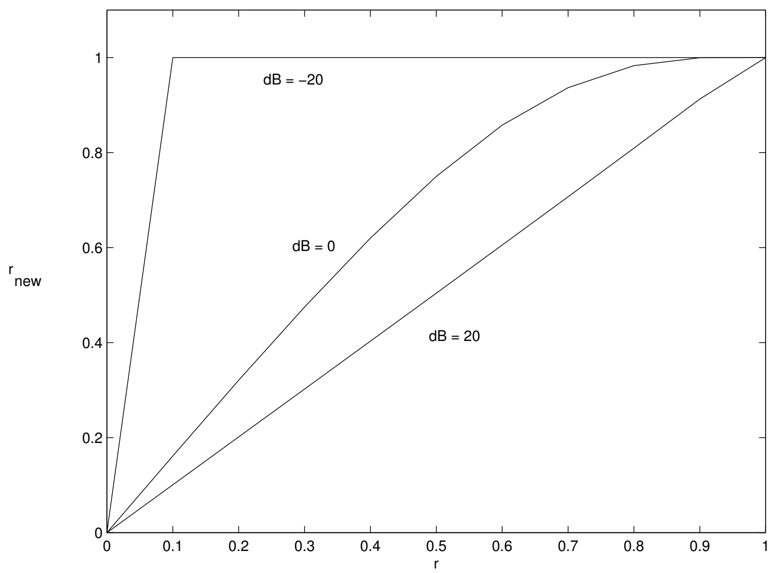


Figure 6.3: Plot of r_{new} as a function of Rician parameter r

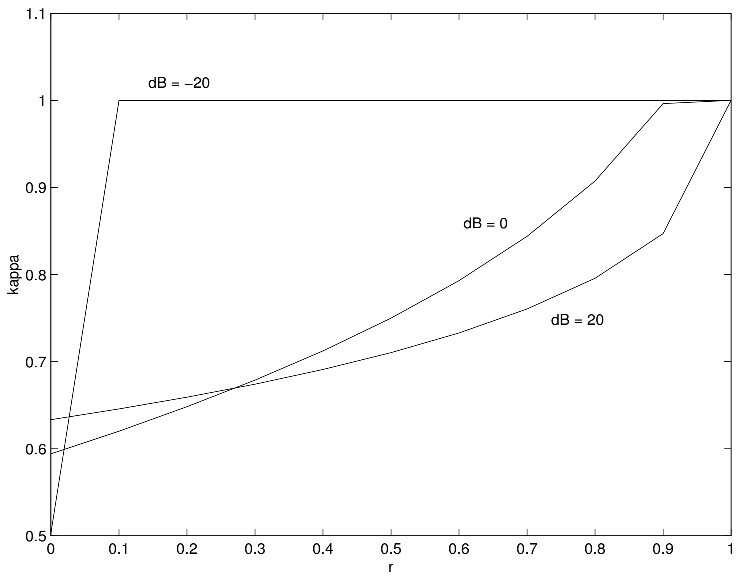


Figure 6.4: Plot of optimal energy allocation κ as a function of Rician parameter r

In Figure 6.5 we plot the training and communication powers for $M = N = 10$ and $dB = 18$ for different values of r. We see that as r goes to 1 less and less

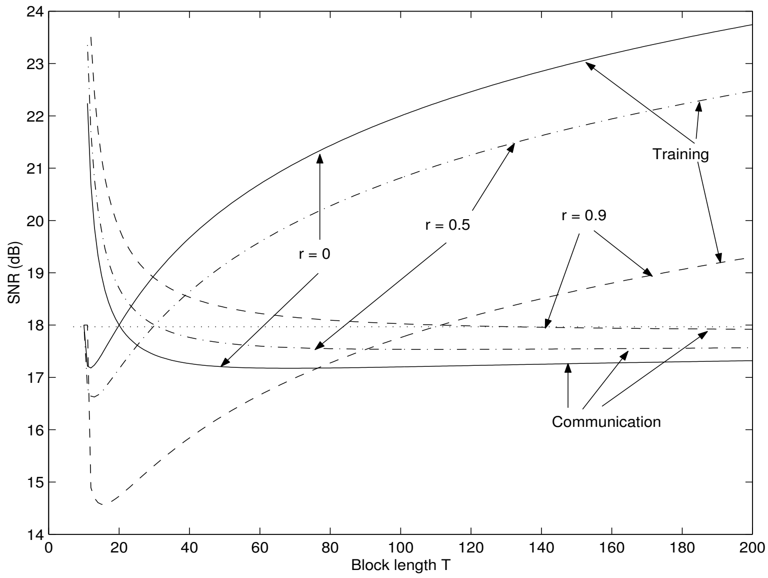


Figure 6.5: Plot of optimal power allocation as a function of T

power is allocated to the training phase. This makes sense as the proportion of the energy through the specular component increases there is less need for the system to estimate the unknown Rayleigh component.

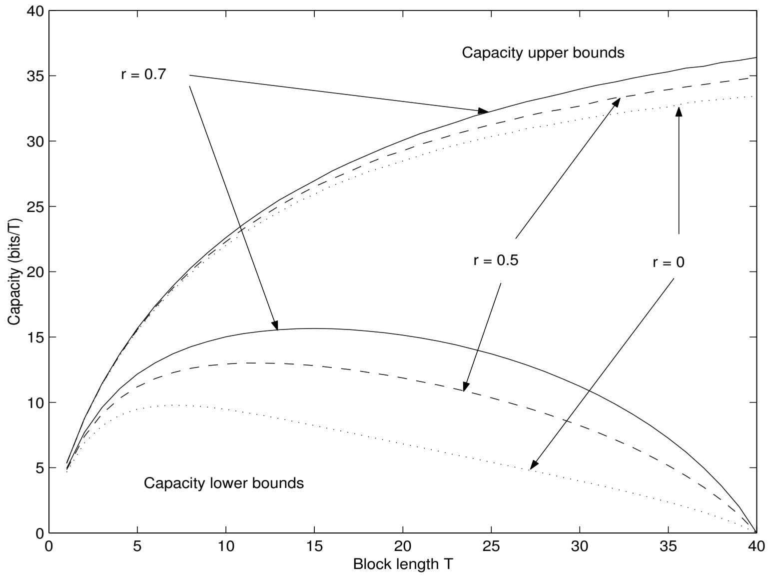


Figure 6.6: Plot of capacity as a function of number of transmit antennas for a fixed T

Figure 6.6 shows capacity as a function of the number of transmit antennas for a fixed block length $T = 40$ when $dB = 0$ and $N = 40$. We can easily see calculate the optimum number of transmit antennas from the figure. In this case, we see that for a fixed T the optimum number of transmit antennas increases as as r increases. This shows that as r goes to 1 there is a lesser need to estimate the unknown Rayleigh part of the channel and this agrees very well with Figure 6.5 and Figure 6.7 as well which shows that the optimal amount of training decreases as r increases. Figure 6.7 shows the optimal training period as a function of the block length for the case of equal transmit and training powers.

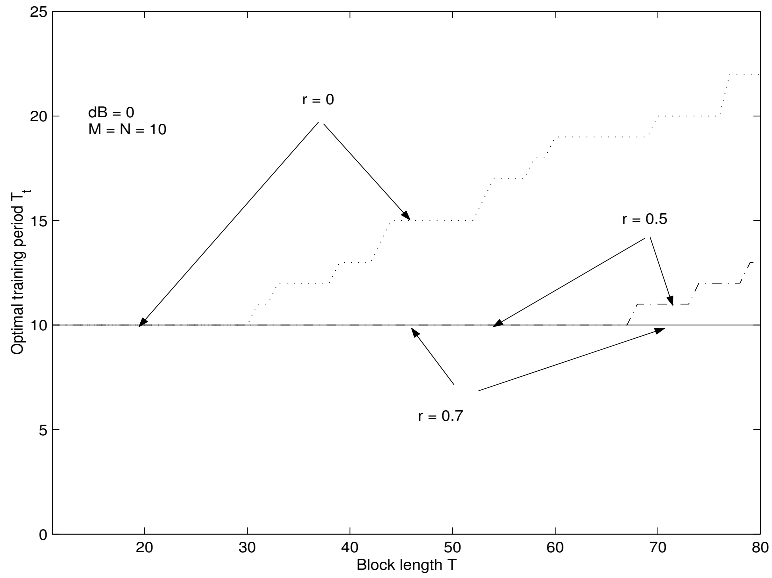


Figure 6.7: Optimal T_t as a function of T for equal transmit and training powers

6.4.4 Effect of Low SNR on Capacity Lower Bound

Let's consider the effect of low SNR on the optimization of κ when $r \neq 0$. For $T_c > (1-r)M$, as $\rho \rightarrow 0$ it is easy to see that $\gamma - \sqrt{\gamma(\gamma - 1 - \eta)} \rightarrow \infty$. Therefore, we conclude that for small ρ we have $\kappa = 1$. Similarly, for $T_c = (1-r)M$ and $T_c < (1-r)M$. Therefore, the lower bound tells us that no energy need be spent on training for small ρ. Also, the form of Λ is known from section 6.3.1.

Evaluating the case where the training and transmission powers are equal we come to a similar conclusion. For small ρ, $\rho_{eff} \approx r\rho$ which is independent of T_t. Therefore, the best value of T_t is $T_t = 0$. Which also means that we spend absolutely no time on training. This is in stark contrast to the case when $r = 0$. In this case, for low SNR $T_t = T/2$ [39] and ρ_{eff} behaves as $O(\rho^2)$.

Note that in both cases of equal and unequal power distribution between training and communication phases the signal distribution during data transmission phase is Gaussian. Therefore, the lower bound behaves as $r\rho\lambda_{max}\{H_m H_m^\dagger\}$. Also, $r_{new} = 1$ for small ρ showing that the channel behaves as a purely Gaussian channel.

All the conclusions above mimic those of the capacity results with Gaussian input in section 6.3.3. The low SNR non-coherent capacity results for the case of a Gaussian input tell us that the capacity behaves as $r\rho\lambda_{max}$ with Gaussian input. Moreover, the results in [39] also agree with the results derived in section 6.3.3. We showed that for purely Rayleigh fading channels with Gaussian input the capacity behaves as ρ^2 which is what the lower bound results in [39] also show. This makes sense because the capacity lower bound assumes that the signaling input during communication period is Gaussian. This shows that the lower bound derived in [39] and extended here is quite tight for low SNR values.

6.4.5 Effect of High SNR on Capacity Lower Bound

For high SNR, γ becomes $\frac{T_c}{T_c-(1-r)M}$ and the optimal power allocation κ becomes

$$\kappa = \frac{\sqrt{T_c}}{\sqrt{T_c} + \sqrt{(1-r)M}}$$

and

$$\rho_{eff} = \frac{T}{(\sqrt{T_c} + \sqrt{(1-r)M})^2}\rho.$$

In the case of equal training and transmit powers, we have for high ρ

$$\rho_{eff} = \rho\frac{T_t}{T_t + M(1-r)}.$$

For high SNR, the channel behaves as if it is completely known to the receiver. Note that in this case $r_{new} = r$ and Λ is an identity matrix for the case $M \leq N$.

From the expressions for ρ_{eff} given above we conclude that unlike the case of low SNR the value of r affects the amount of time and power devoted for training.

Let's look at the capacity lower bound for high SNR. The optimizing Λ in this regime is an identity matrix. We know that at high SNR the optimal training period is M. Therefore, the resulting lower bound is given by

$$C_t/T \geq \frac{T-M}{T} E \log\det\left(I_M + \frac{\rho}{\left(\sqrt{1-\frac{M}{T}} + \sqrt{\frac{(1-r)M}{T}}\right)^2}\frac{HH^\dagger}{M}\right).$$

Note that the lower bound has H figuring in it instead of H_1. That is so because for high SNR, $r_{new} = r$. This lower bound can be optimized over the number of transmit antennas used in which case the lower bound can be rewritten as

$$C_t/T \geq \max_{M' \leq M} \max_{n \leq \binom{M}{M'}} \frac{T-M'}{T} E \log\det\left(I_{M'} + \frac{\rho}{\left(\sqrt{1-\frac{M'}{T}} + \sqrt{\frac{(1-r)M'}{T}}\right)^2}\frac{H^n H^{n\dagger}}{M'}\right)$$

where now H^n is the n^{th} matrix out of a possible M choose M' (the number of ways to choose M' transmit elements out of a maximum M elements) matrices of size $M' \times N$. Let $Q = \min\{M', N\}$ and λ_i^n be an arbitrary nonzero eigenvalue of $\frac{1}{\left(\sqrt{1-\frac{M'}{T}} + \sqrt{\frac{(1-r)M'}{T}}\right)^2}\frac{H^n H^{n\dagger}}{M'}$ then we have

$$C_t/T \geq \max_{M' \leq M} \max_{n \leq \binom{M}{M'}} \left(1 - \frac{M'}{T}\right) \sum_{i=1}^{Q} E \log(1 + \rho\lambda_i^n).$$

At high SNR, the leading term involving ρ in $\sum_{i=1}^{Q} E\log(1+\rho\lambda_i)$ is $Q\log\rho$ which is independent of n. Therefore,

$$C_t/T \geq \max_{M' \leq M} \begin{cases} (1-\frac{M'}{T})M'\log\rho & \text{if } M' \leq N \\ (1-\frac{M'}{T})N\log\rho & \text{if } M > N. \end{cases}$$

The expression $(1-\frac{M'}{T})M'$, is maximized by choosing $M' = T/2$ when $\min\{M,N\} \geq T/2$ and by choosing $M' = \min\{M,N\}$ when $\min\{M,N\} \leq T/2$. This means that the expression is maximized when $M' = \min\{M,N,T/2\}$. This is a similar conclusion drawn in [39] and [85]. Also, the leading term in ρ for high SNR in the lower bound is given by

$$C_t/T \geq (1-\frac{K}{T})K\log\rho$$

where $K = \min\{M,N,T/2\}$. This result suggests that the number of degrees of freedom available for communication is limited by the minimum of the number of transmit antennas, receive antennas and half the length of the coherence interval. Moreover, from the results in section 6.3.4 we see that the lower bound is tight for the case when $M \leq N$ and large T in the sense that the leading term involving ρ in the lower bound is the same as the one in the expression for capacity.

6.4.6 Comparison of the training based lower bound (6.12) with the lower bound derived in section 6.3.2

It is quite natural to use the lower bound to investigate training based techniques as the lower bound to the overall capacity of the system. Actually, using this "training" based lower bound it can be shown that the capacity as $T \to \infty$ converges to the capacity as if the receiver knows the channel. We will see how the new lower bound (6.8) derived in this work compares with this training based lower bound. The three Figures below show that the new lower bound is indeed useful as it does better than the training based lower bound for $r = 0$. The plots are for $M = N = 1$ for different values of SNR.

However, we note that for $r = 1$ the training based lower bound and the lower bound derived in section 6.3.2 agree perfectly with each other and are equal to the upper bound.

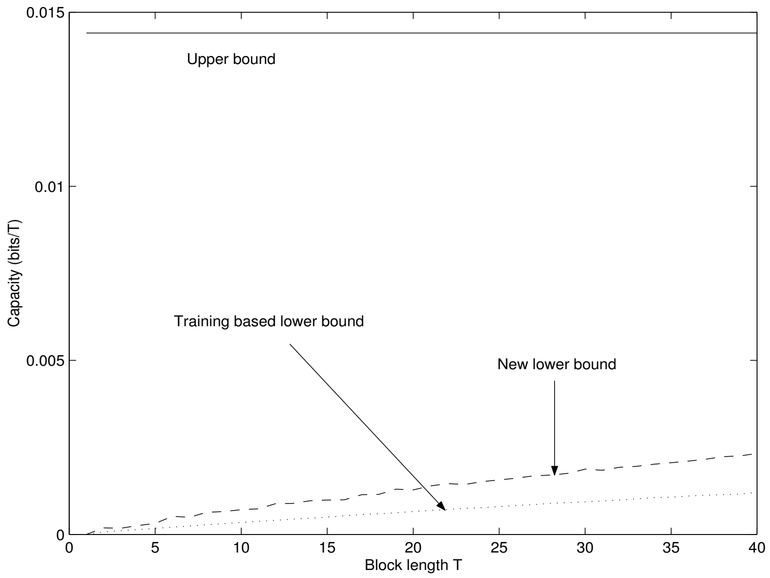


Figure 6.8: Comparison of the two lower bounds for $dB = -20$

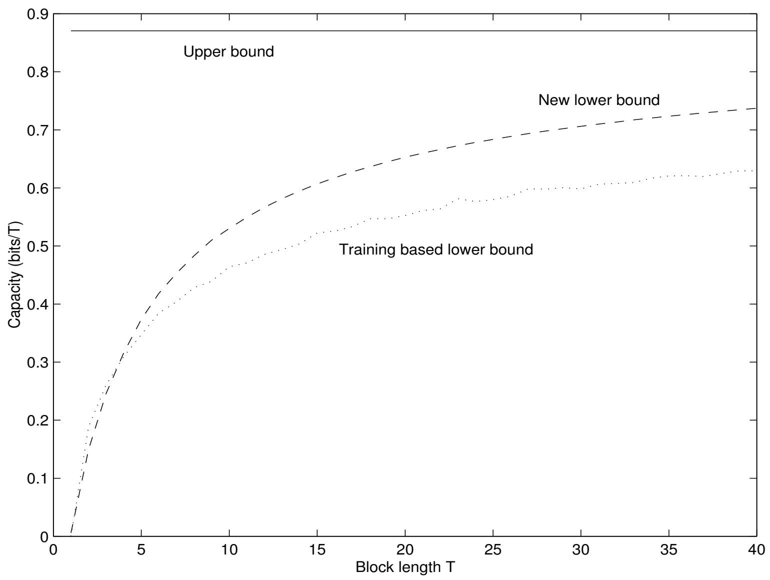


Figure 6.9: Comparison of the two lower bounds for $dB = 0$

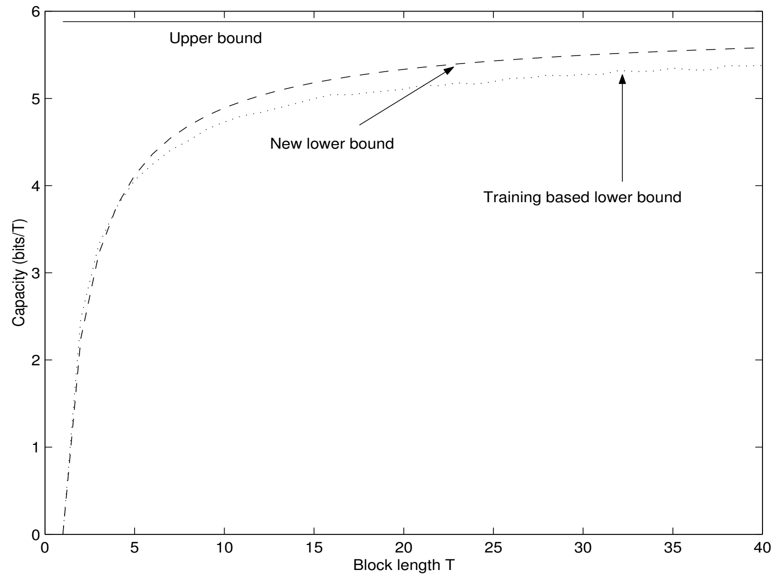


Figure 6.10: Comparison of the two lower bounds for $dB = 20$

6.5 Conclusions and Future Work

In this chapter, we have analyzed the standard Rician fading channel for capacity. Most of the analysis was for a general specular component but, for the special case of a rank-one specular component we were able to show more structure on the signal input. For the case of general specular component, we were able to derive asymptotic closed form expressions for capacity for low and high SNR scenarios.

A big part of the analysis e.g. the non-coherent capacity expression and training based lower bounds can be very easily extended to the non-standard Rician models considered in the previous two chapters.

One important result of the analysis is that for low SNRs beamforming is very desirable whereas for high SNR scenarios it is not. This result is very useful in designing space-time codes. For high SNR scenarios, one could wager that the standard codes designed for Rayleigh fading can work for the case of Rician fading as well.

A lot more work needs to be done such as for the case of $M > N$. We believe that more work along the lines of [85] is possible for the case of Rician fading. We conclude as in [85] that at least for the case $M = N$ the number of degrees of freedom is given by $M\frac{T-M}{T}$. The training based lower bound gives an indication that the number of degrees of freedom of a Rician channel is the same as that of a Rayleigh fading channel $\min\{M, N, T/2\}$ (derived in [85] and [39]). It also seems reasonable that the work in [1] can be extended to the case of Rician fading.

CHAPTER 7

Diversity versus Degrees of Freedom

7.1 Introduction

Two measures of performance gains obtained from using multiple antennas at the transmitter and the receiver in the field of space-time coding are Diversity [71] and Degrees of Freedom [85]. In [71] Diversity (DIV) has been defined as the negative exponent of the signal to noise ratio (SNR) in the probability of error expression for high SNR and in [85] Degrees of Freedom (DOF) has been defined as the co-efficient of $\log \rho$ occuring in the expression for capacity, again for high SNR.

Traditionally, DIV has been thought of as the number of independent channels available for communication [23]. However, Zheng and Tse [85] have shown that DOF is the true indicator of the number of independent channels available in a system. This has given rise to considerable confusion as DIV and DOF for a MIMO system operating in a Rayleigh/Rician fading environment don't necessarily agree with each other. Even though [71] refers to DIV as redundancy in the system it doesn't clarify the difference between the two measures. In this chapter, we attempt to shed some light on this confusing situation. We adopt a more general setting than just a MIMO system operating in a fading environment. We propose that DIV should properly be considered as the redundancy in a particular communication system whereas DOF should be considered as the number of independent channels available for communication again in a particular communication system. A communication system comprises of the channel, the input signaling scheme, coding, decoding etc.

First, we define the following terms. Let H be the channel and S be the input to the channel. Let X be the observed output, i.e. received measurement, which is a corrupted version of the signal component of the received measurement X, Y which is completely determined by S and H. For example, the M-transmit, N-receive antenna MIMO system considered in this thesis can be written as $Y = \sqrt{\frac{\rho}{M}} SH$ and $X = Y + W$ where S is a $T \times M$ matrix, H is an $M \times N$ matrix and X, Y and W are $T \times N$ matrices. W is the "corruption" in the system or the noise in the observation. ρ is the average signal to noise ratio present at each of the receive antennas. Note that we have deviated slightly from the previous chapters in regards to the usage of the word "output". In the previous chapters we have referred to X as the output of the channel. In this chapter Y is the signal dependent output of the channel whereas X is the observed output.

So far in the literature the definitions of DIV and DOF have been fairly ad hoc in the sense that the definitions are particular to the case of MIMO systems operating in Rayleigh/Rician fading environments. We attempt to remedy this by proposing two rigorous definitions of DIV and DOF. For this we require some additional terminology. Let $\mathcal{P} = \{P_1, P_2, \ldots, P_n\}$ be the set of all link parameters in the communication system. For example, P_1 could be M, the number of transmit antennas; P_2 could be N, the number of receive antennas; and P_3 could be t, the number of channel uses. The channel specified by $\mathcal{P}$ in general could be stochastic, for example, H in the Rayleigh fading model. Let $P_{eH}(\rho, P_1, P_2, \ldots, P_n)$ denote the least (best) possible probability of error when the parameters of interest are $(P_1, P_2, \ldots, P_n)$ and the channel H specified by $\mathcal{P}$ is known to the receiver. For example, the $M \times N$ channel matrix H formed by M transmit antennas and N receive antennas is known to the receiver in the case of coherent communications. ρ is the measure of reliability in the channel with the channel becoming more reliable as $\rho \to \infty$. For example, in the MIMO system the reliability measure ρ is the signal to noise ratio. In the binary symmetric channel with crossover probability p, $p \to 0$ is equivalent to $\rho \to \infty$. Let $P_e(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})$ denote the probability of error in a particular communication system $\mathcal{C}$. $\mathcal{C}$ is specified by the input probability distribution $p(S)$, the decoding structure/strategy (coherent vs non-coherent/hard vs soft decoding etc.) and the transmitter structure (coding etc.). Similarly, let $C_H(\rho, P_1, P_2, \ldots, P_n)$ denote the maximum (best) rate of communication when the channel is known to the receiver and $R(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})$ denote the rate of communication in the particular system $\mathcal{C}$. Then, we have the following definitions

$$\mathrm{DIV}_{P_1,P_2,\ldots,P_n}(\mathcal{C}) = \lim_{\rho\to\infty} \frac{\log P_e(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})}{\log P_{eH}(\rho, 1, 1, \ldots, 1)} \tag{7.1}$$

and

$$\mathrm{DOF}_{P_1,P_2,\ldots,P_n}(\mathcal{C}) = \lim_{\rho\to\infty} \frac{R(\rho, P_1, P_2, \ldots, P_n, \mathcal{C})}{C_H(\rho, 1, 1, \ldots, 1)} \tag{7.2}$$

where $P_{eH}(\rho, 1, 1, \ldots, 1)$ and $C_H(\rho, 1, 1, \ldots, 1)$ denote the optimum probability of error and optimum rate evaluated at $P_1 = P_2 = \ldots = P_n = 1$. We will assume that the above limits exist whenever required. DIV and DOF for the channel are defined as

$$\mathrm{DIV}_{P_1,P_2,\ldots,P_n} = \sup_{\mathcal{C}} \mathrm{DIV}_{P_1,P_2,\ldots,P_n}(\mathcal{C}) \tag{7.3}$$

and

$$\mathrm{DOF}_{P_1,P_2,\ldots,P_n} = \sup_{\mathcal{C}} \mathrm{DOF}_{P_1,P_2,\ldots,P_n}(\mathcal{C}). \tag{7.4}$$

Note the following

$$\begin{aligned}
\mathrm{DOF}_{P_1,P_2,\ldots,P_n} &= \sup_{\mathcal{C}} \ \liminf_{\rho\to\infty} \frac{R(\rho,P_1,P_2,\ldots,P_n,\mathcal{C})}{C_H(\rho,1,1,\ldots,1)} \\
&\leq \liminf_{\rho\to\infty} \sup_{\mathcal{C}} \frac{R(\rho,P_1,P_2,\ldots,P_n,\mathcal{C})}{C_H(\rho,1,1,\ldots,1)} \\
&= \liminf_{\rho\to\infty} \frac{C(\rho,P_1,P_2,\ldots,P_n)}{C_H(\rho,1,1,\ldots,1)}.
\end{aligned}$$

But, $\mathrm{DOF}_{P_1,P_2,\ldots,P_n} \geq \lim_{\rho\to\infty} \frac{C(\rho,P_1,P_2,\ldots,P_n)}{C_H(\rho,1,1,\ldots,1)}$ by definition. Therefore,

$$\mathrm{DOF}_{P_1,P_2,\ldots,P_n} = \lim_{\rho\to\infty} \frac{C(\rho,P_1,P_2,\ldots,P_n)}{C_H(\rho,1,1,\ldots,1)}. \tag{7.5}$$

From now on, we will use (7.5) as the definition for DOF of a channel as contrasted with (7.4).

7.2 Examples

We will apply the above definitions to various examples. For some of the examples, we need a little bit of detail on Gallager's error exponents [26, 27]. This introduction was given in Section 5.6 but is reproduced here for convenience. For a system communicating at a rate R the upper bound on probability of error is given as follows

$$P_e \leq \exp\left(-n \max_{p(S)} \max_{0\leq\gamma\leq 1} \left[E_0(\gamma,p(S)) - \gamma R\right]\right) \tag{7.6}$$

where n is the length of codebook used and $E_0(\gamma,p(S))$ is as follows

$$E_0(\gamma,p(S)) = -\log \int \left[\int p(S)p(X|S)^{\frac{1}{1+\gamma}} dS\right]^{\gamma} dX$$

where S is the input to the channel and X is the observed output. Maximization over γ yields a value of γ such that $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma} = R$. If R is chosen equal to $I(X;S)$ (the mutual information between the output and the input when the input probability distribution is $p(S)$) then the value of γ such that $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma} = R$ is zero. In other words $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma} = I(X;S)$ at $\gamma = 0$. If $R > I(X;S)$ then there is no value of γ that satisfies the relation $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma} = R$. If $p(S)$ is chosen to be the capacity achieving signal density then $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma} = C$ at $\gamma = 0$. It has been shown that $\frac{\partial E_0(\gamma,p(S))}{\partial\gamma}$ is a decreasing function of γ [27]. Therefore, if R is small enough then the value of γ that maximizes $E_0(\gamma, p(S)) - \gamma R$ is simply 1 and the error bound is given by

$$P_e \leq \exp\left(-n \max_{p(S)}[E_0(1, p(S)) - R]\right).$$

For more information refer to [27].

Example 1: Consider a single-input single-output discrete-time AWGN channel $x_l = \sqrt{\rho} h s_l + w_l$, $l = 1, \ldots, t$. h is a deterministic complex number and w_l is a complex circular Gaussian random variable with mean zero and variance 1. We will calculate DIV and DOF for t-uses of the channel under the average energy constraint $\sum_{l=1}^{t} E[s_l s_l^*] \leq t$.

First, consider $\mathcal{C}_1$ consisting of a binary input taking values over $\{-1, 1\}$ with equal probability. The code used is a repetition code. That is either $s_l = -1$ for $l = 1, \ldots, t$ or $s_l = 1$ for $l = 1, \ldots, t$. The decoding at the receiver is Maximum A Posteriori (MAP) decoding. Since the probability of error when using the repetition code of length t is given by $Q(t\rho)$ where

$$Q(x) = \frac{1}{2\pi} \int_x^{-\infty} e^{-y^2/2} dy$$

as opposed to $Q(\rho)$ when $t = 1$, we obtain $\mathrm{DIV}_t(\mathcal{C}_1) = t$ Since $R(\rho, t, \mathcal{C}_1) = 1/t$ irrespective of the value of ρ, $\mathrm{DOF}_t(\mathcal{C}_1) = 0$.

Next, consider $\mathcal{C}_2$ where s_l is an i.i.d. complex circular Gaussian random variable with mean zero and unit variance for $l = 1, \ldots, t$. This is the capacity achieving signal density. The decoding at the output is MAP decoding. In this case, $\mathrm{DOF}_t(\mathcal{C}_2) = t$ since $R(\rho, t, \mathcal{C}_2) = t \log(1 + \rho)$. For this value of R ($R = C$, the capacity) the value of γ that maximizes Gallager's error exponent is zero and at $\gamma = 0$, $E_0(\gamma, p(S)) = 1$. Therefore, $\mathrm{DIV}_t(\mathcal{C}_2) = 0$.

Therefore, $\mathrm{DIV}_t \geq t$ and $\mathrm{DOF}_t = t$. It can be obviously seen that $\mathrm{DIV}_t \leq t$. Therefore, $\mathrm{DIV}_t = t$. ■

We believe that DIV_t being equal to DOF_t in this example can cause confusion between DIV and DOF.

Example 2: Consider a MIMO AWGN channel with M inputs and N outputs

$$X_l = \sqrt{\frac{\rho}{M}} H S_l + W_l$$

for $l = 1, \ldots, t$. X_l and W_l are length N column vectors with the elements of W_l i.i.d. complex circular Gaussian random variables with mean zero and variance 1 and S_l is an M-element column vector. Assume that H is a $N \times M$ matrix which has the full rank of $\min(M, N)$ and that $\mathrm{tr}\{HH^\dagger\} = MN$. We propose to calculate DIV and DOF for this system for t channel uses under the average energy constraint on the input $\sum_{l=1}^{t} \mathrm{tr}\{E[S_l S_l^\dagger]\} \leq tM$. Consider three communication systems, $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$.

In $\mathcal{C}_1$, the rate R is fixed for all ρ. In such a case the upper bound on the probability of error can be written as follows

$$P_e \leq \exp\left(- \max_{p(S)} [E_0(1, p(S)) - R] \right).$$

If we fix $p(S)$ to be the capacity achieving signal density $p_G(S)$, then

$$E_0(1, p_G(S)) = -\log \left[\det \left(I_M + \frac{1}{2} \frac{\rho}{M} H^\dagger H \right)^{-t} \right]$$

Since $E_0(1, p_G(S))$ for high SNR tends to $t \min(M, N) \log \rho$, we conclude that

$$\mathrm{DIV}_{M,N,t}(\mathcal{C}_1) = t \min(M, N).$$

In $\mathcal{C}_2$, the input is binary $\{1, -1\}$, and the decoding at the output is MAP decoding. The code is such that when the symbol 1 is chosen the signal transmitted is $S_l = a_l$, $l = 1, \dots, t$ where a_l is an M-element column vector to be specified later. When the symbol -1 is chosen the signal transmitted is simply $S_l = -a_l$, $l = 1, \dots, t$. Let $A = [a_1 \; a_2 \; \dots \; a_t]$ be the $M \times t$ matrices obtained by stacking the t column vectors next to each other. Choose a_l, $l = 1, \dots, t$ such that $AA^\dagger = tI_M$. In this case, the probability of 1 being decoded to -1 for high SNR is given by $Q(\sqrt{d^2(1,-1)\frac{\rho}{4M}})$ with $d^2(1,-1) = 4\text{tr}\{HAA^\dagger H^\dagger\}$. Since $d^2(1,-1) = 4MNt$, the probability of decoding error for high SNR can be upper bounded by $\exp(-\rho t N)$. Therefore, $\text{DIV}_{M,N,t}(\mathcal{C}_2) = Nt$. Since the rate of communication is 1 bit per t channel uses irrespective of ρ, $\text{DOF}_{M,N,t}(\mathcal{C}_2) = 0$.

In $\mathcal{C}_3$, the probability distribution function of S_l is chosen to achieve capacity and the decoding is chosen to be MAP decoding. Then $R(\rho, M, N, t, \mathcal{C}_3) = C(\rho, M, N, t) = t \log\det(I_N + \frac{\rho}{M} H \Lambda_H H^\dagger)$ where Λ_H is the capacity achieving covariance matrix. $R(\rho, M, N, t, \mathcal{C}_3)$ for high SNR tends to $t \min(M, N) \log \rho$. Therefore, $\text{DOF}_{M,N,t}(\mathcal{C}_3) = \text{DOF}_{M,N,t} = t \min(M, N)$. ■

In Example 2 we see that DIV is greater than or equal to DOF.

Example 3: Now, let's consider a MIMO system with M-transmit and N-receive antennas operating in a Rayleigh fading environment:

$$X = \sqrt{\frac{\rho}{M}} SH + W.$$

X is a $T \times N$ matrix, S is a $T \times M$ matrix and W is a $T \times N$ matrix. Note that now H is a $M \times N$ matrix. The elements of H and W are i.i.d complex circular Gaussian random variables with mean zero and variance one. Let the block length of independent fades be T. We will investigate DIV and DOF for $t = T$ channel uses under the average energy constraint $\text{tr}\{E[SS^\dagger]\} \leq tM$. We assume that H is known to the receiver. Let's again consider two systems $\mathcal{C}_1$ and $\mathcal{C}_2$.

In $\mathcal{C}_1$, we use the signaling scheme developed by Tarokh et. al. [71, pp. 747–749]. The decoding at the output is chosen to be MAP decoding. Using our definition for DIV and Tarokh's development for probability of error [71, (10), p. 749] we conclude that $\mathrm{DIV}_{M,N,t}(\mathcal{C}_1) = N \min(M, t)$ which agrees with Tarokh's conclusion about diversity. However, from [71, (18), p. 755] we see that the rate R is bounded above by a constant independent of ρ. Therefore, $\mathrm{DOF}_{M,N,t}(\mathcal{C}_1) = 0$.

In $\mathcal{C}_2$, we choose the elements of the matrix S to be i.i.d. complex circular Gaussian random variables with mean zero and variance one. This is the capacity achieving signal. We choosing the decoding strategy to be MAP. Therefore [50],

$$R(\rho, M, N, t, \mathcal{C}_2) = C(\rho, M, N, t) = tE \log\det(I_M + \frac{\rho}{M} HH^\dagger).$$

This shows that $DOF_{M,N,t}(\mathcal{C}_2) = DOF_{M,N,t} = t \min(M, N)$. ■

In Examples 2 and 3, we see that DIV of a MIMO system is linear in the number of receive antennas. This makes sense intuitively because receive antennas provide natural redundancy in the system. By increasing the number of receive antennas we get many replicas of the transmitted signal and hence greater error protection.

It is quite intuitive to expect that DIV and DOF depend on each other. From the definitions, it is obvious that they are related to each other parametrically through $\mathcal{C}$. Indeed, in the three examples given above we see that when a communication system is operating at maximum diversity ($\sup_{\mathcal{C}} \mathrm{DIV}(\mathcal{C})$), the corresponding DOF is zero whereas when the system is operating at maximum degrees of freedom ($\sup_{\mathcal{C}} \mathrm{DOF}(\mathcal{C})$), the corresponding DIV is zero. The following example illustrates this point further.

Example 4: Consider the same system as in Example 3. A lower bound on the error exponent for this system can be calculated as in [73]. By choosing the input distribution $p(S)$, to be i.i.d. complex circular Gaussian $p_G(S)$ (capacity achieving distribution), the error exponent is:

$$E_0(\gamma, p_G(S), \rho) = -\log E\left[\det\left(I_M + \frac{1}{1+\gamma}\frac{\rho}{M}HH^\dagger\right)^{-\gamma t}\right]$$

where we have chosen to make the dependence of $E_0(\cdot)$ on ρ explicit. Given a rate R, the upper bound on the probability of error is given by

$$P_e \leq \exp\left(-\left[E_0(\gamma, p_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]\right)$$

where γ is chosen so that $R = \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}$. As $\rho \to \infty$, a fixed value of R, for all ρ, corresponds to different values of γ.

Now, instead of fixing R we fix γ. Then we see that as $\rho \to \infty$

$$R(\rho, M, N, t, \gamma) = \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}$$

gives a fixed DOF and varying γ varies DOF. Similarly, as $\rho \to \infty$

$$P_e(\rho, M, N, t, \gamma) \leq \exp\left(-\left[E_0(\gamma, p_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]\right).$$

gives a fixed DIV and varying γ varies DIV. This implies that each value of γ corresponds to a different communication system $\mathcal{C}_\gamma$. That is

$$R(\rho, M, N, t, \mathcal{C}_\gamma) = \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}.$$

and

$$P_e(\rho, M, N, t, \mathcal{C}_\gamma) = \exp\left(-\left[E_0(\gamma, P_G(S), \rho) - \gamma\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}\right]\right).$$

We can plot

$$\mathrm{DOF}_{M,N,t}(\mathcal{C}_\gamma) = \lim_{\rho\to\infty}\frac{\frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial\gamma}}{C_H(\rho, 1, 1, 1, \mathcal{C}_\gamma)}$$

versus

$$\mathrm{DIV}_{M,N,t}(\mathcal{C}_\gamma) = \lim_{\rho\to\infty} \frac{-\left[E_0(\gamma, p_G(S), \rho) - \gamma \frac{\partial E_0(\gamma, p_G(S), \rho)}{\partial \gamma}\right]}{\log P_{eH}(\rho, 1, 1, 1, \mathcal{C}_\gamma)}$$

parameterized by γ. One such plot for $t = T = 5$, $N = 1$ and $M = 3$ is shown in Figure 7.1.

Figure 7.1: DOF as a function of DIV

■

Example 5: Consider the system in Example 4. By evaluating $\frac{\partial E_0(\gamma)}{\partial \gamma}$ at $\gamma = 0$ we obtain $R(\rho, M, N, t, 0) = t \log\det(I_M + \frac{\rho}{M} HH^\dagger) = C(\rho, M, N, t)$. Therefore, $\gamma = 0$ corresponds to a system operating at $\mathrm{DOF}_{M,N,t}$.

If we fix R for all ρ (DOF $= 0$) then as $\rho \to \infty$ the upper bound on probability of error is simply

$$P_e \le \exp(-E_0(1, p_G(S)) - R)$$

where we have chosen the input distribution to be the capacity achieving one. From [73],

$$E_0(1) = -\log E\left[\det\left(I_M + \frac{\rho}{2M} HH^\dagger\right)^{-t}\right].$$

Therefore, we obtain for $\rho \to \infty$

$$\mathrm{DIV}_{M,N,t}(\mathcal{C}_1) = \min\big(\min(M,N)t,\ \min(M,t)N\big).$$

From Figure 7.1 we can indeed see that for $\gamma = 0$, $\mathrm{DOF}_{M,N,t}(\mathcal{C}_\gamma) = t\min(M,N)$ and $\mathrm{DIV}_{M,N,t}(\mathcal{C}_\gamma) = 0$ whereas for $\gamma = 1$, $\mathrm{DOF}_{M,N,t}(\mathcal{C}_\gamma) = 0$ and $\mathrm{DIV}_{M,N,t}(\mathcal{C}_\gamma) = \min\big(\min(M,N)t,\ \min(M,T)N\big)$.

DIV in the communication system corresponding to $\gamma = 1$ is lower than or equal to DIV in the communication system in Example 3 even though DOF in both systems is zero. That is so because the maximization over $p(S)$ in the error exponent (7.6) was performed in Example 3 unlike in the current example where we fixed $p(S) = p_G(S)$.

Example 6 [Rapidly fading channel, [71]]: Now, let's reconsider the MIMO system in Example 3 with the additional constraint that $T = 1$. We will investigate DIV and DOF for $t > 1$ channel uses. Let's again consider the two systems $\mathcal{C}_1$ and $\mathcal{C}_2$.

In $\mathcal{C}_1$, we use the signaling scheme developed by Tarokh et. al [71, pp. 750–751]. Using our definition for DIV and Tarokh's development for probability of error [71, (17), p. 751] we conclude that $\mathrm{DIV}_{M,N,t}(\mathcal{C}_1) = Nt$ which agrees with Tarokh's conclusion about diversity. And similar to Example 3, $\mathrm{DOF}_{M,N,t}(\mathcal{C}_1) = 0$.

In $\mathcal{C}_2$, we choose the elements of the matrix S to be i.i.d. complex circular Gaussian random variables with mean zero and variance one. This is the capacity achieving signal. Therefore [50],

$$R(\rho, M, N, t, \mathcal{C}_2) = C(\rho, M, N, t) = tE\log\det(I_M + \frac{\rho}{M}HH^\dagger).$$

This shows that $DOF_(M,N,t)(\mathcal{C}_2) = DOF_{M,N,t} = t\min(M,N)$.

In this example, we see that DIV $\geq$ DOF. ■

Example 7: Let's reconsider the case of AWGN channel in Example 1 operating at an optimal rate in a communication system with a binary input ($\{-1, 1\}$) and hard decision decoding at the receiver ($\mathcal{C}_1$). In this case, the channel effectively behaves like a binary symmetric channel with a crossover probability $p = Q(\sqrt{\rho}) \approx \frac{c}{\rho}\exp(-\rho/2)$ where c is some constant. We will calculate DOF for t channel uses for $\mathcal{C}_1$. We note that as $\rho \to \infty$, $p \to 0$ and the maximum achievable rate for this system is 1 bit per channel use. That is, $\lim_{\rho\to\infty} R(\rho, t, \mathcal{C}_1) = t$. Therefore, $\mathrm{DOF}_t(\mathcal{C}_1)$ is zero.

Now consider $\mathcal{C}_2$ which is similar to $\mathcal{C}_1$ except that the channel is no longer operating at the optimal rate and the communication system has repetition coding at the transmitter. We will calculate DIV for $\mathcal{C}_2$ corresponding to t, t odd, channel uses. Note that the channel is effectively a binary symmetric channel with crossover probability $p \approx \frac{c}{\rho}\exp(-\rho/2)$. The probability of error when using a repetition code of length t, t odd, is

$$P_e(t, p) = \begin{pmatrix} m \\ \frac{t+1}{2} \end{pmatrix} p^{\frac{t+1}{2}}(1-p)^{\frac{t-1}{2}}$$

Therefore, as $p \to 0$ $P_e(t, p) \approx c' p^{\frac{t+1}{2}}$ where c' is some other constant. Therefore,

$$\mathrm{DIV}_t(\mathcal{C}_2) = \frac{t+1}{2}$$

We see that hard decision decoding at the output reduces DIV to $(t+1)/2$ as opposed to DIV of t in Example 1 that has soft decision decoding at the output. ■

Intuitively, we would expect the DIV $\geq$ DOF for a channel as DOF in a channel can be used to transmit redundant information (repetition coding) thus adding to "natural" redundancy (multiple receive antennas) in the channel. This intuition however, breaks down with the case of multiple antenna channels operating in a coherent Rayleigh fading environment where we have seen the diversity is $\min(M, t)N$ [71] whereas the degrees of freedom is $\min(M, N)t$ [50, 85].

7.3 Conclusions

We have introduced a rigorous definition for diversity and degrees of freedom in a more general setting than MIMO communication system operating in a fading environment that we hope will dispel some of the confusion surrounding these two quantities. We have shown how these definitions agree with the current literature through various examples. We have also given an intuitive definition of these quantities where diversity should be regarded as the maximum amount of redundancy in the channel and degrees of freedom should be regarded as the number of independent channels available for communication.

APPENDICES

APPENDIX A

Appendices for Chapter 3

A.1 Derivation of Stability Condition (3.7)

We will follow the Z-transform method of [46]. Let $\tilde{\xi}(z)$ donate the Z-transform of ξ_k and $\tilde{G}_i(z)$ donate the Z-transform of the i^{th} component of G_k. Then we have the following

$$\begin{aligned}
\tilde{\xi}(z) &= \xi_{min}\frac{1}{1-z^{-1}} + \sum_{i=1}^{N}\tilde{G}_i(z) \\
\tilde{G}_i(z) &= (1-\frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)\tilde{G}_i(z) + \frac{\mu^2}{P}\lambda_i^2 z^{-1}\tilde{\xi}(z) + \\
& G_i(0)
\end{aligned}$$

which leads to

$$\tilde{\xi}(z) = \frac{\xi_{min}\frac{1}{1-z^{-1}} + \sum_{i=1}^{N}\frac{G_i(0)}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}}{1-\sum_{i=1}^{N}\frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}} \tag{A.1}$$

and

$$\tilde{G}_i(z) = \frac{1}{D(z)}\frac{\frac{\mu^2}{P}\lambda_i^2 N(z) + G_i(0)D(z)}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)} \tag{A.2}$$

where $N(z)$ and $D(z)$ denote the numerator and the denominator in (A.1). Therefore, the condition for stability is that the roots of

$$z-(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)=0$$

for $i = 1, \ldots, N$ and

$$\prod_{i=1}^{N} \left[z - (1 - \tfrac{2\mu}{P}\lambda_i + \tfrac{2\mu^2}{P}\lambda_i^2)\right] -$$
$$\sum_{i=1}^{N} \tfrac{\mu^2}{P}\lambda_i^2 \prod_{k\neq i} \left[z - (1 - \tfrac{2\mu}{P}\lambda_k + \tfrac{2\mu^2}{P}\lambda_k^2)\right] = 0$$

should lie within the unit circle.

The reader should note that (A.2) should be used to determine the stability of $\tilde{G}_i(z)$ and not

$$\tilde{G}_i(z) = \frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}\left[\xi_{min}\frac{1}{1-z^{-1}} + \sum_{k\neq i}\tilde{G}_k(z)\right] + G_i(0)}{1 - z^{-1}(1 - \frac{2\mu}{P}\lambda_i + \frac{3\mu^2}{P}\lambda_i^2)}$$

that was used in [46].

Following the rest of the procedure as outlined in [46] exactly, we obtain the conditions for stability to be (3.7).

A.2 Derivation of expression (3.9)

Here we follow the procedure in [22]. Assuming G_k converges we have the expression for G_∞ to be

$$G_\infty = P\ [2\mu\Lambda - 2\mu^2\Lambda - \mu^2\Lambda^2\mathbf{1}\mathbf{1}^\tau]^{-1}\frac{\mu^2}{P}\Lambda^2\mathbf{1}\xi_{min}.$$

Then we have

$$G_{k+1} - G_\infty = F(G_k - G_\infty)$$

where $F = I - \frac{2\mu}{P}\Lambda + \frac{2\mu^2}{P}\Lambda + \frac{\mu^2}{P}\Lambda^2\mathbf{1}\mathbf{1}^\tau$. Since $\xi_k = \text{tr}\{G_k\}$ we have

$$\begin{aligned}\sum_{k=0}^{\infty}(\xi_k - \xi_\infty) &= \text{tr}\{\sum_{k=0}^{\infty}(G_k - G_\infty)\} \\ &= \text{tr}\{\sum_{k=0}^{\infty} F^k(G_0 - G_\infty)\} \\ &= \text{tr}\{(I - F)^{-1}(G_0 - G_\infty)\}\end{aligned}$$

from which (3.9) follows.

A.3 Derivation of the misadjustment factor (3.8)

Here we follow the approach of [46]. The misadjustment numerator and denominator is defined as $M(\mu) = \frac{\xi_\infty - \xi_{min}}{\xi_{min}}$. Since $\xi_\infty = \lim_{z\to 1}(1-z^{-1})\tilde{\xi}(z)$ and the limits of $(1-z^{-1})\xi(z)$ are finite, we have

$$\xi_\infty = \frac{\lim_{z\to 1}\left[\xi_{min} + (1-z^{-1})\sum_{i=1}^{N} \frac{G_i(0)}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)}\right]}{\lim_{z\to 1}\left[1 - \sum_{i=1}^{N} \frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)}\right]},$$

that is

$$\begin{aligned} \xi_\infty &= \frac{\xi_{min}}{1 - \frac{1}{2}\sum_{i=1}^{N}\frac{\mu\lambda_i}{1-\mu\lambda_i}} \\ &= \frac{\xi_{min}}{1-\eta(\mu)}, \end{aligned}$$

from which (3.8) follows.

A.4 Proofs of Lemma 3.1 and Theorem 3.1

Proof of Lemma 3.1: First note that $e_k = -V_k^\dagger X_k$. Next, consider the Lyapunov function $\mathcal{L}_{k+1} = \overline{V_{k+1}^\dagger V_{k+1}}$ where $\overline{\{\cdot\}}$ is as defined in Lemma 3.1. Averaging the following update equation for $V_{k+1}^\dagger V_{k+1}$

$$V_{k+1}^\dagger V_{k+1} = V_k^\dagger V_k - \mu \mathrm{tr}\{V_k V_k^\dagger X_k X_k^\dagger I_i\} - \mu \mathrm{tr}\{V_k V_k^\dagger I_i X_k X_k^\dagger\} + \mu^2 \mathrm{tr}\{V_k V_k^\dagger X_k X_k^\dagger I_i X_k X_k^\dagger\}$$

over all possible choices of S_i, $i = 1, \ldots, P$ we obtain

$$\mathcal{L}_{k+1} = \mathcal{L}_k - \frac{\mu}{P}\mathrm{tr}\{\overline{V_k V_k^\dagger} X_k(2 - \mu X_k X_k^\dagger)X_k^\dagger\}.$$

Since $\sup_k(X_k^\dagger X_k) \leq B < \infty$ the matrix $(2I - \mu X_k X_k^\dagger) - (2I - \mu B I)$ is positive definite. Therefore,

$$\mathcal{L}_{k+1} \leq \mathcal{L}_k - \frac{\mu}{P}(2-\mu B)\mathrm{tr}\{\overline{V_k V_k^\dagger} X_k X_k^\dagger\}.$$

Since $\mu < 2/B$

$$\mathcal{L}_{k+1} \leq \mathcal{L}_k - \mathrm{tr}\{\overline{V_k V_k^\dagger} X_k X_k^\dagger\}$$

Noting that $\overline{e_k^2} = \mathrm{tr}\{\overline{V_k V_k^\dagger} X_k X_k^\dagger\}$ we obtain

$$\mathcal{L}_{k+1} + \sum_{l=0}^{k} \overline{e_k^2} \leq \mathcal{L}_0$$

since $\mathcal{L}_0 < \infty$ we have $\overline{e_k^2} = O(1/k)$ and $\lim_{k\to\infty} \overline{e_k^2} = 0$ ■

Before proving Theorem 3.1 we need Lemmas A.1 and A.2. We reproduce the proof of Lemma A.1 from [63] using our notation because this enables to understand the proof of Lemma A.2 better.

Lemma A.1. *[63, Lemma 6.1 p. 143-144] Let X_k satisfy the persistence of excitation condition in Theorem 3.1. let*

$$\Pi_{k,k+D} = \begin{cases} \prod_{l=k}^{k+D}(I - \frac{\mu}{P} X_l X_l^\dagger) & \textit{if } D \geq 0 \\ 1 & \textit{if } D < 0 \end{cases}$$

and

$$\mathcal{G}_k = \sum_{l=0}^{K} \Pi_{k,k+l-1}^\dagger X_{k+l} X_{k+l}^\dagger \Pi_{k,k+l-1}$$

where K is as defined in Theorem 3.1 then $\mathcal{G}_k - \eta I$ is a positive definite matrix for some $\eta > 0$ and $\forall k$.

Proof: Proof is by contradiction. Suppose not then for some vector ω such that $\omega^\dagger \omega = 1$ we have $\omega^\dagger \mathcal{G}_k \omega \leq c^2$ where c is any arbitrary positive number.

Then

$$\begin{aligned} & \sum_{l=0}^{K} \omega^\dagger \Pi_{k,k+l-1}^\dagger X_{k+l} X_{k+l}^\dagger \Pi_{k,k+l-1} \omega \leq c^2 \\ \Rightarrow \quad & \omega^\dagger \Pi_{k,k+l-1}^\dagger X_{k+l} X_{k+l}^\dagger \Pi_{k,k+l-1} \omega \leq c^2 \quad \text{for } 0 \leq l \leq K. \end{aligned}$$

Choosing $l = 0$ we obtain $\omega^\dagger X_k X_k^\dagger \omega \leq c^2$ or $\|\omega^\dagger X_k\| \leq c$.

Choosing $l = 1$ we obtain $\|\omega^\dagger(I - \frac{\mu}{P}X_kX_k^\dagger)X_{k+1}\| \leq c$. Therefore,

$$\begin{aligned}\|\omega^\dagger X_{k+1}\| &\leq \|\omega^\dagger(I - \frac{\mu}{P}X_kX_k^\dagger)X_{k+1}\| + \frac{\mu}{P}\|\omega^\dagger X_k\|\|X_k^\dagger X_{k+1}\| \\ &\leq c + \frac{\mu}{P}Bc = c(1 + 2/P).\end{aligned}$$

Choosing $l = 2$ we obtain $\|\omega^\dagger(I - \frac{\mu}{P}X_kX_k^\dagger)(I - \frac{\mu}{P}X_{k+1}X_{k+1}^\dagger)X_{k+2}\| \leq c$. Therefore,

$$\begin{aligned}\|\omega^\dagger X_{k+2}\| &\leq \|\omega^\dagger(I - \frac{\mu}{P}X_kX_k^\dagger)(I - \frac{\mu}{P}X_{k+1}X_{k+1}^\dagger)X_{k+2}\| + \frac{\mu}{P}\|\omega^\dagger X_kX_k^\dagger X_{k+2}\| \\ &\quad + \frac{\mu}{P}\|\omega^\dagger X_{k+1}X_{k+1}^\dagger X_{k+2}\| + \frac{\mu^2}{P^2}\|\omega^\dagger X_kX_k^\dagger X_{k+1}X_{k+1}^\dagger X_{k+2}\| \\ &\leq O(c).\end{aligned}$$

Proceeding along similar lines we obtain $\|\omega^\dagger X_{k+l}\| \leq Lc$ for $l = 0, \ldots, K$ where L is some constant. This implies $\omega^\dagger \sum_{l=k}^{k+K} X_lX_l^\dagger\omega \leq (K + 1)L^2c^2$. Since c is arbitrary we obtain that $\omega^\dagger \sum_{l=k}^{k+K} X_lX_l^\dagger\omega < \alpha_1$ which is a contradiction. ■

Lemma A.2. *Let X_k satisfy the persistence of excitation condition in Theorem 3.1. let*

$$\mathcal{P}_{k,k+D} = \begin{cases} \prod_{l=k}^{k+D}(I - \mu I_lX_lX_l^\dagger) & \text{if } D \geq 0 \\ 1 & \text{if } D < 0 \end{cases}$$

where I_l is the randomly chosen masking matrix and let

$$\Omega_k = \sum_{l=0}^{K} \overline{\Pi_{k,k+l-1}^\dagger X_{k+l}X_{k+l}^\dagger \Pi_{k,k+l-1}}$$

where K is as defined in Theorem 3.1 and $\overline{\{\cdot\}}$ is the average over randomly chosen I_l then $\Omega_k - \gamma I$ is a positive definite matrix for some $\gamma > 0$ and $\forall k$.

Proof: Proof is by contradiction. Suppose not then for some vector ω such that $\omega^\dagger\omega = 1$ we have $\omega^\dagger\Omega_k\omega \leq c^2$ where c is any arbitrary positive number.

Then

$$\begin{aligned}&\sum_{l=0}^{K} \omega^\dagger\, \overline{\mathcal{P}_{k,k+l-1}^\dagger X_{k+l}X_{k+l}^\dagger \mathcal{P}_{k,k+l-1}}\, \omega \leq c^2 \\ \Rightarrow\ &\omega^\dagger\, \overline{\mathcal{P}_{k,k+l-1}^\dagger X_{k+l}X_{k+l}^\dagger \mathcal{P}_{k,k+l-1}}\, \omega \leq c^2 \quad \text{for } 0 \leq l \leq K.\end{aligned}$$

Choosing $l = 0$ we obtain $\omega^\dagger X_k X_k^\dagger \omega \leq c^2$ or $\|\omega^\dagger X_k\| \leq c$.

Choosing $l = 1$ we obtain $\omega^\dagger \ \overline{(I - \mu X_k X_k^\dagger I_k) X_{k+1} X_{k+1}^\dagger (I - \mu I_k X_k X_k^\dagger)} \ \omega \leq c^2$. Therefore,

$$\begin{aligned}\omega^\dagger X_{k+1} X_{k+1}^\dagger \omega - \frac{\mu}{P}\omega^\dagger X_k X_k^\dagger X_{k+1} X_{k+1}^\dagger \omega - \frac{\mu}{P}\omega^\dagger X_{k+1} X_{k+1}^\dagger X_k X_k^\dagger \omega + \\ \frac{\mu^2}{P}\omega^\dagger X_k X_k^\dagger \left[\sum_{i=0}^{P} I_i X_{k+1} X_{k+1}^\dagger I_i\right] X_k X_k^\dagger \omega \ \leq \ c^2.\end{aligned}$$

Now

$$\begin{aligned}\|\omega^\dagger X_k X_k^\dagger X_{k+1} X_{k+1}^\dagger \omega\| &\leq \|\omega^\dagger X_k\| \|X_k\| \|X_{k+1}^\dagger X_{k+1}\| \|\omega\| \\ &\leq cB^{3/2}\end{aligned}$$

and

$$\|\omega^\dagger X_k X_k^\dagger \left[\sum_{i=0}^{P} I_i X_{k+1} X_{k+1}^\dagger I_i\right] X_k X_k^\dagger \omega\| \leq c^2 P B^2.$$

Therefore, $\omega^\dagger X_{k+1} X_{k+1}^\dagger \omega = O(c)$ which implies $\|\omega^\dagger X_{k+1}\| = O(c^{1/2})$. Proceeding along the same lines we obtain $\|\omega^\dagger X_{k+1}\| = O(c^{1/L})$ for $l = 0, \ldots, K$ for some constant L. This implies $\omega^\dagger \sum_{l=k}^{k+K} X_l X_l^\dagger \omega = O(c^{2/L})$. Since c is arbitrary we obtain that $\omega^\dagger \sum_{l=k}^{k+K} X_l X_l^\dagger \omega < \alpha_1$ which is a contradiction. ∎

Now, we are ready to Prove Theorem 3.1.

Proof of Theorem 3.1: First, we will prove the convergence of $\overline{V}_k^\dagger \overline{V}_k$. We have $\overline{V}_{k+1} = (I - \frac{\mu}{P} X_k X_k^\dagger)\overline{V}_k$. Proceeding as before, we obtain the following update equation for $\overline{V}_k \overline{V}_k^\dagger$

$$\begin{aligned}\overline{V}_{k+K+1}^\dagger \overline{V}_{k+K+1} &= \overline{V}_{k+K}^\dagger \overline{V}_{k+K} - 2\frac{\mu}{P}\overline{V}_{k+K}^\dagger X_{k+K} X_{k+K}^\dagger \overline{V}_{k+K} \\ &\quad + \frac{\mu^2}{P^2}\overline{V}_{k+K}^\dagger X_{k+K} X_{k+K}^\dagger X_{k+K} X_{k+K}^\dagger \overline{V}_{k+K} \\ &\leq \overline{V}_{k+K}^\dagger \overline{V}_{k+K} - \frac{\mu}{P}\overline{V}_{k+K}^\dagger X_{k+K} X_{k+K}^\dagger \overline{V}_{k+K}.\end{aligned}$$

The last step follows from the fact that $\mu < 2/B$. Using the update equation for $\overline{V_k}$ repeatedly, we obtain

$$\overline{V}^{\dagger}_{k+K+1}\overline{V}_{k+K+1} \;\; \leq \;\; \overline{V}^{\dagger}_{k}\overline{V}_{k} - \frac{\mu}{P}\overline{V}^{\dagger}_{k}\mathcal{G}_k\overline{V}_{k}.$$

From Lemma A.1 we have,

$$\overline{V}^{\dagger}_{k+K+1}\overline{V}_{k+K+1} \leq (1-\frac{\mu}{P}\eta)\overline{V}^{\dagger}_{k}\overline{V}_{k}$$

which ensures exponential convergence of $\mathrm{tr}\{\overline{V}_k\overline{V}^{\dagger}_k\}$.

Next, we prove the convergence of $\overline{V_k^{\dagger}V_k}$. First, we have the following update equation for $\mathrm{tr}\{\overline{V_kV_k^{\dagger}}\}$

$$\mathrm{tr}\{\overline{V_{k+K+1}V_{k+K+1}^{\dagger}}\} \leq \mathrm{tr}\{\overline{V_{k+K}V_{k+K}^{\dagger}}\} - \frac{\mu}{P}\mathrm{tr}\{X_{k+K}X_{k+K}^{\dagger}\overline{V_{k+K}V_{k+K}^{\dagger}}\}. \tag{A.3}$$

Using (A.3) and also

$$\overline{V_{k+1}V_{k+1}^{\dagger}} = \overline{(I-\mu I_kX_kX_k^{\dagger})\overline{V_kV_k^{\dagger}}(I-\mu X_kX_k^{\dagger}I_k)}$$

repeatedly, we obtain the following update equation

$$\mathrm{tr}\{\overline{V_{k+K+1}V_{k+K+1}^{\dagger}}\} \leq \mathrm{tr}\{\overline{V_kV_k^{\dagger}}\} - \mathrm{tr}\{\Omega_k\overline{V_kV_k^{\dagger}}\}.$$

From Lemma A.2 we have

$$\mathrm{tr}\{\overline{V_{k+K+1}V_{k+K+1}^{\dagger}}\} \leq (1-\mu\gamma)\mathrm{tr}\{\overline{V_kV_k^{\dagger}}\}$$

which ensures the exponential convergence of $\mathrm{tr}\{\overline{V_kV_k^{\dagger}}\}$. ■

A.5 Proof of Theorem 3.2 in Section 3.5.1

For the proof, we need some definitions first. We define for $p \geq 1$ the set $\mathcal{M}_p$ of $F = \{F_i\}$ as

$$\mathcal{M}_p = \left\{F : \sup_i \|S_i^{(T)}\|_p = o(T), \text{ as } T \to \infty\right\} \tag{A.4}$$

where $S_i^{(T)} = \sum_{j=iT}^{(i+1)T-1}(F_j - E[F_j])$.

The proof is just a slightly modified version of the proof of Theorem 2 derived in Section IV of [37, pp. 766-769]. The modification takes into account that in the present context F_k is no longer $F_k = X_k X_k^\dagger$ but, $F_k = I_k X_k X_k^\dagger$. Theorem 3.2 is proved in a step by step manner using different lemmas. First, we rewrite

$$X_k = \sum_{j=-\infty}^{\infty} a_j \epsilon(k,j) + \xi_k, \quad \sum_{j=-\infty}^{\infty} a_j < \infty$$

where by definition

$$a_j \stackrel{\text{def}}{=} \sup_k \|A(k,j)\|, \ \ \epsilon(k,j) = a_j^{-1} A(k,j)\epsilon_{k-j}. \tag{A.5}$$

The new process has some simple properties as listed in [37].

Lemma A.3. *if $\{G_k\}$ is a ϕ-mixing $d\times d$-dimensional matrix then so is $\{F_k = I_k G_k\}$.*

Lemma A.4. *Let $\{F_k\}$ be a ϕ-mixing $d \times d$-dimensional matrix process with mixing rate $\{\phi(m)$. Then*

$$\sup_i \|S_i^{(T)}\|_2 \le 2cd \left\{ T \sum_{m=0}^{T-1} \sqrt{\phi(m)} \right\}^{1/2}, \ \forall T \ge 1$$

where $S_i^{(T)}$ is as defined earlier and c is defined by $c \stackrel{\text{def}}{=} \sup_i \|F_i - EF_i\|_2$.

Proof: Proof is the same as the proof of Lemma 1 in [37]. ∎

Lemma A.5. *Let $F_k = I_k X_k X_k^\dagger$, where $\{X_k\}$ is defined by (3.12) with $\sup_k \|\epsilon_k\|_4 < \infty$. Then $\{F_k\} \in \mathcal{M}_2$, where $\mathcal{M}_2$ is defined by (A.4).*

Proof: Proof is practically the same as the proof for Lemma 2 in [37]. All we need to add is that if $\{G_k\}$ is ϕ-mixing then so is $\{I_k G_k\}$. ∎

Lemma A.6. *Let $\sup_k E\|X_k\|_2 < \infty$. Then $\{I_k X_k X_k^\dagger\} \in \mathcal{S}$ if and only if (3.14) holds, where $\mathcal{S}$ is defined by (3.11).*

Proof: Let us first assume that (3.14) is true. Take $\mu^* = (1 + \sup_k E\|X_k\|^2)^{-1}$. Then applying Theorem 2.1 in [36] to the deterministic sequence $A_k = \mu E[I_k X_k X_k^\dagger]$ for any $\mu \in (0, \mu^*]$, it is easy to see that $\{I_k X_k X_k^\dagger\} \in \mathcal{S}(\mu^*)$.

Conversely, if $\{X_k X_k^\dagger\} \in \mathcal{S}$, then there exists $\mu^* \in (0, (1 + \sup_k E\|X_k\|^2)^{-1}]$ such that $\{X_k X_k^\dagger\} \in \mathcal{S}(\mu^*)$. Now, applying Theorem 2.2 in [36] to the deterministic sequence $A_k = \mu^* E[I_k X_k X_k^\dagger]$, it is easy to see that (3.14) holds. This completes the proof. ∎

Lemma A.7. *Let $F_k = I_k X_k X_k^\dagger$, where $\{X_k\}$ is defined by (3.12) with (3.13) satisfied. Then $\{F_k\}$ satisfies condition 1) of Theorem 1.*

Proof: From Lemma 4 in [37] we know that $G_k = X_k X_k^\dagger$ satisfies condition 1) of Theorem 1. Since $\|F_k\| \leq \|G_k\|$ it follows that $\{F_k\}$ satisfies condition 1) of Theorem 1. ∎

Lemma A.8 (Lemma 5 in [37]). *Let $\{z_k\}$ be a nonnegative random sequence such that for some $a > 0, b > 0$ and for all $i_1 < i_2 < \ldots < i_n$, $\forall n \geq 1$*

$$E \exp\left\{\sum_{k=1}^{n} z_{i_k}\right\} \leq \exp\{an + b\}. \tag{A.6}$$

Then for any $L > 0$ and any $n \geq i \geq 0$

$$E \exp\left\{\frac{1}{2}\sum_{j=i+1}^{n} z_j I(z_j \geq L)\right\} \leq \exp\{e^{a-L/2}(n-i) + b\}$$

where $I(\cdot)$ is the indicator function.

Proof: This lemma has been proved in [37]. ∎

Lemma A.9. *Let $F_k = I_k X_k X_k^\dagger$ where $\{X_k\}$ is defined by (3.12) with (3.13) satisfied. Then $\{F_k\}$ satisfies condition 2) of Theorem 1.*

Proof: Set for any fixed k and l

$$z_j \stackrel{\text{def}}{=} z_j(k,l) = \left\| \sum_{t=jT}^{(j+1)T-1} [I_t\epsilon(t,k)\epsilon(t,l)^\dagger - EI_t\epsilon(t,k)\epsilon(t,l)^\dagger] \right\|,$$

where $\epsilon(k,l)$ is as defined in (A.5). Then, we have

$$\sum_{j=i+1}^{n} \|S_j^{(T)}\| \leq \sum_{k,l=-\infty}^{\infty} a_k a_l \sum_{j=i+1}^{n} z_j + 2\sum_{k=-\infty}^{\infty} a_k \sum_{j=i+1}^{n} \left\| \sum_{t=jT}^{(j+1)T-1} I_t\epsilon(t,k)\xi_t^\dagger \right\|. \quad \text{(A.7)}$$

We first consider the second to the last term in the previous equation. By the Hőlder inequality

$$E\exp\left\{\mu \sum_{k,l=-\infty}^{\infty} a_k a_l \sum_{j=i+1}^{n} z_j\right\} \leq \prod_{k,l=-\infty}^{\infty} \left\{E\exp\left\{\mu A^2 \sum_{j=i+1}^{n} z_j\right\}\right\}^{\frac{a_k a_l}{A^2}}$$

where $A \stackrel{\text{def}}{=} \sum_{j=-\infty}^{\infty} a_j$.

Now, let $c = \sum_k E\|\epsilon_k\|^2$, and note that

$$\begin{aligned}
\|I_t\epsilon(t,k)\epsilon(t,l)^\dagger\| &\leq \|\epsilon(t,k)\epsilon(t,l)^\dagger\| \\
&\leq \frac{1}{2}[\|\epsilon(t,k)\|^2 + \|\epsilon(t,l)\|^2] \\
&\leq \frac{1}{2}(\|\epsilon_{t-k}\|^2 + \|\epsilon_{t-l}\|^2)
\end{aligned}$$

and we have

$$z_j \leq \frac{1}{2} \sum_{t=jT}^{(j+1)T-1} (\|\epsilon_{t-k}\|^2 + \|\epsilon_{t-l}\|^2) + \frac{c}{m}T.$$

By this and (3.13) it is easy to prove that the sequence $\{\alpha z_j\}$ satisfies (A.6) with $a = (K + c/m)T$ and $b = \log M$, where α is defined as in (3.13). Consequently, by Lemma 5 in [37] we have for any $L > 0$

$$E\exp\left\{\frac{\alpha}{2} \sum_{j=i+1}^{n} z_j I(z_j \geq LT)\right\} \leq M\exp\left\{e^{(K+c-\frac{\alpha L}{2})T}(n-i)\right\}.$$

Now, in view of the above, taking $\mu < \frac{\alpha A^{-2}}{4}$ and $L > 2\alpha^{-1}(K+c)$, and again applying the Hőlder inequality, we have

$$E\exp\left\{2\mu A^2\sum_{j=i+1}^{n} z_j I(z_j \geq LT)\right\} \leq M\exp\{\mu\delta(T)(n-i)\}$$

where $\delta(T) \to 0$ as $T \to \infty$, where

$$\delta(T) = 4\alpha^{-1}A^2\exp\left\{(K+c-\frac{\alpha L}{2})T\right\}.$$

Next, we consider the term $x_j \stackrel{\text{def}}{=} z_j I(z_j \leq LT)$.

By the inequality $e^x \leq 1+2x$, $0 \leq x \leq \log 2$, we have for small $\mu > 0$

$$\exp\left\{2\mu A^2\sum_{j=i+1}^{n} x_j\right\} \leq \prod_{j=i+1}^{n}(1+4\mu A^2 x_j)$$

$\mu < \min(\log 2/2A^2 x_j)$, $x_j > 0$.

As noted before, for any fixed k and l, the process $\{\epsilon(t,k)\epsilon(t,l)^{\dagger}\}$ is ϕ-mixing with mixing rate $\phi(m-|k-l|)$. Consequently, for any fixed k and l, both $\{z_j\}$ and $\{x_j\}$ are also ϕ-mixing with mixing rate $\phi((m-1)T+1-|k-l|)$. Note also that by Lemma 1 in [37]

$Ex_j \leq Ez_j \leq \|z_j\|_2 \leq f_{kl}(T)$ where $f_{kl}(T) = 2cd\left\{T\sum_{m=0}^{T-1}\sqrt{\phi(m-|k-j|)}\right\}^{\frac{1}{2}}$.

Therefore, applying Lemma 6.2 in [37], we have

$$\begin{aligned} E\prod_{j=i+1}^{n}(1+4\mu A^2 x_j) &\leq 2\{1+8\mu A^2[f_{kl}(T)+2LT\phi(T+1-|k-l|)]\}^{n-i} \\ &\leq 2\exp\{8\mu A^2[f_{kl}(T)+2LT\phi(T+1-|k-l|)](n-i)\}. \end{aligned}$$

Finally, using the Schwartz inequality we get

$$\begin{aligned} E\exp\left\{\mu A^2\sum_{j=i+1}^{n} z_j\right\} &\leq \left\{E\exp\left\{2\mu A^2\sum_{j=i+1}^{n} z_j I(z_j \geq LT)\right\}\right\}^{1/2} \times \\ &\quad \left\{E\exp\left\{2\mu A^2\sum_{j=i+1}^{n} x_j\right\}\right\}^{1/2} \\ &\leq \sqrt{2M}\exp\{\mu[\delta(T)+8A^2 f_{kl}(T)+ \\ &\quad 16LTA^2\phi(T+1-|k-l|)](n-i)\}. \end{aligned}$$

Therefore, it is not difficult to see that there exists a function $g(T) = o(T)$ such that for all small $\mu > 0$

$$E \exp\left\{\mu \sum_{k,l=-\infty}^{\infty} a_k a_l \sum_{j=i+1}^{n} z_j\right\} \leq \sqrt{2M} \exp\{\mu g(T)(n-i)\}.$$

We can similarly bound the second term in (A.7) and we are done. ■

Now we will highlight some of the remarks and corollaries to Theorem 2 in [37]. The remarks and corollaries are pertinent to this Theorem too.

Remark A.1. *By taking $A(k,0) = I$, $A(k,j) = 0$, $\forall k, \forall j \neq 0$, and $\xi_k = 0$, $\forall k$ in (3.12), we see that $\{X_k\}$ is the same as ϵ_k, which means that Theorem 3.2 is applicable to any ϕ-mixing sequences.*

Corollary A.1. *Let the signal process be generated by (3.12), where $\{\xi_k\}$ is a bounded deterministic sequence, and $\{\epsilon_k\}$ is an independent sequence satisfying*

$$\sup_k E[\exp(\alpha \|\epsilon_k\|^2)] < \infty, \text{ for some } \alpha > 0.$$

Then $\{X_k X_k^\dagger\} \in \mathcal{S}_p$ for some $p \geq 1$ if and only if there exists an integer $h > 0$ and a constant $\delta > 0$ such that (3.14) holds.

Proof: Similar to the proof of Corollary 1 in [37]. ■

A.6 Derivation of Expressions in Section 3.8.1

In this section, we will need the following identity

$$\sum_{s=0}^{\infty} s(1-\alpha\mu)^{2s} = \frac{(1-\alpha\mu)^2}{\alpha^2\mu^2(2-\alpha\mu)^2}.$$

First, we have the following expressions for LMS

$$\begin{aligned}
J^0_{k+1} &= \sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s}X_s n_s \\
J^1_{k+1} &= \mu\sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s-1}D_1(k,s+1)X_s n_s \\
J^2_{k+1} &= \mu^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s-2}D_2(k,s+1)X_s n_s
\end{aligned}$$

where

$$\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s)=0 \quad s>k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1)Z_u
\end{aligned}$$

and $Z_u = E[X_u X_u^\dagger] - X_u X_u^\dagger$.

This leads to

$$\lim_{k\to\infty} E[J^0_{k+1}(J^0_{k+1})^\dagger] = \lim_{k\to\infty}\sigma_v^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{2(k-s)}E[X_0X_0^\dagger]$$

and we finally obtain

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] = \frac{\sigma_v^2}{\mu(2-\mu\sigma^2)}I.$$

Similarly,

$$\lim_{k\to\infty} E[J^0_{k+1}(J^1_{k+1})^\dagger] = \mu\sigma_v^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{2s-1}E[D_1(s,1)X_0X_0^\dagger].$$

Now, $E[Z_uX_0X_0^\dagger] = E[X_uX_u^\dagger]E[X_0X_0^\dagger] - E[X_uX_u^\dagger X_0X_0^\dagger] = 0$ which gives

$$E[D_1(s,1)X_0X_0^\dagger] = 0.$$

Thus, $\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] = 0$.

Next,

$$\lim_{k\to\infty} E[J^1_{k+1}(J^1_{k+1})^\dagger] = \sigma_v^2\mu^2\sum_{u=0}^{\infty}(1-\mu\sigma^2)^{2u-2}E[D_1(u,1)X_0X_0^\dagger D_1(u,1)^\dagger]$$

$$
\begin{aligned}
E[Z_v X_0 X_0^\dagger Z_u^\dagger] &= \sigma^6 I - \sigma^2 E[X_v X_v^\dagger X_0 X_0^\dagger] - \sigma^2 E[X_0 X_0^\dagger X_u X_u^\dagger] + \\
&\quad E[X_v X_v^\dagger X_0 X_0^\dagger X_u X_u^\dagger] \\
&= 0 \ \text{ if } v \neq u \\
&= N\sigma^6 I \ \text{ if } v = u.
\end{aligned}
$$

Therefore, $E[D_1(u,1)X_0X_0^\dagger D_1(u,1)^\dagger] = uN\sigma^6 I$ and we obtain

$$
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] = \frac{N\sigma^2\sigma_v^2}{(2-\mu\sigma^2)^2} I.
$$

Next, we have $\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^2)^\dagger] = \mu^2\sigma_v^2 \sum_{s=0}^{\infty} E[D_2(s,1)X_0X_0^\dagger]$. Now,

$$
\begin{aligned}
E[Z_v Z_u X_0 X_0^\dagger] &= \sigma^6 I - \sigma^2 E[X_v X_v^\dagger X_0 X_0^\dagger] - \sigma^2 E[X_u X_u^\dagger X_0 X_0^\dagger] + \\
&\quad E[X_v X_v^\dagger X_u X_u^\dagger X_0 X_0^\dagger] \\
&= 0 \ \text{ if } v \neq u.
\end{aligned}
$$

Therefore, $E[D_2(s,1)X_0X_0^\dagger] = 0$ and consequently $\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^2)^\dagger] = 0$.

Second, we have the following expressions for SPULMS

$$
\begin{aligned}
J_{k+1}^0 &= \sum_{s=0}^{k} (1 - \frac{\mu}{P}\sigma^2)^{k-s} I_s X_s n_s \\
J_{k+1}^1 &= \mu \sum_{s=0}^{k} (1 - \frac{\mu}{P}\sigma^2)^{k-s-1} D_1(k, s+1) I_s X_s n_s \\
J_{k+1}^2 &= \mu^2 \sum_{s=0}^{k} (1 - \frac{\mu}{P}\sigma^2)^{k-s-2} D_2(k, s+1) I_s X_s n_s
\end{aligned}
$$

where

$$
\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k, u+1) Z_u
\end{aligned}
$$

and $Z_u = I_u X_u X_u^\dagger - \frac{1}{P} E[X_u X_u^\dagger]$.

This leads to

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^0)^\dagger] = \lim_{k\to\infty} \sigma_v^2 \sum_{s=0}^{k} (1-\mu\sigma^2)^{2(k-s)} E[I_0 X_0 X_0^\dagger I_0]$$

and

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] = \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P}\sigma^2)} I.$$

Similarly,

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^\dagger] = \mu\sigma_v^2 \sum_{s=0}^{k} (1-\mu\sigma^2)^{2s-1} E[D_1(s,1) I_0 X_0 X_0^\dagger I_0].$$

Now, $E[Z_u I_0 X_0 X_0^\dagger I_0] = E[I_u X_u X_u^\dagger] E[I_0 X_0 X_0^\dagger I_0] - E[I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] = 0$ which gives

$$E[D_1(s,1) X_0 X_0^\dagger] = 0$$

so that $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^\dagger] = 0$.

$$\lim_{k\to\infty} E[J_{k+1}^1(J_{k+1}^1)^\dagger] = \sigma_v^2\mu^2 \sum_{u=0}^{\infty} (1-\mu\sigma^2)^{2u-2} E[D_1(u,1) I_0 X_0 X_0^\dagger I_0 D_1(u,1)^\dagger].$$

Furthermore

$$\begin{aligned}
E[Z_v I_0 X_0 X_0 I_0 Z_u^\dagger] &= \sigma^6 I - \sigma^2 E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0] - \sigma^2 E[I_0 X_0 X_0^\dagger I_0 X_u X_u^\dagger I_u] \\
&\quad + E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0 X_u X_u^\dagger I_u] \\
&= 0 \ \text{ if } v \neq u \\
&= \frac{(N+1)P-1}{P^3} \sigma^6 I \ \text{ if } v = u.
\end{aligned}$$

Therefore, $E[D_1(u,1) X_0 X_0^\dagger D_1(u,1)^\dagger] = u\frac{(N+1)P-1}{P^3}\sigma^6 I$ and we obtain

$$\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] = \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{(2-\frac{\mu}{P}\sigma^2)^2} I.$$

Finally, we consider $\lim_{k\to\infty} E[J^0_{k+1}(J^2_{k+1})^\dagger] = \mu^2\sigma_v^2 \sum_{s=0}^\infty E[D_2(s,1)X_0X_0^\dagger]$. Now

$$\begin{aligned} E[Z_v Z_u X_0 X_0^\dagger] &= \sigma^6 I - \sigma^2 E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0] - \sigma^2 E[I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] \\ &\quad + E[I_v X_v X_v^\dagger I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] \\ &= 0 \ \text{ if } v \neq u. \end{aligned}$$

Therefore, $E[D_2(s,1)X_0X_0^\dagger] = 0$ and consequently $\lim_{k\to\infty} E[J^0_{k+1}(J^2_{k+1})^\dagger = 0$.

A.7 Derivation of Expressions in Section 3.3

In this section, we will need the following identities

$$\begin{aligned} \sum_{v,w=1}^{s} a^{2|v-w|} &= \frac{s(1-a^4) - 2a^2 + 2a^{2(s+1)}}{(1-a^2)^2} \\ \sum_{v,w=1}^{s} a^{|v-w|} a^{v+w} &= \frac{a^2}{(1-a^2)^2}[1 + a^2 - (2s+1)a^{2s} + (2s-1)a^{2s+2}] \\ \sum_{s=0}^{\infty} s(1-\alpha\mu)^{2s} &= \frac{(1-\alpha\mu)^2}{\alpha^2\mu^2(2-\alpha\mu)^2}. \end{aligned}$$

First, we have the following expressions for LMS

$$\begin{aligned} J^0_{k+1} &= \sum_{s=0}^{k}(1-\mu)^{k-s} X_s n_s \\ J^1_{k+1} &= \mu\sum_{s=0}^{k}(1-\mu)^{k-s-1} D_1(k,s+1) X_s n_s \\ J^2_{k+1} &= \mu^2\sum_{s=0}^{k}(1-\mu)^{k-s-2} D_2(k,s+1) X_s n_s \end{aligned}$$

where

$$\begin{aligned} D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\ D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1) Z_u \end{aligned}$$

and $Z_u = E[X_u X_u^\dagger] - X_u X_u^\dagger$.

This leads to

$$\lim_{k\to\infty} E[J^0_{k+1}(J^0_{k+1})^\dagger] = \lim_{k\to\infty} \sigma_v^2 \sum_{s=0}^{k}(1-\mu)^{2(k-s)}E[X_0X_0^\dagger]$$

and as a result

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] = \frac{\sigma_v^2}{\mu(2-\mu)}I.$$

Next,

$$\lim_{k\to\infty} E[J^0_{k+1}(J^1_{k+1})^\dagger] = \mu\sigma_v^2 \sum_{s=0}^{k}(1-\mu)^{2s-1}E[D_1(s,1)X_0X_0^\dagger].$$

Now, $E[Z_uX_0X_0^\dagger] = E[X_uX_u^\dagger]E[X_0X_0^\dagger] - E[X_uX_u^\dagger X_0X_0^\dagger] = -\frac{N}{P^3}\kappa^{2u}$ which gives

$$E[D_1(s,1)X_0X_0^\dagger] = -\frac{N}{P^3}\kappa^2\frac{1-\kappa^{2s}}{1-\kappa^2}.$$

Therefore, $\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] = -\frac{\kappa^2\sigma_v^2N}{2(1-\kappa^2)}I + O(\mu)I$. Next we consider,

$$\lim_{k\to\infty} E[J^1_{k+1}(J^1_{k+1})^\dagger] = \sigma_v^2\mu^2 \sum_{u=0}^{\infty}(1-\mu)^{2u-2}E[D_1(u,1)X_0X_0^\dagger D_1(u,1)^\dagger].$$

Note that,

$$\begin{aligned} E[Z_vX_0X_0^\dagger Z_u^\dagger] &= I - E[X_vX_v^\dagger X_0X_0^\dagger] - E[X_0X_0^\dagger X_uX_u^\dagger] + E[X_vX_v^\dagger X_0X_0^\dagger X_uX_u^\dagger] \\ &= [(N^2+1)\kappa^{v+u}\kappa^{|v-u|} + N\kappa^{2|v-u|}. \end{aligned}$$

Therefore,

$$E[D_1(u,1)X_0X_0^\dagger D_1(u,1)^\dagger] = (N^2+1)\sum_{s=1}^{u}\sum_{t=1}^{u}\kappa^{|v-u|}\kappa^{v+u} + N\sum_{s=1}^{u}\sum_{t=1}^{u}\kappa^{2|v-u|}$$

and consequently

$$\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] = \frac{(1+\kappa^2)\sigma_v^2N}{4(1-\kappa^2)}I + O(\mu)I.$$

Finally, we have $\lim_{k\to\infty} E[J^0_{k+1}(J^2_{k+1})^\dagger] = \mu^2\sigma_v^2\sum_{s=0}^{\infty}E[D_2(s,1)X_0X_0^\dagger]$. Now

$$\begin{aligned} E[Z_vZ_uX_0X_0^\dagger] &= I - E[X_vX_v^\dagger X_0X_0^\dagger] - E[X_uX_u^\dagger X_0X_0^\dagger] + E[X_vX_v^\dagger X_uX_u^\dagger X_0X_0^\dagger] \\ &= [(N^2+1)\kappa^{v+u}\kappa^{|v-u|} + N\kappa^{2|v-u|}. \end{aligned}$$

Therefore,

$$E[D_2(s,1)X_0X_0^\dagger] = (N^2+1)\sum_{s=1}^{u}\sum_{t=s+1}^{u}\kappa^{|v-u|}\kappa^{v+u} + N\sum_{s=1}^{u}\sum_{t=s+1}^{u}\kappa^{2|v-u|}$$

and

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] = \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)}I + O(\mu)I.$$

Second, we have the following expressions for SPULMS

$$\begin{aligned}
J_{k+1}^0 &= \sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s}I_sX_sn_s \\
J_{k+1}^1 &= \mu\sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s-1}D_1(k,s+1)I_sX_sn_s \\
J_{k+1}^2 &= \mu^2\sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s-2}D_2(k,s+1)I_sX_sn_s
\end{aligned}$$

where

$$\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1)Z_u
\end{aligned}$$

and $Z_u = I_uX_uX_u^\dagger - \frac{1}{P}E[X_uX_u^\dagger]$.

This leads to

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^0)^\dagger] = \lim_{k\to\infty}\sigma_v^2\sum_{s=0}^{k}(1-\mu)^{2(k-s)}E[I_0X_0X_0^\dagger I_0]$$

and therefore,

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^\dagger] = \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P})}I.$$

Next,

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^\dagger] = \mu\sigma_v^2\sum_{s=0}^{k}(1-\mu)^{2s-1}E[D_1(s,1)I_0X_0X_0^\dagger I_0].$$

Furthermore, $E[Z_u I_0 X_0 X_0^\dagger I_0] = E[I_u X_u X_u^\dagger] E[I_0 X_0 X_0^\dagger I_0] - E[I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] = -\frac{N}{P^3}\kappa^{2u}$ which gives

$$E[D_1(s,1)X_0X_0^\dagger] = -\frac{N}{P^3}\kappa^2\frac{1-\kappa^{2s}}{1-\kappa^2}$$

and as a result

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^\dagger] = -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)P}I + O(\mu)I.$$

Next, consider

$$\lim_{k\to\infty} E[J_{k+1}^1(J_{k+1}^1)^\dagger] = \sigma_v^2\mu^2\sum_{u=0}^{\infty}(1-\mu)^{2u-2}E[D_1(u,1)I_0X_0X_0^\dagger I_0 D_1(u,1)^\dagger].$$

Since,

$$\begin{aligned}
E[Z_v I_0 X_0 X_0 I_0 Z_u^\dagger] &= I - E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0] - E[I_0 X_0 X_0^\dagger I_0 X_u X_u^\dagger I_u] \\
&\quad + E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0 X_u X_u^\dagger I_u] \\
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|} + N\kappa^{2|v-u|}]I \quad \text{if } v \neq u \\
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|} + N\kappa^{2|v-u|}]I \\
&\quad + \frac{P-1}{P^3}[(N+1) + \frac{N^2+2N+1}{P}\kappa^{2u}]I \quad \text{if } v = u
\end{aligned}$$

we have $\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^\dagger] = \frac{\sigma_v^2}{4}[\frac{N}{P}\frac{1+\kappa^2}{1-\kappa^2} + (N+1)\frac{P-1}{P}]I + O(\mu)I$.

Finally, we have $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^2)^\dagger] = \mu^2\sigma_v^2\sum_{s=0}^{\infty}E[D_2(s,1)X_0X_0^\dagger]$. Furthermore,

$$\begin{aligned}
E[Z_v Z_u X_0 X_0^\dagger] &= \sigma^6 I - \sigma^2 E[I_v X_v X_v^\dagger I_0 X_0 X_0^\dagger I_0] - \sigma^2 E[I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] \\
&\quad + E[I_v X_v X_v^\dagger I_u X_u X_u^\dagger I_0 X_0 X_0^\dagger I_0] \\
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|} + N\kappa^{2|v-u|}]I \quad \text{if } v \neq u
\end{aligned}$$

which leads to $\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^\dagger] = \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)P}I + O(\mu)I$.

APPENDIX B

Appendices for Chapter 6

B.1 Capacity Optimization in Section 6.2.1

We have the following expression for the capacity

$$C = E \log \det(I_N + \frac{\rho}{M} H^\dagger \Lambda H)$$

where Λ is of the form

$$\Lambda = \begin{bmatrix} M-(M-1)d & l\underline{1}^\tau_{M-1} \\ l\underline{1}_{M-1} & dI_{M-1} \end{bmatrix}.$$

Let l^r denote the real part of l and l^i the imaginary part. We can find the optimal value of d and l iteratively by using the method of steepest descent as follows

$$\begin{aligned} d_{k+1} &= d_k + \mu \frac{\partial C}{\partial d_k} \\ l^r_{k+1} &= l^r_k + \mu \frac{\partial C}{\partial l^r_k} \\ l^i_{k+1} &= l^i_k + \mu \frac{\partial C}{\partial l^i_k} \end{aligned}$$

where d_k, l^r_k and l^i_k are the values of d, l^r and l^i respectively at the k^{th} iteration. We use the following identity (Jacobi's formula) to calculate the partial derivatives.

$$\frac{\partial \log \det A}{\partial d} = \mathrm{tr}\{A^{-1}\frac{\partial A}{\partial d}\}.$$

Therefore, we obtain

$$\frac{\partial C}{\partial d} = E\text{tr}\{[I_N + \frac{\rho}{M}H^\dagger \Lambda H]^{-1}\frac{\rho}{M}H^\dagger \frac{\partial \Lambda}{\partial d}H\}$$

and similarly for l^r and l^i where

$$\frac{\partial \Lambda}{\partial d} = \begin{bmatrix} -(M-1) & \underline{0}_{M-1}^\tau \\ \underline{0}_{M-1} & I_{M-1} \end{bmatrix}$$

$$\frac{\partial \Lambda}{\partial l^r} = \begin{bmatrix} 0 & \underline{1}_{M-1}^\tau \\ \underline{1}_{M-1} & \mathbf{0}_{M-1} \end{bmatrix}$$

$$\frac{\partial \Lambda}{\partial l^i} = \begin{bmatrix} 0 & \underline{1}_{M-1}^\tau \\ -\underline{1}_{M-1} & \mathbf{0}_{M-1} \end{bmatrix}.$$

The derivative can be evaluated using monte carlo simulation.

B.2 Non-coherent Capacity for low SNR values under Peak Power Constraint

In this section, we will use the notation introduced in Section 6.3.3. Here we concentrate on calculating the capacity under the constraint $\text{tr}\{SS^\dagger\} \leq TM$.

Theorem B.1. *Let the channel H be Rician (6.6) and the receiver have no knowledge of G. For fixed M, N and T under the peak power constraint*

$$C = rT\rho\lambda_{max}(H_m H_m^\dagger) + O(\rho^{3/2}).$$

Proof: First, Define $p(\tilde{X}) = E[p(\tilde{X}|\tilde{S})]$ where

$$p(\tilde{X}|\tilde{S}) = \frac{1}{\pi^{TN}\Lambda_{\tilde{X}|\tilde{S}}} e^{-(\tilde{X}-\sqrt{r\frac{\rho}{M}}\hat{H}_m\tilde{S})^\dagger \ \Lambda_{\tilde{X}|\tilde{S}}^{-1} \ (\tilde{X}-\sqrt{r\frac{\rho}{M}}\hat{H}_m\tilde{S})}$$

Now

$$\mathcal{H}(\tilde{X}) = E_{\|\tilde{X}\|<(\frac{M}{\rho})^\gamma}[\log p(\tilde{X})] + E_{\|\tilde{X}\|\geq(\frac{M}{\rho})^\gamma}[\log p(\tilde{X})]$$

$E_{\|\tilde{X}\| \geq (\frac{M}{\rho})^\gamma}$ is defined by (6.7). Since $P(\|\tilde{X}\| \geq (\frac{M}{\rho})^\gamma) < O(e^{-(\frac{M}{\rho})^\gamma/TM})$ where we have chosen γ such that $1 - 2\gamma > 1/2$ or $\gamma < 1/4$. We have

$$\mathcal{H}(\tilde{X}) = E_{\|\tilde{X}\| < (\frac{M}{\rho})^\gamma}[\log p(\tilde{X})] + O(e^{-\frac{1}{TM}(\frac{M}{\rho})^\gamma}).$$

For $\|\tilde{X}\| < (\frac{M}{\rho})^\gamma$

$$\begin{aligned}
p(\tilde{X}|\tilde{S}) \;=\; & \frac{1}{\pi^{TN}} e^{-\tilde{X}^\dagger \tilde{X}} \left[1 + \sqrt{r\frac{\rho}{M}}(\tilde{X}^\dagger \hat{H}_m \tilde{S} + \tilde{S}^\dagger \hat{H}_m^\dagger \tilde{X}) - \right. \\
& \frac{\rho}{M}\left(\mathrm{tr}\{(1-r)SS^\dagger \otimes I_N\} + \mathrm{tr}\{r\tilde{S}^\dagger \hat{H}_m^\dagger \hat{H}_m \tilde{S}\}\right) + \\
& (1-r)\frac{\rho}{M}\tilde{X}^\dagger SS^\dagger \otimes I_N \tilde{X} + r\frac{1}{2}\frac{\rho}{M}\left(\tilde{S}^\dagger \hat{H}_m^\dagger \tilde{X}\tilde{S}^\dagger \hat{H}_m^\dagger \tilde{X} + \tilde{S}^\dagger \hat{H}_m^\dagger \tilde{X}\tilde{X}^\dagger \hat{H}_m \tilde{S} + \right. \\
& \left.\left. \tilde{X}^\dagger \hat{H}_m \tilde{S}\tilde{S}^\dagger \hat{H}_m^\dagger \tilde{X} + \tilde{X}^\dagger \hat{H}_m \tilde{S}\tilde{X}^\dagger \hat{H}_m \tilde{S}\right) + O(\rho^{3/2-3\gamma})\right].
\end{aligned}$$

Since the capacity achieving signal has zero mean, for $\|\tilde{X}\| < (\frac{M}{\rho})^\gamma$

$$\begin{aligned}
p(\tilde{X}) \;=\; & \frac{1}{\pi^{TN}} e^{-\tilde{X}^\dagger \tilde{X}} \left[1 - \frac{\rho}{M}\left(\mathrm{tr}\{(1-r)E[SS^\dagger] \otimes I_N\} + \mathrm{tr}\{r\hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger\}\right) + \right. \\
& \left. \frac{\rho}{M}((1-r)\tilde{X}^\dagger E[SS^\dagger] \otimes I_N \tilde{X} + r\tilde{X}^\dagger \hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger \tilde{X} + O(\rho^{3/2-3\gamma})\right] \\
=\; & \frac{1}{\pi^{TN}\det(\Lambda_{\tilde{X}})} e^{-\tilde{X}^\dagger \Lambda_{\tilde{X}}^{-1}\tilde{X}} + \frac{1}{\pi^{TN}} e^{-\tilde{X}^\dagger \tilde{X}}[O(\rho^{3/2-3\gamma})]
\end{aligned}$$

where $\Lambda_{\tilde{X}} = I_{TN} + \frac{\rho}{M}(1-r)E[SS^\dagger] \otimes I_N + \frac{\rho}{M}r\hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger$. Also,

$$\begin{aligned}
\mathcal{H}(\tilde{X}) \;=\; & \log\det(I_{TN} + \frac{\rho}{M}(1-r)E[SS^\dagger] \otimes I_N + \frac{\rho}{M}r\hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger) + O(\rho^{3/2-3\gamma}) \\
=\; & \frac{\rho}{M}\mathrm{tr}\{(1-r)E[SS^\dagger] \otimes I_N + r\hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger\} + O(\rho^{3/2-3\gamma}).
\end{aligned}$$

Since $P(\|S\|^2 > TM) = 0$ we can show $\mathcal{H}(\tilde{X}|\tilde{S}) = (1-r)\frac{\rho}{M}\text{tr}\{E[SS^\dagger]\otimes I_N\}+O(\rho^2)$. Since $0 < \gamma < 1/4$, $I(X;S) = r\frac{\rho}{M}\text{tr}\{\hat{H}_m E[\tilde{S}\tilde{S}^\dagger]\hat{H}_m^\dagger\} + O(\rho^{3/2})$. It is very clear that to maximize C we need to choose $E[\tilde{S}\tilde{S}^\dagger]$ in such a way that all the energy is concentrated in the direction of the maximum eigenvalues of $H_m H_m^\dagger$. So that we obtain, $C = r\frac{\rho}{M}\lambda_{max}(H_m H_m^\dagger)\text{tr}E[\tilde{S}\tilde{S}^\dagger]+O(\rho^{3/2})$. $\text{tr}E[\tilde{S}\tilde{S}^\dagger]$ is maximized by choosing $\text{tr}\{\tilde{S}\tilde{S}^\dagger\}$ to be the maximum possible which is TM. Therefore,

$$C = r\rho T\lambda_{max}(H_m H_m^\dagger) + O(\rho^{3/2}).$$

■

Corollary B.1. *For purely Rayleigh fading channels* $\lim_{\rho\to 0} C/\rho = 0$

B.3 Proof of Lemma 6.3 in Section 6.3.4

In this section we will show that as $\sigma^2 \to 0$ or as $\rho \to \infty$ for the optimal input $(s_i^{(\sigma)}, i = 1,\dots,M)$, $\forall\delta,\epsilon > 0$, $\exists\sigma_0$ such that for all $\sigma < \sigma_0$

$$P(\frac{\sigma}{\|s_i^\sigma\|} > \delta) < \epsilon \tag{B.1}$$

for $i = 1,\dots,M$. $s_i^{(\sigma)}$ denotes the optimum input signal being transmitted over antenna i, $i = 1\dots,M$ when the noise power at the receiver is σ^2. Also, throughout we use ρ to denote the average signal to noise ratio M/σ^2 present at each of the receive antennas.

The proof in this section has basically been reproduced from [85] except for some minor changes to account for the deterministic specular component (H_m) present in the channel.

The proof is by contradiction. We need to show that if the distribution P of a source $s_i^{(\sigma)}$ satisfies $P(\frac{\sigma}{\|s_i\|} > \delta) > \epsilon$ for some ϵ and δ and for arbitrarily small σ^2, there exists σ^2 such that $s_i^{(\sigma)}$ is not optimal. That is, we can construct another input distribution that satisfies the same power constraint, but achieves higher mutual information. The steps in the proof are as follows

1. We show that in a system with M transmit and N receive antennas, coherence time $T \geq 2N$, if $M \leq N$, there exists a finite constant $k_1 < \infty$ such that for any fixed input distribution of S, $I(X; S) \leq k_1 + M(T - M) \log \rho$. That is, the mutual information increases with SNR at a rate no higher than $M(T-M) \log \rho$.

2. For a system with M transmit and receive antennas, if we choose signals with significant power only in M' of the transmit antennas, that is $\|s_i\| \leq C\sigma$ for $i = M' + 1, \ldots, M$ and some constant C, we show that the mutual information increases with SNR at rate no higher than $M'(T - M') \log \rho$.

3. We show that for a system with M transmit and receive antennas if the input distribution doesn't satisfy (B.1), that is, has a positive probability that $\|s_i\| \leq C\sigma$, the mutual information achieved increases with SNR at rate strictly lower than $M(T - M) \log \rho$.

4. We show that in a system with M transmit and receive antennas for constant equal norm input $P(\|s_i\| = \sqrt{T}) = 1$, for $i = 1, \ldots, M$, the mutual information increases with SNR at rate $M(T - M) \log \rho$. Since $M(T - M) \geq M'(T - M')$ for any $M' \leq M$ and $T \geq 2M$, any input distribution that doesn't satisfy (B.1) yields a mutual information that increases at lower rate than constant equal norm input, and thus is not optimal at high enough SNR level.

Step 1 For a channel with M transmit and N receive antennas, if $M < N$ and $T \geq 2N$, we write the conditional differential entropy as

$$\mathcal{H}(X|S) = N \sum_{i=1}^{M} E[\log((1 - r)\|s_i\|^2 + \sigma^2)] + N(T - M) \log \pi e \sigma^2.$$

Let $X = \Phi_X \Sigma_X \Psi_X^\dagger$ be the SVD for X then

$$\begin{aligned}\mathcal{H}(X) &\leq \mathcal{H}(\Phi_X) + \mathcal{H}(\Sigma_X|\Psi) + \mathcal{H}(\Psi) + E[\log J_{T,N}(\sigma_1, \ldots, \sigma_N)] \\ &\leq \mathcal{H}(\Phi_X) + \mathcal{H}(\Sigma_X) + \mathcal{H}(\Psi) + E[\log J_{T,N}(\sigma_1, \ldots, \sigma_N)] \\ &= \log|R(N,N)| + \log|R(T,N)| + \mathcal{H}(\Sigma_X) + E[\log J_{T,N}(\sigma_1, \ldots, \sigma_N)]\end{aligned}$$

where $R(T,N)$ is the Steifel Manifold for $T \geq N$ [85] and is defined as the set of all unitary $T \times N$ matrices. $|R(T,N)|$ is given by

$$|R(T,N)| = \prod_{i=T-N+1}^{T} \frac{2\pi^i}{(i-1)!}.$$

$J_{T,N}(\sigma_1, \ldots, \sigma_N)$ is the Jacobian of the transformation $X \to \Phi_X \Sigma_X \Psi_X^\dagger$ [85] and is given by

$$J_{T,N} = (\frac{1}{2\pi})^N \prod_{i<j\leq N} (\sigma_i^2 - \sigma_j^2)^2 \prod_{i=1}^{N} \sigma_i^{2(T-M)+1}.$$

We have also chosen to arrange σ_i in decreasing order so that $\sigma_i > \sigma_j$ if $i < j$. Now

$$\begin{aligned}\mathcal{H}(\Sigma_X) &= \mathcal{H}(\sigma_1, \ldots, \sigma_M, \sigma_{M+1}, \ldots, \sigma_N) \\ &\leq \mathcal{H}(\sigma_1, \ldots, \sigma_M) + \mathcal{H}(\sigma_{M+1}, \ldots, \sigma_N)\end{aligned}$$

Also,

$$\begin{aligned}
E[\log J_{T,N}(\sigma_1,\ldots,\sigma_N)] &= \log\frac{1}{(2\pi)^N} + \sum_{i=1}^{N} E[\log\sigma_i^{2(T-N)+1}] + \sum_{i<j\le N} E[\log(\sigma_i^2-\sigma_j^2)^2] \\
&= \log\frac{1}{(2\pi)^M} + \sum_{i=1}^{M} E[\log\sigma_i^{2(T-N)+1}] + \\
&\quad \sum_{i<j\le M} E[\log(\sigma_i^2-\sigma_j^2)^2] + \sum_{i\le M, M<j\le N} E[\underbrace{\log(\sigma_i^2-\sigma_j^2)^2}_{\le\log\sigma_i^4}] + \\
&\quad \log\frac{1}{(2\pi)^{N-M}} + \\
&\quad \sum_{i=M+1}^{N} E[\log\sigma_i^{2(T-N)+1}] + \sum_{M<i<j\le N} E[\log(\sigma_i^2-\sigma_j^2)^2] \\
&\le E[\log J_{N,M}(\sigma_1,\ldots,\sigma_M)] \\
&\quad +E[\log J_{T-M,N-M}(\sigma_{M+1},\ldots,\sigma_N)] \\
&\quad +2(T-M)\sum_{i=1}^{M} E[\log\sigma_i^2].
\end{aligned}$$

Next define $C_1 = \Phi_1\Sigma_1\Psi_1^\dagger$ where $\Sigma_1 = \text{diag}(\sigma_1,\ldots,\sigma_M)$, Φ_1 is a $N\times M$ unitary matrix, Ψ_1 is a $M\times M$ unitary matrix. Choose Σ_1, Φ_1 and Ψ_1 to be independent of each other. Similarly define C_2 from the rest of the eigenvalues. Now

$$\begin{aligned}
\mathcal{H}(C_1) &= \log|R(M,M)| + \log|R(N,M)| + \mathcal{H}(\sigma_1,\ldots,\sigma_M) + E[\log J_{N,M}(\sigma_1,\ldots,\sigma_M)] \\
\mathcal{H}(C_2) &= \log|R(N-M,N-M)| + \log|R(T-M,N-M)| \\
&\quad +\mathcal{H}(\sigma_{M+1},\ldots,\sigma_N) + E[\log J_{T-M,N-M}(\sigma_{M+1},\ldots,\sigma_N)].
\end{aligned}$$

Substituting in the formula for $\mathcal{H}(X)$, we obtain

$$\begin{aligned}
\mathcal{H}(X) &\le \mathcal{H}(C_1) + \mathcal{H}(C_2) + (T-M)\sum_{i=1}^{M} E[\log\sigma_i^2] + \log|R(T,N)| + \log|R(N,N)| \\
&\quad -\log|R(N,M)| - \log|R(M,M)| - \log|R(N-M,N-M)| - \\
&\quad \log|R(T-M,N-M)| \\
&= \mathcal{H}(C_1) + \mathcal{H}(C_2) + (T-M)\sum_{i=1}^{M} E[\log\sigma_i^2] + \log|G(T,M)|.
\end{aligned}$$

Note that C_1 has bounded total power

$$\mathrm{tr}\{E[C_1 C_1^\dagger]\} = \mathrm{tr}\{E[\sigma_i^2]\} = \mathrm{tr}\{E[XX^\dagger]\} \leq NT(M+\sigma^2).$$

Therefore, the differential entropy of C_1 is bounded by the entropy of a random matrix with entries iid Gaussian distributed with variance $\frac{T(M+\sigma^2)}{M}$ [13, p. 234, Theorem 9.6.5]. That is

$$\mathcal{H}(C_1) \leq NM \log\left[\pi e \frac{T(M+\sigma^2)}{M}\right].$$

Similarly, we bound the total power of C_2. Since $\sigma_{M+1}, \ldots, \sigma_N$ are the $N-M$ least singular values of X, for any $(N-M)\times N$ unitary matrix Q.

$$\mathrm{tr}\{E[C_2 C_2^\dagger]\} \leq (N-M)T\sigma^2.$$

Therefore, the differential entropy is maximized if C_2 has independent iid Gaussian entries and

$$\mathcal{H}(C_2) \leq (N-M)(T-M)\log\left[\pi e \frac{T\sigma^2}{T-M}\right].$$

Therefore, we obtain

$$\begin{aligned}\mathcal{H}(X) \;\; \leq \;\; & \log|G(T,M)| + NM\log\left[\pi e\frac{T(M+\sigma^2)}{M}\right] + (T-M)\sum_{i=1}^{M} E[\log \sigma_i^2] \\ & + (N-M)(T-M)\log \pi e \sigma^2 + (N-M)(T-M)\log\frac{T}{T-M}.\end{aligned}$$

Combining with $\mathcal{H}(X|S)$, we obtain

$$\begin{aligned} I(X;S) \;\; \leq \;\; & \underbrace{\log|G(T,M)| + NM\log\frac{T(M+\sigma^2)}{M} + (N-M)(T-M)\log\frac{T}{T-M}}_{\alpha} \\ & + \underbrace{(T-M-N)\sum_{i=1}^{M} E[\log\sigma_i^2]}_{\beta} + \\ & \underbrace{N\left(\sum_{i=1}^{M} E[\log\sigma_i^2] - \sum_{i=1}^{M} E[\log((1-r)\|s_i\|^2+\sigma^2)]\right)}_{\gamma} \\ & - M(T-M)\log \pi e\sigma^2. \end{aligned}$$

By Jensen's inequality

$$\begin{aligned}\sum_{i=1}^{M} E[\log \sigma_i^2] &\leq M \log(\frac{1}{M}\sum_{i=1}^{M} E[\sigma_i^2]) \\ &= M \log \frac{NT(M+\sigma^2)}{M}.\end{aligned}$$

For γ it will be shown that

$$\sum_{i=1}^{M} E[\log \sigma_i^2] - \sum_{i=1}^{M} E[\log((1-r)\|s_i\|^2 + \sigma^2)] \leq k$$

where k is some finite constant.

Given S, X has mean $\sqrt{r}SH_m$ and covariance matrix $I_N \otimes ((1-r)SS^\dagger + \sigma^2 I_T)$. If $S = \Phi V \Psi^\dagger$ then

$$\begin{aligned}X^\dagger X &= H^\dagger S^\dagger S H + W^\dagger S H + H^\dagger S^\dagger W + W^\dagger W \\ &\stackrel{d}{=} H_1^\dagger V^\dagger V H_1 + W^\dagger V^\dagger H_1 + H_1^\dagger V W + W^\dagger W\end{aligned}$$

where H_1 has the covariance matrix as H but mean is given by $\sqrt{r}\Psi^\dagger H_m$. Therefore, $X^\dagger X = X_1^\dagger X_1$ where $X_1 = VH_1 + W$

Now, X_1 has the same distribution as $((1-r)VV^\dagger + \sigma^2 I_T)^{1/2} Z$ where Z is a random Gaussian matrix with mean $\sqrt{r}((1-r)VV^\dagger + \sigma^2 I_T)^{-1/2}\Psi^\dagger H_m$ and covariance I_{NT}. Therefore,

$$X^\dagger X \stackrel{d}{=} Z^\dagger((1-r)VV^\dagger + \sigma^2 I_T)Z.$$

Let $(X^\dagger X|S)$ denote the realization of $X^\dagger X$ given S then

$$(X^\dagger X|S) \stackrel{d}{=} Z^\dagger \begin{bmatrix} (1-r)\|s_1\|^2+\sigma^2 & & & & & \\ & \ddots & & & & \\ & & (1-r)\|s_M\|^2+\sigma^2 & & & \\ & & & \sigma^2 & & \\ & & & & \ddots & \\ & & & & & \sigma^2 \end{bmatrix} Z.$$

Let $Z = [Z_1|Z_2]$ be the partition of Z such that

$$(X^\dagger X|S) \stackrel{d}{=} Z_1^\dagger((1-r)V^2 + \sigma^2 I_M)Z_1 + \sigma^2 Z_2^\dagger Z_2$$

where Z_1 has mean $\sqrt{r}((1-r)V^2+\sigma^2 I_M)^{-1/2}V\Psi^\dagger H_m$ and covariance I_{NM} and Z_2 has mean 0 and covariance $I_{N(T-M)}$

We use the following Lemma from [45]

Lemma B.1. *If C and B are both Hermitian matrices, and if their eigenvalues are both arranged in decreasing order, then*

$$\sum_{i=1}^{N}(\lambda_i(C) - \lambda_i(B))^2 \leq \|C-B\|_2^2$$

where $\|A\|_2^2 \stackrel{\text{def}}{=} \sum A_{ij}^2$*,* $\lambda_i(A)$ *denotes the* i^{th} *eigenvalue of Hermitian matrix* A*.*

Applying this Lemma with $C = (X^\dagger X|S)$ and $B = Z_1^\dagger(V^2 + \sigma^2 I_M)Z_1$ we obtain

$$\lambda_i(C) \leq \lambda_i(B) + \sigma^2\|Z_2^\dagger Z_2\|_2$$

for $i = 1, \ldots, M$ Note that $\lambda_i(B) = \lambda_i(B')$ where $B' = ((1-r)V^2 + \sigma^2 I_M)Z_1 Z_1^\dagger$. Let $k = E[\|Z_2^\dagger Z_2\|_2]$ be a finite constant. Now, since Z_1 and Z_2 are independent matrices (covariance of $[Z_1|Z_2]$ is a diagonal matrix)

$$\begin{aligned}
\sum_{i=1}^{M} E[\log\sigma_i^2|S] &\leq \sum_{i=1}^{M} E[\log(\lambda_i(((1-r)V^2+\sigma^2 I_M)Z_1Z_1^\dagger) + \sigma^2\|Z_2^\dagger Z_2\|_2)] \\
&= \sum_{i=1}^{M} E[E[\log(\lambda_i(((1-r)V^2+\sigma^2 I_M)Z_1Z_1^\dagger) + \sigma^2\|Z_2^\dagger Z_2\|_2) \mid Z_1]] \\
&\leq \sum_{i=1}^{M} E[\log(\lambda_i(((1-r)V^2+\sigma^2 I_M)Z_1Z_1^\dagger) + \sigma^2 k)] \\
&= E[\log\det(((1-r)V^2+\sigma^2 I_M)Z_1Z_1^\dagger + k\sigma^2 I_M)] \\
&= E[\log\det Z_1Z_1^\dagger] + E[\log\det((1-r)V^2 + \sigma^2 I_M + k\sigma^2(Z_1Z_1^\dagger)^{-1})]
\end{aligned}$$

where the second inequality follows from Jensen's inequality and taking expectation over Z_2. Using Lemma B.1 again on the second term, we have

$$\begin{aligned}\sum_{i=1}^{M} E[\log\sigma_i^2|S] &\leq E[\log\det Z_1 Z_1^\dagger] + E[\log\det((1-r)V^2+\sigma^2 I_M \\ &\quad +k\sigma^2\|(Z_1 Z_1)^{-1}\|_2 I_M)] \\ &\leq E[\log\det Z_1 Z_1^\dagger] + E[\log\det((1-r)V^2+k'\sigma^2 I_M)]\end{aligned}$$

where $k' = 1 + kE[\|Z_1 Z_1^\dagger\|_2]$ is a finite constant. Next, we have

$$\begin{aligned}\sum_{i=1}^{M} E[\log\sigma_i^2|S] - \sum_{i=1}^{M}\log((1-r)\|s_i\|^2+\sigma^2) &\leq E[\log\det Z_1 Z_1^\dagger] + \\ &\quad \sum_{i=1}^{M}\log\frac{(1-r)\|s_i\|^2+k'\sigma^2}{(1-r)\|s_i\|^2+\sigma^2} \\ &\leq E[\log\det Z_1 Z_1^\dagger] + k''\end{aligned}$$

where k'' is another constant. Taking Expectation over S, we have shown that $\sum_{i=1}^{M} E[\log\sigma_i^2] - \sum_{i=1}^{M} E[\log((1-r)\|s_i\|^2+\sigma^2)]$ is bounded above by a constant.

Note that as $\|s_i\| \to \infty$, $Z_1 \to \sqrt{\frac{1}{1-r}}H_1$ so that $E[Z_1 Z_1^\dagger] \to \frac{1}{1-r}E[H_1 H_1^\dagger] = \frac{1}{1-r}E[HH^\dagger]$.

Step 2 Now assume that there are M transmit and receive antennas and that for $N - M' > 0$ antennas, the transmitted signal has bounded energy, that is, $\|s_i\|^2 < C\sigma^2$ for some constant C. Start from a system with only M' transmit antennas, the extra power we send on the rest $M - M'$ antennas accrues only a limited capacity gain since the SNR is bounded. Therefore, we conclude that the mutual information must be no more than $k_2 + M'(T - M')\log\rho$ for some finite k_2 that is uniform for all SNR level and all input distributions.

Particularly, if $M' = M - 1$, ie we have at least 1 transmit antenna to transmit signal with finite SNR, under the assumption that $T \geq 2M$ (T greater than twice the number of receivers), we have $M'(T - M') < M(T - M)$. This means that the mutual information achieved has an upper bound that increases with log SNR at rate $M'(T-M')\log\rho$, which is a lower rate than $M(T-M)\log\rho$.

Step 3 Now we further generalize the result above to consider the input which on at least 1 antennas, the signal transmitted has finite SNR with a positive probability, that is $P(\|s_M\|^2 < C\sigma^2) = \epsilon$. Define the event $E = \{\|s_M\|^2 < C\sigma^2\}$, then the mutual information can be written as

$$\begin{aligned} I(X;S) &\leq \epsilon I(X;S|E) + (1-\epsilon)I(X;S|E^c) + I(E;X) \\ &\leq \epsilon(k_1 + (M-1)(T-M+1)\log\rho) + (1-\epsilon)(k_2 + M(T-M)\log\rho) + \\ &\quad \log 2 \end{aligned}$$

where k_1 and k_2 are two finite constants. Under the assumption that $T \geq 2M$, the resulting mutual information thus increases with SNR at rate that is strictly less than $M(T-M)\log\rho$.

Step 4 Here we will show that for the case of M transmit and receive antennas, the constant equal norm input $P(\|s_i\| = \sqrt{T}) = 1$ for $i = 1, \ldots, M$, achieves a mutual information that increases at a rate $M(T - M\log\rho$.

Lemma B.2. *For the constant equal norm input,*

$$\lim\inf_{\sigma^2\to 0}[I(X;S) - f(\rho)] \geq 0$$

where $\rho = M/\sigma^2$, *and*

$$f(\rho) = \log|G(T,M)| + (T-M)E[\log\det HH^\dagger] + M(T-M)\log\frac{T\rho}{M\pi e} - M^2\log[(1-r)T]$$

where $|G(T,M)|$ *is as defined in Lemma 6.2.*

Proof: Consider

$$\begin{aligned}
\mathcal{H}(X) &\geq \mathcal{H}(SH) \\
&= \mathcal{H}(QVH) + \log|G(T,M)| + (T-M)E[\log\det H^\dagger \Psi V^2 \Psi^\dagger H] \\
&= \mathcal{H}(QVH) + \log|G(T,M)| + M(T-M)\log T + (T-M)E[\log\det HH^\dagger] \\
\mathcal{H}(X|S) &\leq \mathcal{H}(QVH) + M\sum_{i=1}^{M} E[\log((1-r)\|s_i\|^2 + \sigma^2)] + M(T-M)\log \pi e \sigma^2 \\
&\approx \mathcal{H}(QVH) + M^2\log[(1-r)T] + M^2\frac{\sigma^2}{(1-r)T} + M(T-M)\log \pi e\sigma^2.
\end{aligned}$$

Therefore,

$$\begin{aligned}
I(X;S) &\geq \log|G(T,M)| + (T-M)E[\log\det HH^\dagger] - M(T-M)\log \pi e\sigma^2 + \\
&\quad M(T-M)\log T - M^2\log[(1-r)T] - M^2\frac{\sigma^2}{(1-r)T} \\
&= f(\rho) - M^2\frac{\sigma^2}{(1-r)T} \to f(\rho).
\end{aligned}$$

■

Combining the result in step 4 with results in Step 3 we see that for any input that doesn't satisfy (B.1) the mutual information increases at a strictly lower rate than for the equal norm input. Thus at high SNR, any input not satisfying (B.1) is not optimal and this completes the proof of Lemma 6.3.

B.4 Convergence of Entropies

The main results in this section are Theorems B.2 and B.3. Lemma B.3 is useful in establishing the proof of Theorem B.2.

Let $\chi_P(x)$ denote the characteristic function over a set P defined as $\chi_P(x) = 0$ if $x \notin P$ and $\chi_P(x) = 1$ if $x \in P$.

Lemma B.3. *Let $g : \mathbb{C}^P \to R$ be a positive bounded function whose region of support, support(g), is compact. If there exists a constant L such that $\int g(x)dx \leq L < 1/e$*

then $|\int g(x)\log g(x)dx| \leq \max\{|L\log L| + |L\log vol(support(g))|, |L\log A|\}$ *where* $A = \sup g(x)$.

Proof: First, $\int g(x)\log g(x)dx \leq \int g(x)\log A dx \leq L\log A$. Let $\int g(x)dx = I_g$. Consider the probability density function $g(x)/I_g$. We know that $\int \frac{g(x)}{I_g}\log\frac{g(x)}{I_g f(x)}dx \geq 0$ for all probability density functions $f(x)$. If

$$f(x) = \frac{\chi_{\text{support}(g)}}{\text{vol}(\text{support}(g))}$$

then

$$\int g(x)\log g(x)dx \geq \int g(x)\log(I_g f(x)) = I_g \log\frac{I_g}{\text{vol}(\text{support}(g))}.$$

This implies

$$\begin{aligned}
|\int g(x)\log g(x)| &\leq \max\{|L\log A|, |I_g\log\frac{I_g}{\text{vol}(\text{support}(g))}|\} \\
&\leq \max\{|L\log A|, |I_g\log I_g| + |I_g\log\text{vol}(\text{support}(g))|\} \\
&\leq \max\{|L\log A|, |L\log L| + |L\log\text{vol}(\text{support}(g))|\}.
\end{aligned}$$

The last inequality follows from the fact that for $x < 1/e$, $|x\log x|$ is an increasing function of x. ■

Theorem B.2. *Let* $\{X_i \in \mathbb{C}^P\}$ *be a sequence of continuous random variables with probability density functions,* $\{f_i\}$ *and* $X \in \mathbb{C}^P$ *be a continuous random variable with probability density function* f *such that* $f_i \to f$ *pointwise. If 1)* $\max\{f_i(x), f(x)\} \leq A < \infty$ *for all* i *and 2)* $\max\{\int \|x\|^\kappa f_i(x)dx, \int \|x\|^\kappa f(x)dx\} \leq L < \infty$ *for some* $\kappa > 1$ *and all* i *then* $\mathcal{H}(X_i) \to \mathcal{H}(X)$. $\|x\| = \sqrt{x^\dagger x}$ *denotes the Euclidean norm of* x.

Proof: The proof is based on showing that given an $\epsilon > 0$ there exists an R such that for all i

$$|\int_{\|x\|>R} f_i(x)\log f_i(x)dx| < \epsilon.$$

This R also works for $f(x)$.

Since $y \log y \to 0$ as $y \to 0$ we have $\max_{f(x) \le A} |f(x) \log f(x)| \le \max\{A \log A, e\} \stackrel{\text{def}}{=} K$. Therefore, $f_i(x) \log f_i(x)$ is bounded above by an L^1 function ($g = K\chi_{\|x\| \le R}$) and by the dominated convergence theorem we have

$$-\int_{\|x\| \le R} f_i(x) \log f_i(x) dx \to -\int_{\|x\| \le R} f(x) \log f(x) dx.$$

Now, to show that the integral outside of $\|x\| \le R$ is uniformly bounded for all f_i and f. Let g denote either f_i or f. We have $\int \|x\|^\kappa g(x) dx \le L$. Therefore, by Markov's inequality $\int_{R < \|x\| \le R+1} g(x) dx = I^R \le L/R^\kappa$. Choose R large enough so that for all $l > R$: $I^l < 1/e$. Now

$$|\int_{\|x\| > R} g(x) \log g(x) dx| \le \int_{\|x\| > R} |g(x) \log g(x)| dx = \sum_{l=R}^{\infty} \int_{B_l} |g(x) \log g(x)| dx$$

where $B_l = \{x : l < \|x\| \le l + 1\}$.

Consider the term $\int_{B_l} |g(x) \log g(x)| dx = G_l$. Also, define $A_+ = \{x : -\log g(x) > 0\}$ and $A_- = \{x : -\log g(x) < 0\}$ Now,

$$\begin{aligned} G_l &= \int_{A_+ \cap B_l} |g(x) \log g(x)| dx + \int_{A_- \cap B_l} |g(x) \log g(x)| dx \\ &= |\int_{A_+ \cap B_l} g(x) \log g(x) dx| + |\int_{A_- \cap B_l} g(x) \log g(x) dx|. \end{aligned}$$

From Lemma B.3, we have

$$G_l \le 2 \max\{|I^l \log I^l| + |I^l \log \text{vol}(\{B_l\})|, |I^l \log A|\}.$$

We know $\text{vol}(\{x : B_l\}) = o(l^{2P})$. Therefore,

$$\int_{B_l} |g(x) \log g(x)| dx \le \frac{Q}{l^\kappa} \log l$$

where Q is some sufficiently large constant. Therefore, we have

$$\int_{\|x\| > R} |g(x) \log g(x)| dx \le \sum_{l=R}^{\infty} \frac{Q}{l^\kappa} \log l = O(\log R / R^{\kappa - 1}).$$

Finally, as $\kappa > 1$ we can choose R sufficiently large to have $|\int_{\|x\|>R} g(x)\log g(x)dx| < \epsilon$. ■

Theorem B.3. *Let $\{X_i \in \mathbb{C}^P\}$ be a sequence of continuous random variables with probability density functions, f_i and $X \in \mathbb{C}^P$ be a continuous random variable with probability density function f. Let $X_i \xrightarrow{P} X$. If 1) $\int \|x\|^\kappa f_n(x)dx \leq L$ and $\int \|x\|^\kappa f(x)dx \leq L$ for some $\kappa > 1$ and $L < \infty$ 2) $f(x)$ is bounded then $\limsup_{i\to\infty} \mathcal{H}(X_i) \leq \mathcal{H}(X)$.*

Proof: We will prove this by constructing a density function g_i corresponding to f_i that maximizes the entropy at stage i then show that $\limsup_{i\to\infty} \mathcal{H}_{g_i} \leq \mathcal{H}(X)$ thus concluding $\limsup \mathcal{H}(X_i) \leq \mathcal{H}(X)$ where $\mathcal{H}_{g_i} \stackrel{\text{def}}{=} -\int g_i(x)\log g_i(x)dx$.

First we will show that for all g_i defined above there exists a single real number $R > 0$ such that $-\int_{\|x\|>R} g_i(x)\log g_i(x)dx \leq \epsilon$. Note that this is different from the condition in Theorem B.2 where we show $|\int_{\|x\|>R} g_i(x)\log g_i(x)dx| \leq \epsilon$. As in Theorem B.2 choose R large enough so that $I^l < 1/e$. Also define the two sets A_+ and A_- as in Theorem B.2 then

$$\begin{aligned}
-\int_{\|x\|>R} g(x)\log g(x)dx &= -\int_{A_+} g(x)\log g(x)dx - \int_{A_-} g(x)\log g(x)dx \\
&= -\sum_{l=R}^{\infty}\int_{B_l\cap A_+} g(x)\log g(x)dx - \int_{A_-} g(x)\log g(x)dx
\end{aligned}$$

where B_l is as defined in Theorem B.2. The last line follows from the Monotone Convergence Theorem. From the proof of Lemma B.3 we have $-\int_{B_l\cap A_+} g(x)\log g(x)dx \leq -I^l\log I^l + I^l\log \mathrm{vol}(B_l)$ Therefore

$$\begin{aligned}
-\int_{\|x\|>R} g(x)\log g(x)dx &\leq \sum_{l=R}^{\infty}[-I^l\log I^l + I^l\log \mathrm{vol}(B_l)] - \int_{A_-} g(x)\log g(x)dx \\
&\leq \sum_{l=R}^{\infty}[-I^l\log I^l + I^l\log \mathrm{vol}(B_l)]
\end{aligned}$$

and the sum in the last line is bounded above by $\sum_{l=R}^{\infty} \frac{Q}{l^{\kappa}} \log l = O(\log R/R^{\kappa-1})$. Therefore,

$$\max_g \{ -\int_{\|x\|>R} g(x) \log g(x) dx \ \} \leq O(\log R/R^{\kappa-1}).$$

From the proof of Theorem B.2, $|\int_{\|x\|>R} f(x) \log f(x) dx| = O(\log R/R^{\kappa-1})$.

Now let's concentrate on upperbounding $-\int_{\|x\|\leq R} f_i(x) \log f_i(x) dx$. Let $A = \sup f(x)$. For each n partition the region $\{\|x\| \leq R\}$ into n regions $P_m, m = 1, \ldots, n$ such that $A\frac{m-1}{n} \leq f(x) < A\frac{m}{n}$ for $x \in P_m$, $m < n$ and $A\frac{n-1}{n} \leq f(x) \leq A$ for $x \in P_n$. Now for each n, there exists a number M_n such that $\max_m |\int_{P_m} (f_i(x) - f(x)) dx| < \frac{1}{n} \min_m \int_{P_m} f(x) dx$ for all $i \geq M_n$. If $M_n \leq M_{n-1}$ set $M_n = M_{n-1} + 1$. Now, define the function $M(i)$ such that

$$M(i) = \begin{cases} 1, & 1 \leq i \leq M_2 \\ 2, & M_2 < i \leq M_3 \\ 3, & M_3 < i \leq M_4 \\ \vdots & \end{cases}$$

For each i, divide the region $\{\|x\| \leq R\}$ into $M(i)$ parts as defined in the previous paragraph: $P_n, n = 1, \ldots, M(i)$, and define $g_i(x)$ over $\{\|x\| \leq R\}$ as

$$g_i(x) = \sum_{n=1}^{M(i)} \chi_{P_n}(x) I_{n,i}/V_n$$

where $I_{n,i} = \int_{P_n} f_i(x) dx$, $V_n = \mathrm{vol}(P_n)$.

Now, it is easy to see that $-\int_{\|x\|\leq R} f_i(x) \log f_i(x) \leq -\int_{\|x\|\leq R} g_i(x) \log g_i(x)$. Also, note that $g_i(x) \to f(x)$ pointwise. Since $f(x)$ is bounded there exists a number N and a constant K such that $g_i(x) \leq K$ for all values of $i > N$, also $f(x) \leq K$. Therefore, using Theorem B.2 we conclude that $\lim -\int_{\|x\|\leq R} g_i(x) \log g_i(x) dx \to -\int_{\|x\|\leq R} f(x) \log f(x) dx$.

Therefore, $\limsup \mathcal{H}(X_i) \leq \limsup \mathcal{H}_{g_i} \leq \mathcal{H}(X)$. ∎

Lemma B.4 is useful for applying Theorem B.2 to applications in which it is known that the cummulative distribution functions are converging.

Lemma B.4. *Let a sequence of cummulative distribution functions $F_n(x)$ having continuous derivatives converge to a cummulative distribution function $F(x)$ which also has a continuous derivative. If $f_n(x)$ are uniformly continuous then $f_n(x)$ converge to $f(x)$.*

Proof: Since $F_n(x)$ is absolutely continuous and converging to $F(x)$, we have for all y

$$P_n(A_y = \{x : |y - x| < \delta\}) = \int_{A_y} f_n(x)dx \to \int_{A_y} f(x)dx = P(A_y).$$

Since $f_n(x)$ are uniformly continuous, given ϵ there exists a single δ for all n such that $|f_n(x + \Delta x) - f_n(x)| < \epsilon$ for all $|\Delta x| < \delta$.

We have $|f_n(x) - f_n(y)| < \epsilon \; \forall x \in A_y$ and $|\int_{A_y} f_n(x)dx - f_n(y)\text{vol}(A_y)| < \epsilon \; \text{vol}(A_y)$. Since $|\int_{A_y} f(x)dx - f(y)\text{vol}(A_y)| < \epsilon \; \text{vol}(A_y)$ and $\int_{A_y} f_n(x)dx \to \int_{A_y} f(x)dx$ we have $|\lim f_n(y) - f(y)| < 2\epsilon \; \text{vol}(A_y)$. Since ϵ is arbitrary we have $\lim f_n(y) = f(y)$ for all y. ■

B.5 Convergence of $\mathcal{H}(X)$ for $T > M = N$ needed in the proof of Theorem 6.4 in Section 6.3.4

First, we will show convergence for the case $T = M = N$ needed for Theorem 6.4 and then use the result to to show convergence for the general case of $T > M = N$. We need the following lemma to establish the result for $T = M = N$.

Lemma B.5. *If $\lambda_{min}(SS^\dagger) \geq \lambda > 0$ then $\forall n$ there exists an $\mathcal{M}$ such that $|f(X) - f(Z)| < \mathcal{M}\delta$ if $|X - Z| < \delta$.*

Proof: Let $Z = X + \Delta X$ with $|\Delta X| < \delta$ and $[\sigma^2 I_T + (1 - r)SS^\dagger] = D$. First, we will fix S and show that for all S, $f(X|S)$ satisfies the above property. Therefore, it will follow that $f(X)$ also satisfies the same property. Consider $f^0(X|S)$ the density defined with zero mean which is just a translated version of $f(X|S)$.

$$f(X+\Delta X|S) = f(X|S)[1 - \mathrm{tr}[D^{-1}(\Delta X X^\dagger + X\Delta X^\dagger + O(\|\Delta X\|_2^2))]]$$

then

$$|f(X+\Delta X|S) - f(X|S)| \leq f(X|S)|\mathrm{tr}[D^{-1}(\Delta X X^\dagger + X\Delta X^\dagger)] + \mathrm{tr}[D^{-1}\|\Delta X\|_2^2]|.$$

Now

$$f(X|S) \leq \frac{1}{\pi^{TN}\det^N[D]} \cdot \min\{\frac{1}{\sqrt{\mathrm{tr}[D^{-1}XX^\dagger]}}, 1\}.$$

Next, make use of the following inequalities

$$\begin{aligned} \mathrm{tr}\{D^{-1}XX^\dagger\} &\geq \mathrm{tr}\{\lambda_{min}(D^{-1})XX^\dagger\} \\ &\geq \lambda_{min}(D^{-1})\lambda_{max}(XX^\dagger) = \lambda_{min}(D^{-1})\|X\|_2^2. \end{aligned}$$

Also,

$$\begin{aligned} |\mathrm{tr}\{D^{-1}(X\Delta X^\dagger + \Delta X X^\dagger + O(\|\Delta X\|_2^2)\}| &\leq \sum_i |\lambda_i(D^{-1}[\Delta X X^\dagger + X\Delta X^\dagger])| + \\ &\quad \|D^{-1}\|_2\|\Delta X\|_2^2 \\ &\leq T\|D^{-1}\|_2\|X\|_2\|\Delta X\|_2 + \\ &\quad T\|D^{-1}\|_2\|\Delta X\|_2^2. \end{aligned}$$

Therefore,

$$\begin{aligned} |f(X+\Delta X|S) - f(X|S)| &\leq \frac{1}{\pi^{TN}\det^N[D]} \cdot \min\{\frac{1}{\sqrt{\lambda_{min}(D^{-1})}\|X\|_2}, 1\} \cdot \\ &\quad T\|D^{-1}\|_2\|\Delta X\|_2(\|X\|_2 + \|\Delta X\|_2). \end{aligned}$$

Since, we have restricted $\lambda_{min}(SS^\dagger) \geq \lambda > 0$ we have for some constant $\mathcal{M}$

$$|f(X+\Delta X|S) - f(X|S)| \leq \mathcal{M}\|\Delta X\|_2.$$

From which the Lemma follows. Note that $\det[D]$ compensates for $\sqrt{\lambda_{min}(D^{-1})}$ in the denominator. ■

Let's consider the $T \times N$ random matrix $X = SH + W$. The entries of $M \times N$ matrix H, $T = M = N$, are independent circular complex Normal random variables with non-zero mean and unit variance whereas the entries of W are independent circular complex Normal random variables with zero-mean and variance σ^2.

Let S be a random matrix such that $\lambda_{min}(SS^\dagger) \geq \lambda > 0$ with distribution, $F_{max}(S)$ chosen in such a way to maximize $I(X;S)$. For each value of $\sigma^2 = 1/n$, n an integer $\rightarrow \infty$, the density of X is

$$f(X) = E_S \left[\frac{e^{-\mathrm{tr}\{[\sigma^2 I_T + (1-r)SS^\dagger]^{-1}(X - \sqrt{rNM}SH_m)(X-\sqrt{rNM}SH_m)^\dagger\}}}{\pi^{TN}\det^N[\sigma^2 I_T + (1-r)SS^\dagger]} \right].$$

where the expectation is over $F_{max}(S)$. It is easy to see that $f(X)$ as a function of σ^2 is a continuous function of σ^2. As $\lim_{\sigma^2 \rightarrow 0} f(X)$ exists, let's call this limit $g(X)$.

Since we have imposed the condition that $\lambda_{min}(SS^\dagger) \geq \lambda > 0$ w.p. 1, $f(X)$ is bounded above by $\frac{1}{(\lambda\pi)^{TN}}$. Thus $f(X)$ satisfies the condition for Theorem B.2. From Lemma B.5 we also have that for all n there exists a common δ such that $|f(X) - f(Z)| < \epsilon$ for all $|X - Z| < \delta$. Therefore, $\mathcal{H}(X) \rightarrow \mathcal{H}_g$. Since λ is arbitrary we conclude that for all optimal signals with the restriction $\lambda_{min}(SS^\dagger) > 0$, $\mathcal{H}(X) \rightarrow \mathcal{H}_g$. Now, we claim that the condition $\lambda_{min} > 0$ covers all optimal signals. Otherwise, if $\lambda_{min}(SS^\dagger) = 0$ with finite probability then for all σ^2 we have $\min \|s_i\|^2 \leq L\sigma^2$ for some constant L with finite probability. This is a contradiction of the condition (6.10). This completes the proof of convergence of $\mathcal{H}(X)$ for $T = M = N$. ■

Now, we show convergence of $\mathcal{H}(X)$ for $T > M = N$. We will show that $\mathcal{H}(X) \approx \mathcal{H}(SH)$ for small values of σ where $S = \Phi V \Psi^\dagger$ with Φ independent of V and Ψ.

Let $S_0 = \Phi_0 V_0 \Psi_0^\dagger$ denote a signal with its density set to the limiting optimal density of S as $\sigma^2 \to 0$.

$$\mathcal{H}(X) \geq \mathcal{H}(Y) = \mathcal{H}(Q\Sigma_Y \Psi_Y^\dagger) + \log|G(T,M)| + (T-M)E[\log\det\Sigma_Y^2]$$

where $Y = SH$ and Q is an isotropic matrix of size $N \times M$. Let

$$Y_Q = QV\Psi^\dagger H$$

Then $\mathcal{H}(Q\Sigma_Y\Psi_Y^\dagger) = \mathcal{H}(Y_Q)$.

From the proof of the case $T = M = N$, we have $\lim_{\sigma^2\to 0}\mathcal{H}(Y_Q) = \mathcal{H}(QV_0\Psi_0^\dagger H)$. Also,

$$\lim_{\sigma^2\to 0} E[\log\det\Sigma_Y^2] = E[\log\det\Sigma_{Y_0}^2]$$

where $Y_0 = S_0 H$ Therefore, $\liminf_{\sigma^2\to 0}\mathcal{H}(X) \geq \lim_{\sigma^2\to 0}\mathcal{H}(Y) = \mathcal{H}(S_0H)$.

Now, to show $\lim_{\sigma^2\to 0}\mathcal{H}(X) \leq \mathcal{H}(S_0H)$. From before

$$\mathcal{H}(X) = \mathcal{H}(Q\Sigma_X\Psi_X^\dagger) + |G(T,N)| + (T-M)E[\log\det\Sigma_X^2].$$

Now $Q\Sigma_X\Psi_X^\dagger$ converges in distribution to $QV_0\Psi_0^\dagger H$. Since the density of $QV_0\Psi_0^\dagger H$ is bounded, from Theorem B.3 we have $\limsup_{\sigma^2\to 0}\mathcal{H}(Q\Sigma_X\Psi_X^\dagger) \leq \mathcal{H}(QV_0\Psi_0^\dagger H)$. Also, note that $\lim_{\sigma^2\to 0} E[\log\det\Sigma_X^2] = E[\log\det\Sigma_{Y_0}^2] = \lim_{\sigma^2\to 0} E[\log\det\Sigma_Y^2]$. Which leads to $\limsup_{\sigma^2\to 0}\mathcal{H}(X) \leq \mathcal{H}(S_0H) = \lim_{\sigma^2\to 0}\mathcal{H}(SH)$.

Therefore, $\lim_{\sigma^2\to 0}\mathcal{H}(X) = \lim_{\sigma^2\to 0}\mathcal{H}(SH)$ and for small σ^2, $\mathcal{H}(X) \approx \mathcal{H}(SH)$.

B.6 Proof of Theorem 6.7 in Section 6.4.1

First we note that $\sigma_{\hat{G}}^2 = 1 - \sigma_{\tilde{G}}^2$. This means that

$$\rho_{eff} = \frac{\kappa T\rho + T_c}{(1-r)\kappa T\rho\sigma_{\tilde{G}}^2 + T_c} - 1.$$

Therefore, to maximize ρ_{eff} we just need to minimize $\sigma_{\tilde{G}}^2$. Now,

$$\sigma_{\tilde{G}}^2 = \frac{1}{NM}\mathrm{tr}\{E[\tilde{\tilde{G}}\tilde{\tilde{G}}^\dagger]\}$$

where

$$E[\tilde{\bar{G}}\tilde{\bar{G}}^\dagger] = (I_M + (1-r)\frac{\rho}{M}S_t^\dagger S_t)^{-1} \otimes I_N$$

where $\rho = \frac{M}{\sigma^2}$. Therefore, the problem is the following

$$\min_{S_t:\mathrm{tr}\{S_t^\dagger S_t\} \leq (1-\kappa)TM} \frac{1}{M}\mathrm{tr}\{\left(I_M + (1-r)\frac{\rho}{M}S_t^\dagger S_t\right)^{-1}\}.$$

The problem above can be restated as

$$\min_{\lambda_1,\ldots,\lambda_M:\sum\lambda_m \leq (1-\kappa)TM} \frac{1}{M}\sum_{m=1}^{M}\frac{1}{1+(1-r)\frac{\rho}{M}\lambda_m}$$

where λ_m, $m = 1,\ldots,M$ are the eigenvalues of $S_t^\dagger S_t$. The solution to the above problem is $\lambda_1 = \ldots = \lambda_M = (1-\kappa)T$. Therefore, the optimum S_t satisfies $S_t^\dagger S_t = (1-\kappa)TI_M$.

This gives $\sigma_{\tilde{G}}^2 = \frac{1}{1+(1-r)\frac{\rho}{M}(1-\kappa)T}$. Also, for this choice of S_t we obtain the elements of $\hat{G}$ to be zero mean independent with Gaussian distribution. This gives

$$\rho_{eff} = \frac{\kappa T\rho[Mr + \rho(1-\kappa)T]}{T_c(M+\rho(1-\kappa)T) + (1-r)\kappa T\rho M}.$$

B.7 Proof of Theorem 6.8 in Section 6.4.1

First, from Theorem 6.7

$$\begin{aligned}
\rho_{eff} &= \frac{\kappa T\rho[Mr + \rho(1-\kappa)T]}{T_c(M+\rho(1-\kappa)T) + (1-r)\kappa T\rho M} \\
&= \frac{T\rho}{T_c - (1-r)M}\frac{(1-\kappa)\kappa + \kappa\frac{rM}{T\rho}}{\frac{MT_c + T\rho T_c}{T\rho[T_c-(1-r)M]} - \kappa} \qquad T_c \neq (1-r)M \\
&= \frac{T^2\rho^2}{T_c(M+T\rho)}[(1-\kappa)\kappa + \kappa\frac{rM}{T\rho}] \qquad T_c = (1-r)M.
\end{aligned}$$

Consider the following three cases for the maximization of ρ_{eff} over $0 \leq \kappa \leq 1$.

Case 1. $T_c = (1-r)M$:

We need to maximize $(1-\kappa)\kappa + \kappa\frac{rM}{T\rho}$ over $0 \leq \kappa < 1$. The maximum occurs at $\kappa = \kappa_0 = \min\{\frac{1}{2} + \frac{rM}{2T\rho}, 1\}$. In this case

$$\rho_{eff} = \frac{T^2\rho^2}{(1-r)M(M+T\rho)}[\kappa_0\frac{rM}{T\rho} + \kappa_0(1-\kappa_0)].$$

Case 2. $T_c > (1-r)M$:

In this case,

$$\rho_{eff} = \frac{T\rho}{T_c - (1-r)M}\frac{(1-\kappa)\kappa + \kappa\eta}{\gamma - \kappa}$$

where $\eta = \frac{rM}{T\rho}$ and $\gamma = \frac{MT_c + T\rho T_c}{T\rho[T_c-(1-r)M]} > 1$. We need to maximize $\frac{(1-\kappa)\kappa+\kappa\eta}{\gamma-\kappa}$ over $0 \leq \kappa \leq 1$ which occurs at $\kappa = \min\{\gamma - \sqrt{\gamma^2 - \gamma - \eta\gamma}, 1\}$. Therefore,

$$\rho_{eff} = \frac{T\rho}{T_c - (1-r)M}(\sqrt{\gamma} - \sqrt{\gamma - 1 - \eta})^2$$

when $\kappa < 1$. When $\kappa = 1$ we obtain $T_c = T$. Substituting $\kappa = 1$ in the expression for ρ_{eff}

$$\rho_{eff} = \frac{\kappa T\rho[Mr + \rho(1-\kappa)T]}{T_c(M + \rho(1-\kappa)T) + (1-r)\kappa T\rho M}$$

we obtain $\rho_{eff} = \frac{rT\rho}{T+(1-r)T\rho}$.

Case 3. $T_c < (1-r)M$:

In this case,

$$\rho_{eff} = \frac{T\rho}{(1-r)M - T_c}\frac{(1-\kappa)\kappa + \kappa\eta}{\kappa - \gamma}$$

where $\gamma = \frac{MT_c + T\rho T_c}{T\rho[T_c-(1-r)M]} < 0$. Maximizing $\frac{(1-\kappa)\kappa+\kappa\eta}{\gamma-\kappa}$ over $0 \leq \kappa \leq 1$ we obtain $\kappa = \min\{\gamma + \sqrt{\gamma^2 - \gamma - \gamma\eta}, 1\}$. Therefore, when $\kappa < 1$

$$\rho_{eff} = \frac{T\rho}{T_c - (1-r)M}(\sqrt{-\gamma} - \sqrt{-\gamma + 1 + \eta})^2$$

Similar to the case $T_c < (1-r)M$, when $\kappa = 1$ we obtain $T_c = T$ and $\rho_{eff} = \frac{rT\rho}{T+(1-r)T\rho}$.

B.8 Proof of Theorem 6.9 in Section 6.4.1

Note that optimization over T_c makes sense only when $\kappa < 1$. If $\kappa = 1$ then T_c obviously has to be set equal to T. First, we examine the case $T_c > (1-r)M$. The other two cases are similar. Let $Q = \min\{M, N\}$ and let λ_i denote the i^{th} non-zero eigenvalue of $\frac{H_1 H_1^\dagger}{M}$, $i = 1, \ldots, Q$. Then we have

$$C_t \geq \sum_{i=1}^{Q} \frac{T_c}{T} E \log(1 + \rho_{eff}\lambda_i).$$

Let C_l denote the RHS in the expression above. The idea is to maximize C_l as a function of T_c. We have

$$\frac{dC_l}{dT_c} = \sum_{i=1}^{Q} \left\{ \frac{1}{T} E \log(1 + \rho_{eff}\lambda_i) + \frac{T_c}{T}\frac{d\rho_{eff}}{dT_c} E\left[\frac{\lambda_i}{1 + \rho_{eff}\lambda_i}\right]\right\}.$$

Now, ρ_{eff} for $T_c > (1-r)M$ is given by

$$\rho_{eff} = \frac{T\rho}{T_c - (1-r)M}(\sqrt{\gamma} - \sqrt{\gamma - 1 - \eta})^2$$

where $\gamma = \frac{MT_c + T\rho T_c}{T\rho[T_c - (1-r)M]}$ and $\eta = \frac{rM}{T\rho}$. It can be easily verified that

$$\frac{d\rho_{eff}}{dT_c} = \frac{T\rho(\sqrt{\gamma} - \sqrt{\gamma - 1 - \eta})^2}{[T_c - (1-r)M]^2}\left[\sqrt{\frac{(1-r)M(M + T\rho)}{T_c(T_c + T\rho + rM)}} - 1\right].$$

Therefore,

$$\begin{aligned}\frac{dC_l}{dT_c} &= \frac{1}{T}\sum_{i=1}^{Q} E\Bigg[\log(1 + \rho_{eff}\lambda_i) - \\ &\qquad \frac{\rho_{eff}\lambda_i}{1 + \rho_{eff}\lambda_i}\frac{T_c}{T_c - (1-r)M}\left[1 - \sqrt{\frac{(1-r)M(M + T\rho)}{T_c(T_c + T\rho + rM)}}\right]\Bigg].\end{aligned}$$

Since, $\frac{T_c}{T_c - (1-r)M}\left[1 - \sqrt{\frac{(1-r)M(M+T\rho)}{T_c(T_c + T\rho + rM)}}\right] < 1$ and $\log(1 + x) - x/(1 + x) \geq 0$ for all $x \geq 0$ we have $\frac{dC_l}{dT_c} > 0$. Therefore, we need to increase T_c as much as possible to maximize C_l or $T_c = T - M$.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] I. Abou-Faycal and B. Hochwald, “Coding requirements for multiple-antenna channels with unknown rayleigh fading,” *Bell Labs Technical Memorandum*, Mar. 1999.

[2] I. C. Abou-Faycal, M. D. Trott, and S. S. (Shitz), “The capacity of discrete-time memoryless rayleigh fading channels,” *To Appear in IEEE Transactions on Information Theory.*

[3] T. Aboulnasr and K. Mayyas, “Selective coefficient update of gradient-based adaptive algorithms,” *Signal Processing*, vol. 47, no. 5, pp. 1421–1424, May 1999.

[4] R. Aguech, E. Moulines, and P. Priouret, “On a perturbation approach for the analysis of stochastic tracking algorithms,” *SIAM Journal on Control and Optimization*, vol. 39, no. 3, pp. 872–899, 2000.

[5] S. M. Alamouti, “A simple transmitter diversity scheme for wireless communications,” *IEEE Journal of Selected Areas in Communication*, pp. 1451–1458, 1998.

[6] T. S. Alexander, *Adaptive Signal Processing: Theory and Applications*, Springer-Verlag and New York, 1986.

[7] M. R. Andrews, P. P. Mitra, and R. deCarvalho, “Tripling the capacity of wireless communications using electromagnetic polarization,” *Nature*, vol. 409, pp. 316–318, Jan. 2001.

[8] S. L. Ariyavisitakul, “Turbo space-time processing to improve wireless channel capacity,” *IEEE Communications Magazine*, vol. 48, no. 8, pp. 1347–1359, Aug. 2000.

[9] N. Benvenuto, P. Bisaglia, A. Salloum, and L. Tomba, “Worst case equalizer for noncoherent hiperlan receivers,” *IEEE Trans. on Communications*, vol. 48, no. 1, pp. 28–36, January 2000.

[10] E. Biglieri, J. Proakis, and S. S. (Shitz), “Fading channels: Information-theoretic and communications aspects,” *IEEE Trans. on Inform. Theory*, vol. 44, no. 6, pp. 2619–2692, Oct. 1998.

[11] P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, 1995.

[12] G. Caire, P. D. Torino, and E. Biglieri, “Capacity of multi-antenna block fading channels,” *Communication Theory mini Conference*, pp. 11–15, 1999.

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley Series in Telecommunications, 1995.

[14] I. Csiszár and J. Kőrner, *Information Theory: Coding Theorems For Discrete Memoryless Systems*, New York and NY: Academic Press, 1981.

[15] J. Cui, D. D. Falconer, and A. U. H. Sheikh, “Performance evaluation of optimum combining in the presence of co-channel interference and channel correlation for wireless communication systems,” *Mobile Networks and Applications*, vol. 2, no. 4, pp. 315–324, Jan 1998.

[16] S. C. Douglas, "Analysis and implementation of the max-nlms adaptive filter," *Proceedings of ASIMOLAR-29*, vol. 1, pp. 659l–663, 1996.

[17] S. C. Douglas, "A family of normalized lms algorithms," *IEEE Signal Processing Letters*, vol. 1, no. 3, pp. 49–51, March 1994.

[18] S. C. Douglas, "Adaptive filters employing partial updates," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-44, pp. 209–216, March 1997.

[19] P. F. Driessen and G. J. Foschini, "On the capacity formula for multiple input-multiple output wireless channels: A geometric interpretation," *IEEE Trans. on Communications*, vol. 47, no. 2, pp. 173–176, February 1999.

[20] R. B. Ertel and P. Cardieri, "Overview of spatial channel models for antenna array communication systems," *IEEE Personal Communications*, pp. 10–22, Feb. 1998.

[21] F. R. Farrokhi, G. J. Foschini, A. Lozano, and R. A. Valenzuela, "Link-optimal space-time processing with multiple transmit and receive antennas," *To appear in IEEE Communication Letters.*

[22] A. Feuer and E. Weinstein, "Convergence analysis of lms filter with uncorrelated gaussian data," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-33, no. 1, pp. 222–229, Feb. 1985.

[23] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[24] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, March 1998.

[25] B. Friedlander and A. J. Weiss, "Effects of model errors on signal reconstruction using a sensor array," *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, vol. 4, pp. 552–555, 1993.

[26] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968.

[27] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. on Inform. Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.

[28] A. Ganesan and A. M. Sayeed, "Bandwidth-efficient exploitation of the degrees of freedom in a multipath fading channel," *Proc. of IEEE Symposium on Information Theory*, p. 161, June 2000.

[29] E. Gautheir, A. Yongaçoglu, and J.-Y. Chouinard, "Capacity of multiple antenna systems in rayleigh fading channels," *Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 275–279, 2000.

[30] M. J. Gingell, B. G. Hay, and L. D. Humphrey, "A block mode update echo canceller using custom lsi," *GLOBECOM Conference Record*, vol. 3, pp. 1394–1397, Nov. 1983.

[31] L. C. Godara, "Applications of antenna arrays to mobile communications and part ii: Beamforming and direction-of-arrival considerations," *IEEE Proceedings*, vol. 85, no. 8, pp. 1195–1245, August 1997.

[32] L. C. Godara, "Applications of antenna arrays to mobile communications and part i: Performance improvement and feasibility and and system considerations," *IEEE Proceedings*, vol. 85, no. 7, pp. 1031–1060, July 1997.

[33] M. Godavarti, "Implementation of a g.165 line echo canceller on texas instruments' tms320c3x and tms320c54x chips," *ICSPAT Conference Proceedings*, pp. 65–69, Sep. 1997.

[34] A. Gorokhov, "Transmit diversity versus sdma: Analytic and numerical comparisons," *Proc. of the IEEE Internation Conference on Communications*, vol. 2, pp. 1020–1024, 2000.

[35] A. Gorokhov, "Capacity of multiple-antenna rayleigh channel with a limited transmit diversity," *Proc. of IEEE Symposium on Information Theory*, p. 411, June 2000.

[36] L. Guo, "Stability of recursive stochastic tracking algorithms," *SIAM Journal on Control and Optimization*, vol. 32, pp. 1195–1225, 1994.

[37] L. Guo, L. Ljung, and G. J. Wang, "Necessary and sufficient conditions for stability of lms," *IEEE Trans. Automatic Control*, vol. 42, no. 6, , 1997.

[38] B. Hassibi and B. Hochwald, "Linear dispersion codes," *Proc. of IEEE Symposium on Information Theory*, p. 325, 2001.

[39] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?," *Bell Labs Technical Memorandum*, 2000.

[40] B. Hassibi and M. Khorrami, "Fully-diverse multiple-antenna signal constellations and fixed-point-free lie groups," *Proc. of IEEE Symposium on Information Theory*, p. 199, 2001.

[41] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall and Englewood Cliffs and NJ, 1996.

[42] A. O. Hero and T. L. Marzetta, "On computational cut-off rate for space-time coding," *Submitted to IEEE Transactions on Information Theory.*

[43] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading," *IEEE Trans. on Inform. Theory*, vol. 46, no. 2, pp. 543–564, March 2000.

[44] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, and W. Sweldens, "Systematic design of unitary space-time constellations," *IEEE Trans. on Inform. Theory*, vol. 46, no. 6, pp. 1962–1973, Sep. 2000.

[45] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1996.

[46] L. L. Horowitz and K. D. Senne, "Performance advantage of complex lms for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-29, no. 3, pp. 722–736, June 1981.

[47] B. L. Hughes, "Differential space-time modulation," *Submitted to IEEE Transactions on Information Theory*, Feb. 2000.

[48] B. L. Hughes, "Optimal space-time constellations from groups," *Submitted to IEEE Transactions on Information Theory*, March 2000.

[49] A. R. Kuruc, "Lower bounds on multiple-source direction finding in the presence of direction-dependent antenna-array-calibration errors," *Technical Report 799 and MIT Lincoln Laboratory*, Oct. 1989.

[50] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading channel," *IEEE Trans. on Inform. Theory*, vol. 45, no. 1, pp. 139–157, Jan 1999.

[51] D. Messerschmitt, D. Hedberg, C. Cole, A. Haoui, and P. Winship, "Digital voice echo canceller with a tms32020," *Application report: SPRA129 and Texas Instruments.*

[52] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time coding and signal processing for high data rate wireless communications," *IEEE Signal Processing Magazine*, pp. 77–92, May 2000.

[53] A. Narula, "Information theoretic analysis of multiple-antenna transmission diversity," *Ph.D. Dissertation and MIT and Cambridge and MA*, 1997.

[54] A. Narula, M. D. Trott, and G. W. Wornell, "Performance limits of coded diversity methods for transmitter antenna arrays," *IEEE Trans. on Inform. Theory*, vol. 45, no. 7, pp. 2418–2433, Nov. 1999.

[55] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw Hill, Inc., 1991.

[56] S. Rice, "Mathematical analysis of random noise," *Bell Systems Technical Journal*, vol. 23, , 1944.

[57] A. M. Sengupta and P. P. Mitra, "Capacity of multivariate channels with multiplicative noise: I-random matrix techniques and large-n expansions for full transfer matrices," *Bell Laboratories Technical Report*.

[58] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency division duplex (fdd) transmission systems using transmitter antenna diversity," *Proc. of the IEEE Vehicular Technology Conference*, pp. 508–511, 1993.

[59] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.

[60] N. Sharma, H.-J. Su, and E. Geraniotis, "A novel approach for multi-antenna systems," *Proc. of the IEEE Internation Conference on Communications*, vol. 3, pp. 1264–1269, 2000.

[61] A. Shokrollahi, "Double antenna diagonal space-time codes and continued fractions," *Proc. of IEEE Symposium on Information Theory*, p. 108, 2001.

[62] A. Shokrollahi, "Group characters and unitary space-time codes," *Proc. of IEEE Symposium on Information Theory*, p. 107, 2001.

[63] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms: Stability and Performance*, Prentice-Hall and Englewood Cliffs and NJ, 1995.

[64] P. Stoica, B. Ottersten, M. Viberg, and R. L. Moses, "Maximum likelihood array processing for stochastic coherent sources," *IEEE Trans. on Signal Processing*, vol. 44, no. 1, pp. 96–105, Jan. 1996.

[65] P. Stoica, M. Viberg, K. M. Wong, and Q. Wu, "Maximum-likelihood bearing estimation with partly calibrated arrays in spatially correlated noise fields," *IEEE Trans. on Signal Processing*, vol. 44, no. 4, pp. 888–899, Apr. 1996.

[66] A. Swindlehurst, "Maximum a posteriori approach to beamforming in the presence of calibration errors," *Proc. of the IEEE Workshop on Statistical Signal and Array Processing*, pp. 82–85, 1996.

[67] A. Swindlehurst, "Detection and estimation in the presence of signals with uncalibrated spatial structure," *Signal Processing*, vol. 50, pp. 29–37, Apr. 1996.

[68] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *Signal Processing Letters*, vol. 1, no. 2, pp. 29–31, Feb. 1994.

[69] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array – part i: Algorithms," *IEEE Trans. on Signal Processing*, vol. 44, no. 5, pp. 1184–1197, May 1996.

[70] V. Tarokh, A. Naguib, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criteria in the presence of channel estimation errors and mobility and and multiple paths," *IEEE Trans. on Inform. Theory*, vol. 47, no. 2, pp. 199–207, Feb. 1999.

[71] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. on Communications*, vol. 44, no. 2, pp. 744–765, March 1998.

[72] G. Tarrico, E. Biglieri, and G. Caire, "Limiting performance of block-fading channels with multiple antennas," *Proc. of the IEEE Information Theory and Communications Workshop*, pp. 27–29, 1999.

[73] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, Nov/Dec 1999.

[74] A.-J. van der Veen, S. Talwar, and A. Paulraj, "Blind estimation of multiple digital signals transmitted over fir channels," *Signal Processing Letters*, vol. 2, no. 5, pp. 99–102, May 1995.

[75] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. on Inform. Theory*, vol. 36, no. 5, pp. 1019–1030, September 1990.

[76] P. Viswanath, D. Tse, and V. Ananthram, "Asymptotically optimal waterfilling in multiple antenna multiple access channels," *Proc. of IEEE Symposium on Information Theory*, p. 466, June 2000.

[77] V. Weerackody, "Diversity of the direct-sequence spread spectrum system using multiple transmit antennas," *Proc. of the IEEE Internation Conference on Communications*, pp. 1775–1779, 1993.

[78] J. H. Winters, "Optimum combining in digital mobile radio with co-channel interference," *IEEE Journal of Selected Areas in Communication*, vol. 2, pp. 528–539, 1984.

[79] J. H. Winters, "Optimum combining for indoor radio systems with multiple users," *IEEE Trans. on Communications*, vol. 35, pp. 1222–1230, 1987.

[80] J. H. Winters and J. Salz, "Upper bounds on the bit error rate of optimum combining in wireless systems," *Proc. of the IEEE Vehicular Technology Conference*, vol. 2, pp. 942–946, 1994.

[81] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. on Communications*, vol. 42, pp. 1740–1751, 1994.

[82] A. Wittenben, "Basestation modulation diversity for digital simulcast," *Proc. of the IEEE Vehicular Technology Conference*, pp. 848–853, 1991.

[83] J. Yang, S. DAAS, and A. Swindlehurst, "Improved signal copy with partially known or unknown array response," *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.*, vol. 4, pp. 265–268, 1994.

[84] J. Yang and A. L. Swindlehurst, "The effects of array calibration errors on df-based signal copy performance," *IEEE Trans. on Signal Processing*, vol. 43, no. 11, pp. 2724–2732, Nov. 1995.

[85] L. Zheng and D. N. C. Tse, "Packing spheres in the grassmann manifold: A geometric approach to the non-coherent multi-antenna channel,".