ABSTRACT

ZHOU, QUAN. Wireless Communications with MIMO Systems: Analysis and Practice. (Under the direction of Dr. Huaiyu Dai).

Multiple input multiple output (MIMO) systems using multiple transmit and receive antennas are widely considered as the vital breakthrough that will allow future wireless systems to achieve higher date rates and link reliability with limited bandwidth and power resources. In this dissertation, we address four interesting topics in the wireless MIMO systems, in both point-to-point and multiuser environments. First, in a point-to-point MIMO spatial diversity system, usually the probability distribution function (PDF) of the received SNR is rather involved, which leads to the difficulty in analyzing the average symbol error rate (SER). We provide a succinct result at the high SNR region. Second, in point-to-point wireless MIMO communications, in order to protect the transmitted data against random channel impairment, we consider the problem of link adaptation, including rate adaptation and power control to improve the system performance and guarantee certain quality of service. Third, in a multiuser MIMO wireless network, there is another form of diversity called multiuser diversity which can be exploited to increase the system throughput. By analyzing the scheduling gain (defined as the rate difference between the opportunistic scheduling and round-robin scheduling scheme), we provide a complete analysis on the interaction between the spatial diversity and multiuser diversity. Fourth, in a multiuser MIMO wireless network, we propose a crosslayer-based scheduling scheme that exploits Tomlinson-Harashima Precoding (THP) at the physical (PHY) layer to reduce the multiuser scheduling burden at the medium access control (MAC) layer. Compared with some existing

scheduling schemes, the proposed scheme greatly reduces the scheduling complexity while simultaneously improves overall system performance.

WIRELESS COMMUNICATIONS WITH MIMO SYSTEMS:

ANALYSIS AND PRACTICE

by

QUAN ZHOU

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

ELECTRICAL ENGINEERING

Raleigh

2006

APPROVED BY:

Huaiyu Dai *Chair of Advisory Committee*	Brian L. Hughes
Keith Townsend	Alexandra Duel-Hallen
Jack W. Silverstein	

To my family

BIOGRAPHY

Quan Zhou received his B.E. degree in Communication and Control from Northern Jiaotong University, Beijing, China, in 1998. And he received his M.E. degree in Electronic Engineering from Tsinghua University, Beijing, China in 2001. In August 2001, he started his graduate studies at North Carolina State University in the Department of Electrical and Computer Engineering. His research interests are in the areas of wireless MIMO communications with emphasis on link adaptation and multiuser MIMO network. His current research focuses on nonlinear optimization in wireless communication systems with emphasis on crosslayer optimization.

The author intends to pursue a career in the area of Telecomunication and Information Technology.

ACKNOWLEDGMENTS

My graduate study at North Carolina State University is one of the most memorable periods of my life. I have benefited tremendously from my interactions with many extraordinary individuals. First of all, I would like to thank my advisor Professor Huaiyu Dai for his supervision, support and encouragement throughout my Ph.D. research. I sincerely appreciate his help in suggesting research topics, revising papers and solving some critical problems during my research. I am very glad to be one of his first students.

I would also like to thank my committee members Professor Keith Townsend, Professor Brian Hughes and Professor Alexandar Duel-Hallen for their expertise and advice. I learned the fundamentals of communication and signal processing theory through their classes. Their way of teaching showed me an excellent example of a systematic and straightforward fashion in presenting technical ideas. I would also like to thank Professor Jack. W. Silverstein and Professor Zhidong Bai for their helps in providing some proofs in chapter 4.

It was a great pleasure to have closely worked with my colleagues: Dr. Jia Liu, Hongyuan Zhang, Wenjun Li, Xingying Yu and Li Ma. I thank them for the numerous suggestions, guidance, comments, and enlightening discussions.

Last but not least, I would like to thank my family members for their support and encouragement during my graduate studies. Mere words cannot express how much I love and appreciate them. This thesis is a dedication for their love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

1.1 Overview of wireless MIMO communications

The use of multiple antennas at both the transmitter and receiver side, so as to form a multiple-input multiple-output (MIMO) antenna system, is an emerging technology that makes building both reliable and high data rate wireless networks a reality [17][56]. Compared with the conventional single-input single-output (SISO) system, MIMO system creates multiple spatial dimensions that can be exploited to improve the performance[1] of the wireless link. More specifically, such performance improvement comes from the array gain, diversity gain, multiplexing gain and interference cancellation introduced by MIMO systems, which are illustrated below.

Array gain is achieved by using multiple antennas at the transmitter and receiver so that the received signals can add coherently. To exploit the transmit/receive array gain, transmitter/receiver needs to have the channel state information (CSI). The transmit/receive array gain is proportional to the number of transmit/receive antennas.

Diversity can improve the reliability of the received signal strength and is achieved by transmitting the signal over multiple independent fading subchannels. Before the introduction of MIMO techniques, multiple subchannels are usually created by time division and frequency division (correspondingly such diversities are usually called time diversity and frequency diversity). MIMO systems provide a new form of diversity without additional cost in time or frequency, i.e., spatial diversity by means of transmitting the signal over multiple

[1] The performance is often measured as the average bit rate or average bit error rate of the wireless link.

independently fading paths created by multiple antennas. A well-known example to exploit the spatial diversity gain without CSI at the transmitter side is space-time coding, like the Alamouti code for two transmit antennas [2]. For a MIMO system with M transmit antennas and N receive antennas, a maximum spatial diversity order of $M \times N$ can be achieved.

In a rich scattering environment, MIMO channels can offer a linear ($\min(M, N)$) increase in capacity with the number of antennas without increasing the transmission power or bandwidth. The spatial multiplexing gain can be realized by transmitting and receiving parallel independent streams across the multiple antennas at both ends; a famous example is the pioneer Bell Labs Layered Space-Time (BLAST) architecture proposed by Foschini in [18].. An interesting tradeoff between the spatial diversity gain and multiplexing gain is revealed in [74].

In cellular communications, cochannel interference arises due to frequency reuse. Multiple antennas can be used in cellular communications to mitigate the cochannel interference thus increasing the reuse factor and improving the system capacity. The basic idea is to make use of the spatial channel response (usually a vector) between the desired user and interference users, and design the receive weighting vector to maximize the signal power to interference power ratio.

Due to the promising advantages brought by MIMO system, MIMO technology is being adopted by international standards organizations, such as high-speed packet data mode of third-generation cellular systems, high-speed wireless local area networks (WLAN, IEEE 802.11n) and high-speed wireless metropolitan area network (WMAN, IEEE802.16).

1.2 Outline of the Thesis

Our thesis focuses on the four interesting topics on MIMO systems, which is organized as follows.

In chapter 2, through the analysis of the probability distribution function (PDF) of the received signal to noise ratio (SNR), we study the asymptotic symbol error rate of point-to-point MIMO spatial diversity system. The results reveal a simple connection with system parameters, providing good insights for the design of MIMO diversity systems.

In chapter 3, in order to protect the transmitted data against random channel impairment in wireless MIMO communications, we consider link adaptation, such as rate adaptation and power control to improve the system performance and guarantee certain quality of service. We propose a joint antenna subset selection and link adaptation study for MIMO systems, including both uncorrelated and correlated MIMO channels. Specifically, we propose one simplified antenna selection and link adaptation rule based on the expected optimal number of active antennas for uncorrelated MIMO with Rayleigh fading, and one for correlated MIMO channels only based on the slowly varying channel correlation information. Our proposed algorithms demonstrate significant gains over traditional MIMO signaling while feasible for practical implementation through numerical results.

In chapter 4, through asymptotic analysis of the average system capacity and scheduling gain, we investigate the cross-layer interaction between the spatial diversity and multiuser diversity in wireless networks. Rigorous proofs and necessarily stronger results in terms of convergence are provided for some intuitions in this area. Equally important, explicit expressions of scheduling gain and average system capacity in various circumstances

that reveal inter-connections and fundamental tradeoffs among key system parameters are given, which afford us some insights in real system design.

In chapter 5, we propose a crosslayer approach that explores Tomlinson-Harashima Precoding (THP) at the physical layer to reduce the multiuser scheduling burden at the MAC layer, and improves the sum rate of the downlink multiuser MIMO systems. Our proposed scheme is further evaluated with imperfect feedback, obtained by the long range prediction (LRP) technique. Compared to some existing scheduling schemes, the proposed scheme approaches the performance upper bound in certain scenarios, while incurring much less computation complexity. Significant gains are still maintained with imperfect channel state information (CSI), fed back at a rate much lower than the data rate.

Finally, in chapter 6, we propose some open problems and future work of our thesis.

1.3 List of author's publication

Below are the publications during the author's Ph.D. research:

Journal publications:

[J1] **Q. Zhou** and H. Dai, "Joint Antenna Selection and Link Adaptation for MIMO Systems," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 1, pp.243-255, Jan. 2006.

[J2] **Q. Zhou** and H. Dai, "Asymptotic Analysis on the Interaction between Spatial Diversity and Multiuser Diversity in Wireless Network," submitted to *IEEE Transactions on Signal Processing*, Dec. 2005 (under second review).

[J3] **Q. Zhou** and H. Dai, "Asymptotic Analysis in MIMO MRT/MRC systems," submitted to *EURASIP Journal on Wireless Communications and Networking*, Jan. 2006.

Conference publications:

[C1] H .Dai and **Q. Zhou**, "Scheduling gain in spatial diversity systems: Asymptotic analysis," *2006 IEEE International Symposium on Information Theory* (ISIT), Seattle, WA, July 2006.

[C2] **Q. Zhou** and H. Dai, "Asymptotic Analysis on Spatial Diversity versus Multiuser Diversity in Wireless Networks," *Proc. IEEE International Conference on Communications* (ICC), Istanbul, Turkey, June 2006.

[C3] **Q. Zhou**, H. Dai and H. Zhang, "Joint Tomlinson-Harashina Precoding and Scheduling for Multiuser MIMO Downlink with Imperfect Feedback," *Proc. IEEE Wireless Communications and Networking Conference* (WCNC), Las Vegas, NE, Apr. 2006.

[C4] H**.** Dai, and **Q. Zhou**, "Asymptotic Analysis in MIMO Diversity Systems," *Proc. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, HK, CHINA, Dec. 2005.

[C5] **Q. Zhou** and H. Dai, "Joint Antenna Selection and Link Adaptation for MIMO Systems," *Proc. 2004 Fall IEEE Conference on Vehicular Technology (VTC)*, Los Angels, CA, Sept. 2004.

[C6] **Q. Zhou** and H. Dai, "Adaptive Spatial Multiplexing Techniques for Distributed MIMO Systems," *Proc. 2004 Conference on Information Sciences and Systems(CISS)*, Princeton University, NJ, March 2004.

Chapter 2

Asymptotic Symbol Error Rate of MIMO Spatial Diversity System

2.1 Background

Multi-input multi-output (MIMO) systems can be exploited for spatial multiplexing or diversity gains. For diversity usage, joint maximum ratio transmission (MRT) and maximum ratio combining (MRC) provides the optimal performance reference [42]-[55]. With the assumption of identical transmit and receive beamforming vectors, the average output signal-to-noise ratio (SNR) of a MRT/MRC system is upper and lower bounded in [42], based on which the average symbol error rate (SER) and diversity order for a BPSK system are approximately derived. With the restraining assumptions in [42] removed, it is well known that (for white Gaussian noise) the optimal transmit and receive beamformer are given by the principal right and left singular vector of the channel matrix $\mathbf{H}$, respectively; and the MIMO channel is transformed into a single-input single-output link with equivalent channel gain $s_{\max}$, the largest singular value of $\mathbf{H}$. For Rayleigh fading channels, the distribution of $s_{\max}^2$, already derived in [37], is revisited in [13] and expressed in an alternative form – a linear combination of Gamma functions. Based on this expression, the exact system SER is derived for general modulation schemes in [13], which nonetheless still involves complex numerical calculations. The distribution of $s_{\max}^2$ for Ricean fading is obtained in [36]. Unfortunately, results in [13] and [36] don't easily lead one to an insightful understanding of the impact of the system parameters, including the number of transmit and receive antennas M and N, on performance. For example, in [13], the authors make two observations on MIMO MRT/MRC systems through simulations: one is that when $M+N$

keeps fixed, the antennas distribution with $|M-N|$ minimized will provide the lowest SER, while the other is that when $M \times N$ is fixed, a distribution with maximum $M+N$ gives the best performance. But the authors do not provide a rigorous justification for both observations. Some similar observations are also made in [36].

This chapter is organized as follows. In Section 2.2, we will give our model for MIMO MRT/MRC systems. Then we provide our asymptotic analysis for the average SER in Section 2.3 and Section 2.4 respectively, together with some numerical results for illustration purpose. Final conclusion is made in Section 2.5.

2.2 System Model

We assume a narrowband MIMO diversity system with M transmit antennas and N receive antennas, modeled as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{w}_t u + \mathbf{n}, \tag{2.1}$$

where $\mathbf{w}_t \in \mathbb{C}^{M\times 1}$ is the unit-norm transmit weight vector and u is the transmitted symbol with power P_T, $\mathbf{y} \in \mathbb{C}^{N\times 1}$ is the received vector, $\mathbf{H} \in \mathbb{C}^{N\times M}$ is the channel matrix, and $\mathbf{n} \in \mathbb{C}^{N\times 1}$ is a zero-mean circularly symmetric complex Gaussian noise vector with variance $s_n^2/2$ per real dimension. We define $g_t = P_T / s_n^2$ the average transmit SNR. For illustration purpose, independent and identically distributed Rayleigh fading is considered for $\mathbf{H}$, but our analysis can be readily extended to other fading scenarios when appropriate distributions are available. When multiple MIMO users are involved, their channels are assumed independent.

At the receiver side a weight vector $\mathbf{w}_r \in \mathbb{C}^{N\times 1}$ is applied on $\mathbf{y}$ to obtain a decision statistic for u. For a MIMO MRT/MRC system, $\mathbf{w}_t$ and $\mathbf{w}_r$ are chosen to be the principal

right and left singular vector corresponding to the largest singular value $s_{\max}$ of **H** to maximize the output SNR.

The cumulative distribution function (CDF) of $x = s_{\max}^2$ is given by [37]

$$F_g^{MRT/MRC}(x) = \frac{\left|\mathbf{\Psi}_c(x)\right|}{\Pi_{k=1}^{s}\Gamma(t-k+1)\Gamma(s-k+1)}, x \in (0,+\infty), \tag{2.2}$$

where $s = \min(M,N)$, $t = \max(M,N)$, and $\mathbf{\Psi}_c(x)$ is an $s \times s$ Hankel matrix function with the (i,j)th entry given by $\{\mathbf{\Psi}_c(x)\}_{i,j} = g(t-s+i+j-1,x)$, for $i, j = 1,2,...,s$. Here $g(a,b)$ is the incomplete Gamma function defined as $g(a,b) = \int_0^b e^{-t}t^{a-1}dt$, and $\Gamma(a)$ is the Gamma function defined as $\Gamma(a) = g(a,+\infty)$. The probability distribution function (PDF) of x can be derived as

$$f_g^{MRT/MRC}(x) = F(x)tr(\mathbf{\Psi}_c^{-1}(x)\mathbf{\Phi}_c(x)), x \in (0,+\infty), \tag{2.3}$$

where $\mathbf{\Phi}_c(x)$ is an $s \times s$ matrix whose (i,j)th entry is given by $\{\mathbf{\Phi}_c(x)\}_{i,j} = x^{t-s+i+j-2}e^{-x}$.

It is seen that the distribution functions for MIMO MRT/MRC systems are quite involved, which makes relevant math expressions (such as SER) quite complex. In what follows, we will turn to asymptotic analysis to obtain some succinct and insightful results. At the same time, we will compare the MIMO MRT/MRC system with two other popular MIMO diversity schemes. One such scheme is employing space-time block coding at the transmitter and MRC at the receiver (STBC/MRC). In this case, the transmitter doesn't need channel state information (CSI) and the transmit power is equally allocated among the transmit antennas. Another interesting case is to conduct selection combing at both the transmitter and receiver (SC/SC), i.e., one transmit antenna and one receive antenna are

selected so that the resultant channel gain is maximized. This scheme requires less feedback than the MIMO MRT/MRC.

In the remainder part of this chapter, we write $g(x) \sim f(x)$ if $\lim_{\substack{x\to\infty \\ or\ x\to 0}} \frac{g(x)}{f(x)} = 1$.

2.3 Asymptotic Average Symbol Error Rate

In this section, we will derive a succinct expression for average SER at high SNR. The conditional SER for lattice-based modulations can be represented by the Gaussian tail *Q*-function as $P_s(\mathbf{H}) = M_n Q(\sqrt{kg_t x})$, where M_n is the number of the nearest neighboring constellation points, and k is a positive fixed constant determined by the modulation and coding schemes [55]. At high transmit SNR g_t, the system performance will be dominated by the low-probability event that x becomes small [63]. Therefore, only the behavior of $f_g^{MRT/MRC}(x)$ at $x \to 0^+$ determines high transmit SNR performance. To find the asymptotic expression for $P_s = E\{P_s(\mathbf{H})\}$ at high g_t, we need the following result for the behavior of $f_g^{MRT/MRC}(x)$ at the origin.

***Lemma 1*:** $f_g^{MRT/MRC}(x) \sim \frac{MN\prod_{k=0}^{s-1} k!}{\prod_{k=0}^{s-1}(t+k)!} x^{MN-1}$, as $x \to 0^+$.

Proof: See Appendix A.

With Lemma 1, we establish the following result for the asymptotic average SER for MIMO MRT/MRC systems following Proposition I in [63].

Proposition 1: For MIMO MRT/MRC systems, the asymptotic average SER is given by

$$P_s = \frac{2^q M_n a^{(MRT/MRC)} \Gamma(q^{(MRT/MRC)} + \frac{3}{2})}{\sqrt{p}\,(q^{(MRT/MRC)} + 1)} \times (kg_t)^{-(q^{(MRT/MRC)}+1)} + o(g_t^{-(q^{(MRT/MRC)}+1)}). \quad (2.4)$$

where

$$a^{(MRT/MRC)} = \frac{MN \prod_{k=0}^{s-1} k!}{\prod_{k=0}^{s-1} (t+k)!}, q^{(MRT/MRC)} = MN - 1. \quad (2.5)$$

The validity of (2.4) is demonstrated in Figure 2.1 for uncoded BPSK systems. Based on (2.4), one readily concludes that the optimal diversity order for MIMO diversity systems is $M \times N$. Therefore, if we keep $M + N$ fixed (a measure of system cost), even distribution of

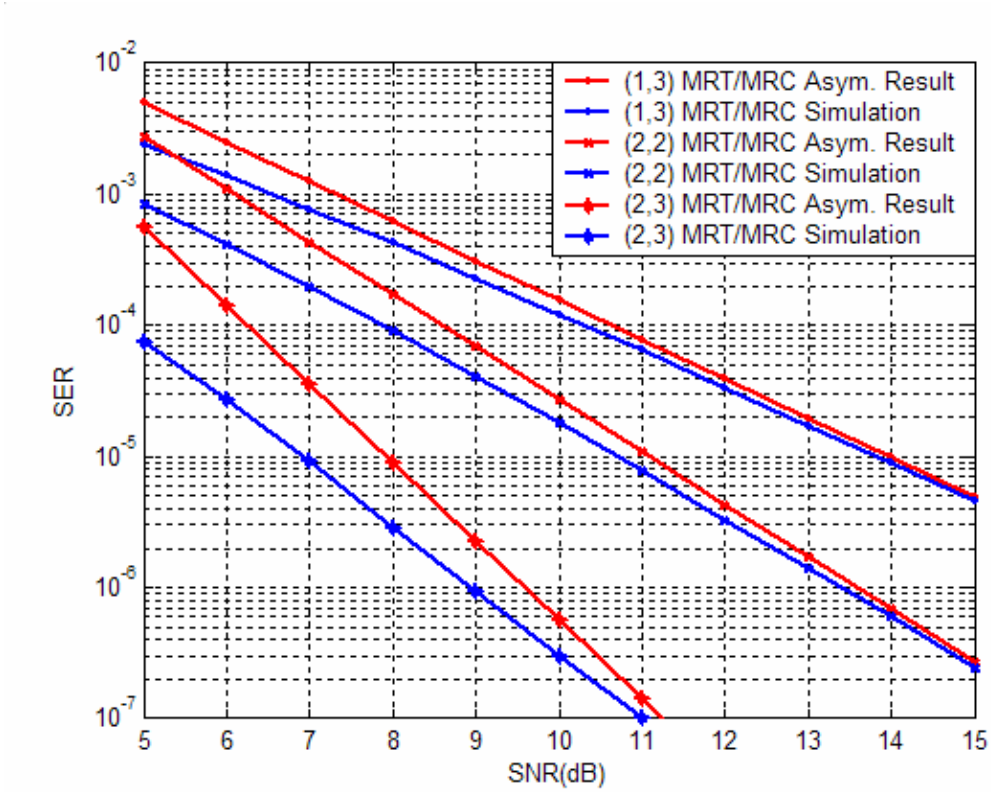

Figure 2.1 Comparision between asymptotic and simulated results for BPSK under different antenna configurations

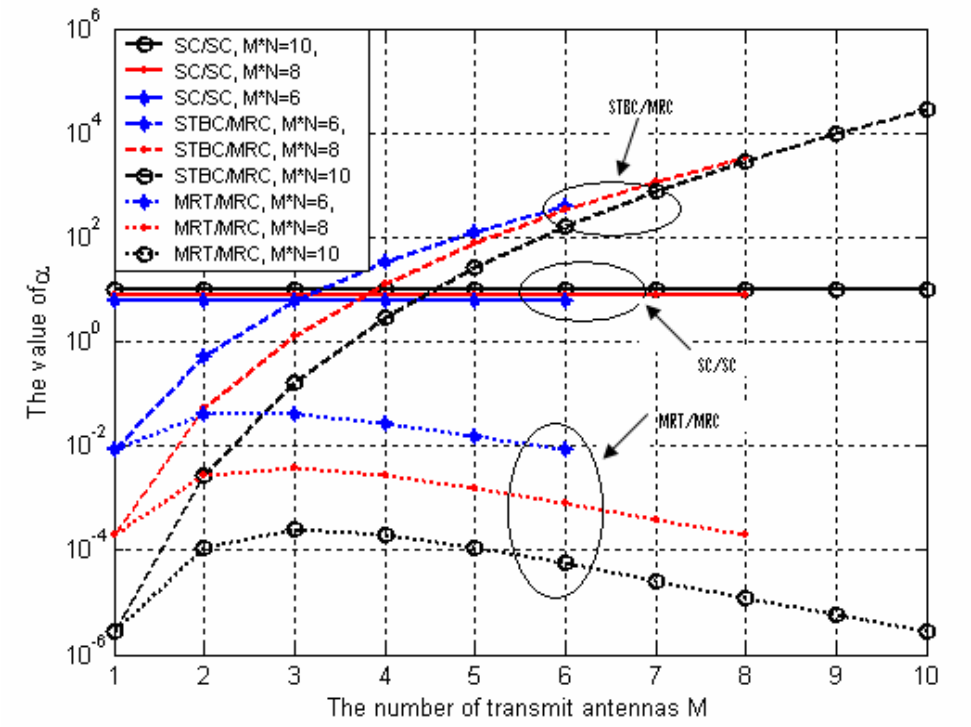

Figure 2.2 The value of a varies with number of transmit antennas under the same $M \times N$

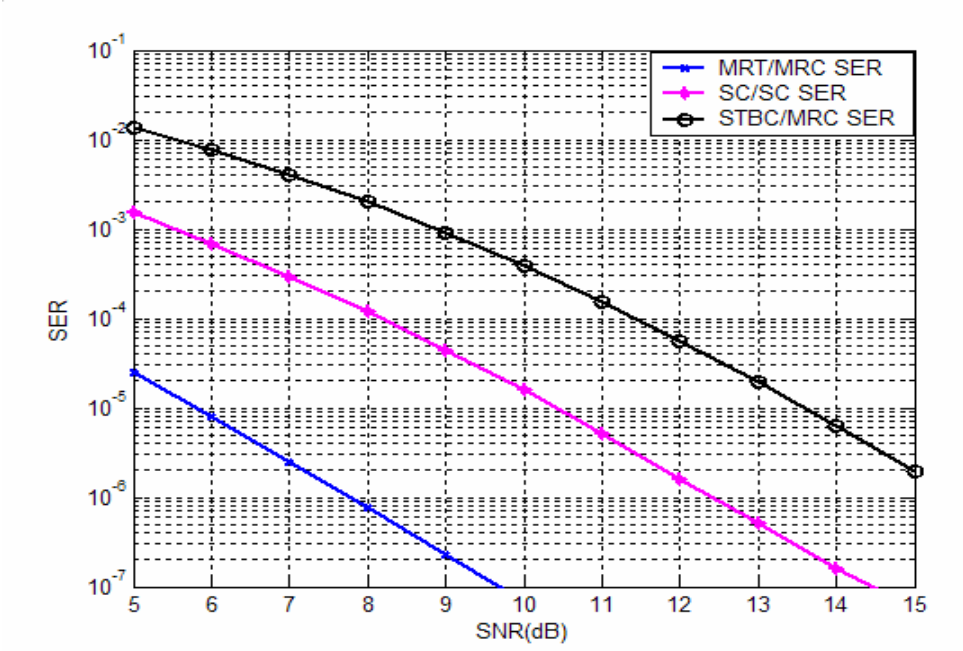

Figure 2.3 Symbol error rate of the three spatial diversity schemes under BPSK

the number of transmit and receive antennas (more precisely a smallest $|M-N|$) maximizes $M \times N$, thus minimizing the system SER at high SNR. On the other hand, when comparing two MIMO MRT/MRC systems with the same diversity order $M \times N$, the one with smaller $a^{(MRT/MRC)}$ yields larger coding gain and thus smaller SER (in this case, $q^{(MRT/MRC)}$ is a constant). We can conclude that in this scenario, the sum of transmit and receive antennas should be made as large as possible, with the optimum achieved at $s=1$ and $t=M \times N$. This conclusion is based on the following result regarding $a^{(MRT/MRC)}$ as a function of M and N (or equivalently of s and t).

***Lemma 2*:** Given four positive integers s_1, t_1, s_2, t_2, assume $s_1 \times t_1 = s_2 \times t_2$, $s_1 < t_1$, $s_2 < t_2$, and $s_1 + t_1 > s_2 + t_2$, then $a^{(MRT/MRC)}(s_1,t_1) < a^{(MRT/MRC)}(s_2,t_2)$.

Proof: see Appendix A.

From the asymptotic SER expression in (2.4), we have verified the two observations made in [13] rigorously at high SNR. In what follows, we will compute the corresponding parameters for the coding gain and diversity order for MIMO STBC/MRC and SC/SC systems (whose asymptotic average SERs assume the same forms as (2.4)).

STBC/MRC

Without loss of generality, we assume that the adopted space-time block coding scheme achieves the full rate and the transmit power is equally allocated among the transmit antennas. In this case, the normalized effective link SNR for a generic user is given by $g = \frac{1}{M}\sum_{i=1}^{N}\sum_{j=1}^{M}\left|h_{i,j}\right|^2$, whose PDF admits:

$$f_g^{STBC/MRC}(x) = \frac{M^{MN}}{(MN-1)!} x^{MN-1} e^{-Mx}, x \geq 0. \quad (2.6)$$

Therefore the corresponding parameters for the coding gain and diversity order for MIMO STBC/MRC systems can be obtained following a similar approach as above as

$$a^{(STBC/MRC)} = \frac{M^{MN}}{(MN-1)!}, q^{(STBC/MRC)} = MN - 1. \quad (2.7)$$

SC/SC

In this spatial diversity scheme, both the user and the base station choose one optimal antenna such that the resultant channel gain is maximized. Thus the normalized effective link SNR at the receiver is $g = \max_{1 \leq i \leq N, 1 \leq j \leq M} \left(\left| h_{i,j} \right|^2 \right)$, whose PDF can be easily obtained as

$$f_g^{SC/SC}(x) = MNe^{-x}(1-e^{-x})^{MN-1}, x \geq 0. \quad (2.8)$$

We can obtain the corresponding parameters for the coding gain and diversity order for MIMO SC/SC systems as

$$a^{(SC/SC)} = MN, q^{(SC/SC)} = MN - 1. \quad (2.9)$$

Comparing (2.5), (2.7) and (2.9) we can see that all these MIMO diversity schemes achieve the same diversity order. Nonetheless, their error performances could still be dramatically different owing to different coding gains, as exhibited in Figure 2.2. For example, when $M = 6$ and $N = 1$, our formulas predict a SNR gap of 4.7 dB between MRT/MRC ($a^{(MRT/MRC)} = 1/120$) and SC/SC ($a^{(SC/SC)} = 6$), and 7.8 dB between MRT/MRC and STBC/MRC ($a^{(STBC/MRC)} = 388.8$) for uncoded BPSK systems at high SNR, which agree

well with simulation results (see Figure 2.3 at SER 10^{-5}). It is also observed that for the same diversity order, the performance of STBC worsens with the increase of transmit antennas.

2.4 Summary

In this chapter, through the analysis of the distribution of the squared largest singular value of a complex Gaussian matrix at the origin, we obtain the asymptotic error performance in the single-user scenario at high transmit SNR. Our results are rigorous and succinct, which provide a performance reference for MIMO diversity systems and facilitate various tradeoff studies in terms of system parameters and designs.

Chapter 3

Joint Antenna Selection and Link Adaptation for MIMO Systems

3.1 Background

The use of multiple antennas at both the transmitter and receiver side, so as to form a multiple-input multiple-output (MIMO) antenna system, is an emerging technology that makes building high data rate wireless networks a reality [19][58]. Transmitting independent data streams simultaneously from different antennas through spatial multiplexing (see, e.g., [1]) effectively realizes the high spectral efficiency promised by MIMO systems, but leaves the transmitted data unprotected from random channel impairment. Therefore, it is often desirable to consider link adaptation, such as rate adaptation and power control to improve the system performance and guarantee certain quality of service [8][76][14][49].

One of the drawbacks with an MIMO system is the increased complexity and hardware cost due to the expensive RF chains required by each active antenna. It is of increasing research interest recently to find a good antenna selection scheme that can significantly reduce such cost while incurring little performance loss. Generally, there are two goals for antenna subset selection in MIMO systems: one aims to maximize the channel capacity [26][46], the other aims to minimize the bit error rate for spatial multiplexing systems when some practical signaling schemes are used [24][25][32][47].

It is interesting to notice that link adaptation and antenna selection problems are actually coupled for MIMO systems, when practical signal processing techniques such as zero-forcing successive interference cancellation (ZF-SIC) (as used in V-BLAST) are employed at the receiver for data decoupling and detection. This is because the decoupled

subchannel gains (post-detection signal-to-noise ratio (SNR)) are determined by the active antenna subset, while some weak subchannels are naturally dropped during link adaptation process. Motivated by this fact, we propose a joint antenna subset selection and link adaptation study for MIMO systems.

In a real propagation environment, the capacity of a MIMO system may be lower than what is predicted with rich scattering assumption due to fading correlation [51][10]. Meanwhile, link adaptation and antenna selection are expected to achieve more gains in correlated MIMO channels due to more prominent subchannel discrepancies. Furthermore, fading correlation information varies much more slowly, hence it is feasible and advantageous to implement antenna selection and link adaptation only based on the correlation information rather than on the instantaneous channel information. The author in [48] also proposed some simplified rules for joint antenna selection and link adaptation based on the channel correlation information, aiming to maximize some lower bounds of the minimum post SNR. Therefore the performance of these rules depends on how tight the lower bounds would be. Furthermore, the exhaustive search entailed there might make these rules still complex in implementation.

In this chapter, we consider the problem of joint antenna selection and link adaptation for an uncoded spatial multiplexing system with a ZF-SIC receiver, for both uncorrelated and correlated MIMO channels. Our goal is to minimize the bit error rate given a throughput and power constraint. We allow all the available resources, including the number of active transmit antennas, symbol constellation size and transmit power dynamically adapted to the channel conditions.

This chapter is organized as follows. In Section 3.2, we introduce the MIMO system model with transmit antenna selection, and formulate the problem of joint antenna subset selection and link adaptation. In Section 3.3, we develop incremental and decremental antenna selection rules with link adaptation for uncorrelated MIMO channels. We also propose a simplified rule based on the expected optimal number of active antennas to further reduce complexity. In Section 3.4, we develop an antenna selection rule with link adaptation for correlated MIMO channels only based on the slowly varying channel correlation information. Simulation results are given and analyzed in Section 3.5. Finally, in Section 3.6, we make conclusions and propose some future work.

3.2 Problem Formulation

3.2.1 MIMO Systems with Transmit Antenna Selection

Without loss of generality, we assume a narrowband MIMO system with total K_t transmit and N_r receive antennas, with the channel between K_t transmit and N_r receive antennas denoted by $\mathbf{H}$. In our study, the antenna selection is only carried out at the transmitter side, and it is easily shown that the best performance is achieved when all receive antennas are active [50]. With N_t out of K_t transmit antennas to be chosen, we denote the selected subset of transmit antennas by p and the channel matrix between the selected N_t transmit antennas and N_r receive antennas by $\mathbf{H}(p)$, whose columns correspond to the selected antennas. The received signals are then given by

$$\mathbf{y} = \mathbf{H}(p)\mathbf{x} + \mathbf{n}, \tag{3.1}$$

where $\mathbf{x} = (x_1, x_2, ..., x_{N_t})^T$ is the transmitted signal vector, $\mathbf{y} = (y_1, y_2, ..., y_{N_r})^T$ is the received signal vector, and $\mathbf{n} = (n_1, n_2, ..., n_{N_t})^T$ is assumed to be i.i.d Gaussian with zero mean and variance of s_n^2. For ease of description, we will drop the index p in (3.1) in the following discussion when no ambiguity incurs. All through this paper we assume $N_r \geq N_t$.

3.2.2 ZF-SIC with QR Decomposition Interpretation

The zero-forcing successive interference cancellation, widely used in MIMO detection, can be simply interpreted by matrix QR decomposition. With $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is a unitary matrix and $\mathbf{R}$ is an upper triangular matrix, we can apply $\mathbf{Q}^H$ to the received vector to obtain $\tilde{\mathbf{y}} = \mathbf{Q}^H\mathbf{y} = \mathbf{Rx} + \tilde{\mathbf{n}}$, detailed as

$$\begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_{N_t} \end{pmatrix} = \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,N_t} \\ 0 & r_{2,2} & \cdots & r_{2,N_t} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & .. & r_{N_t,N_t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_t} \end{pmatrix} + \begin{pmatrix} \tilde{n}_1 \\ \tilde{n}_2 \\ \vdots \\ \tilde{n}_{N_t} \end{pmatrix}, \tag{3.2}$$

from which the transmitted symbols $x_{N_t}, x_{N_t-1}, ..., x_1$ can be detected successively. Assume no error propagation during interference cancellation process[2], it is clear that QR decomposition decomposes an $N_r \times N_t$ MIMO channel matrix $\mathbf{H}$ into N_t subchannels with $|r_{i,i}|$ being the gain for the *ith* subchannel.

3.2.3 Joint Antenna Selection and Link Adaptation

As mentioned in the introduction, the link adaptation problem and the antenna selection problem are often coupled for a MIMO system. Furthermore, it is often beneficial

[2] This assumption is reasonable at sufficiently high SNR regimes and commonly adopted in relevant study to simplify analysis. Our simulation results validate its effectiveness.

to use only a good subset of antennas in MIMO communications to reduce hardware complexity and energy consumption. The induced performance loss is often negligible when judicious antenna selection is made and link adaptation is employed. To this end, we propose to jointly consider the antenna selection and link adaptation for wireless MIMO communications. Antenna selection and link adaptation can be realized either at the transmitter or at the receiver, depending on the availability of channel state information. In the latter case, the receiver will only feed back the selected active antenna subset and corresponding communication modes to the transmitter.

In this chapter we assume QAM modulation for illustration purpose. For square *M*-ary QAM with average power g , the minimum Euclidean distance d is

$$d = \sqrt{\frac{6g}{M-1}}, \tag{3.3}$$

which is also a good approximation for energy efficient "non-square" QAM in a large range of interest [1].

Assume there are N_t active antennas in use. For the ith subchannel with gain $|r_{i,i}|$, the square of the minimum Euclidean distance of the output constellation is given as

$$d_{i,out}^2 = \frac{6|r_{i,i}|^2 g_i}{M_i - 1}, i = 1, 2, ..., N_i, \tag{3.4}$$

where g_i and M_i are the power and constellation size allocated to the *i*th substream. As with many other multi-channel communications, the performance of a spatial multiplexing system is usually limited by the weakest link. Thus the optimization problem can be sensibly formulated as:

$$\max_{(N_t,p,b_i,g_i):\sum_{i=1}^{N_t} b_i=b_T,\sum_{i=1}^{N_t} g_i=g_T} \left\{ \min_{i\in\{1,\ldots N_t\}} d_{i,out}^2 \right\}, \tag{3.5}$$

where $b_i = \log_2 M_i$ is the number of bits allocated to the $i-th$ subchannel, b_T and g_T are the total throughput and power constraints imposed on the system.

In (3.5) we want to find an optimal antenna subset together with its optimal bit and power allocation, subject to the total throughput and power constraints. The number of active antennas N_t can also be an optimization parameter, thus further complicating the problem. To our best knowledge, the global optimal solution is open and often a thorough search has to be resorted to, which is typically infeasible for practical implementations. Therefore we take some effective steps to decouple the original problem into some suboptimal ones, which will be shown to yield excellent performance nonetheless.

First, assuming the set of active antennas and associated bit allocation are given, as the system performance is limited by the worst subchannel, to maximize the aggregate performance we would like to allocate power so as to achieve the same output minimum Euclidean distances for all subchannels, i.e., $d_{1,out}^2 = \ldots = d_{N_t,out}^2 = d_{(e)}^2$, given as

$$d_{(e)}^2 = \frac{g_T}{\sum_{j=1}^{N_t} (\left|r_{j,j}\right|^2 \times \frac{6}{M_j - 1})^{-1}} = \frac{6g_T}{\sum_{j=1}^{N_t} \left|r_{j,j}\right|^{-2} \times (M_j - 1)}. \tag{3.6}$$

Thus our optimization goal is simplified as

$$\min \sum_{j=1}^{N_t} \left|r_{j,j}\right|^{-2} \times (M_j - 1) = \min \left\langle \mathbf{g}(N_t), \mathbf{m}(N_t) \right\rangle, 1 \le N_t \le K_t, \tag{3.7}$$

$$\text{subject to } b_T = \log_2(M_1) + \log_2(M_2) + \ldots + \log_2(M_{N_t}),$$

where $\mathbf{g}(N_t) = (|r_{1,1}|^{-2}, |r_{2,2}|^{-2}, ..., |r_{N_t,N_t}|^{-2})^T$ is named the antenna gain vector, while $\mathbf{m}(N_t) = (M_1 - 1, M_2 - 1, ..., M_{N_t} - 1)^T$ is named the bit allocation vector, and $\langle \mathbf{g} \rangle$ denotes the inner product between them. Our target is to find an optimal pair of $(\mathbf{g}(N_t), \mathbf{m}(N_t))$ for a given N_t, and further choose the best pair among $1 \le N_t \le K_t$, when the number of active antennas is not given beforehand.

Given N_t, the optimal pair of $(\mathbf{g}(N_t), \mathbf{m}(N_t))$ can be found through a thorough search in principle, which is still not an easy task when N_t and K_t are large. We further decouple the antenna selection and bit allocation problems by exploiting the discrete and finite-alphabet nature of the bit allocation vector $\mathbf{m}(N_t)$.When the total throughput and the modulation set are given, the possible choices of the bit allocation vector can be determined in advance by a lookup table. Furthermore, by Lemma 3 given in the appendix B, in order to minimize (3.7), only one permutation (decreasing order) of the elements in the bit allocation vector needs to be considered for each possible combination. With this decoupling, the optimization problem is finally approximated as an antenna selection problem to find a suitable $\mathbf{g}(N_t)$ followed by table lookup to find a matching $\mathbf{m}(N_t)$. Some simple recursive algorithms are proposed in the next section to avoid exhaustive search while incurring little performance degradation.

Finally, note that our proposed algorithms can be readily extended to other modulation schemes. For example, when PSK is employed, the minimum Euclidean distance

of M-ary PSK with power g is given by $d = \sqrt{2\sin^2(\frac{p}{M})g}$. Correspondingly, (3.7) becomes $\min \sum_{j=1}^{N_t} \left| r_{j,j} \right|^{-2} \times \csc^2(\frac{p}{M_j})$ subject to the same constraint.

3.3 Joint Antenna Selection and Link Adaptation for uncorrelated MIMO Channels

We first consider the uncorrelated MIMO channels where the channel matrix **H** can be modeled with i.i.d. complex Gaussian entries. Two basic recursive algorithms are proposed to choose the desired antenna gain vector $\mathbf{g}(N_t)$: incremental selection means the "desired" antennas are recursively added to an initially empty active antenna set while decremental selection means "undesired" antennas are recursively removed from an initially full antenna set[3]. When $N_t << K_t$, we can use the incremental selection rule described in Section 3.3.1, while we can use the decremental rule in Section 3.3.2, when N_t is close to K_t. In a general link adaptation problem where N_t is unknown in advance, we can search over all possible $1 \le N_t \le K_t$ to find the optimal one. To reduce complexity, we propose an adaptive selection rule based on estimation of N_t in Section 3.3.3.

3.3.1 Incremental Selection Rule with Link Adaptation

Intuitively, we want $\left| r_{1,1} \right|$, $\left| r_{2,2} \right|$, …, $\left| r_{N_t,N_t} \right|$ as large as possible. Our incremental recursive rule works as follows: starting from a column of **H** ($N_r \times K_t$) which results in maximum $\left| r_{1,1} \right|$ (corresponding to the largest vector norm), we successively choose from the remaining columns of **H** such that the next subchannel gain is maximized. The subchannel

[3] Same notations are used in [26] and other literature with different problem settings.

gain of the newly added antenna can be obtained in a closed-form solution, which is described by the following lemma.

Lemma4.a Assume the QR decomposition of a matrix $\mathbf{H}^{(k)}$ with k independent columns is $\mathbf{H}^{(k)} = \mathbf{Q}(k)\mathbf{R}(k)$. Then for the enhanced matrix $\mathbf{H}^{(k+1)} = \left[\mathbf{H}^{(k)} \quad \mathbf{h}\right]$ with QR decomposition $\mathbf{H}^{(k+1)} = \mathbf{Q}(k+1)\mathbf{R}(k+1)$, the first k diagonal elements of $\mathbf{R}(k+1)$ keep the same with those of $\mathbf{R}(k)$, while the $(k+1)th$ one is given by $\sqrt{\mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{Q}(k)\mathbf{Q}(k)^H\mathbf{h}}$; similarly, the first k column vectors of $\mathbf{Q}(k+1)$ keep the same with those of $\mathbf{Q}(k)$, while the $(k+1)th$ one is given by $\mathbf{Q}(:,k+1) = \mathbf{h} - \sum_{l=1}^{k}\mathbf{Q}(:,l)^H\mathbf{h}\mathbf{Q}(:,l)$.

Proof: see the appendix B.

Based on Lemma 4.a, assume in the kth step, $\mathbf{H}^{(k)}$ stores the k selected columns of $\mathbf{H}$ and the QR decomposition of $\mathbf{H}^{(k)}$ is $\mathbf{Q}(k)\mathbf{R}(k)$, then in the $(k+1)th$ step, we choose the column vector $\mathbf{h}$ from $\mathbf{H} \setminus \mathbf{H}^{(k)}$ (which represents the remaining columns of $\mathbf{H}$) in such a way that $r_{k+1,k+1} = \sqrt{\mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{Q}(k)\mathbf{Q}^H(k)\mathbf{h}}$ is maximized. Furthermore, it can also be shown as follows that the successively generated antenna gains are already ordered.

Lemma4.b In the above incremental selection rule for uncorrelated MIMO, $|r_{1,1}| \geq |r_{2,2}| \geq .. \geq |r_{k,k}| \geq ... \geq |r_{N_t,N_t}|$.

Proof: see the appendix B.

Lemma 4.b shows that the elements in the selected antenna gain vector $\mathbf{g}(N_t) = (|r_{1,1}|^{-2}, |r_{2,2}|^{-2}, ..., |r_{N_t,N_t}|^{-2})^T$ are already in an increasing order. Thus we only need to

arrange the elements of candidate bit allocation vectors $\mathbf{m}(N_t)$ in a deceasing order in the lookup table according to Lemma 1, which saves storage space and increases the matching speed for (3.7). We further assume $\hat{\mathbf{m}}(N_t)$ is the optimal bit allocation vector that minimizes $\langle \mathbf{g}(N_t), \mathbf{m}(N_t) \rangle$ for a given $\mathbf{g}(N_t)$.

The incremental selection rule with link adaptation for uncorrelated MIMO is summarized in the following table.

Table 3.1 Incremental antenna selection rule with link adaptation for uncorrelated MIMO

Set $\mathbf{I} = \{1,2,3,..,K_t\}$ *and* $\mathbf{p} = \Phi$ *(empty set),* $\mathbf{g} = \Phi$ *(empty set),* $\mathbf{Q} = \Phi$ *(empty set)*

for $i = 1$ *to* K_t

$a_i = \left\| \mathbf{H}(:,i) \right\|^2$;

end

$\mathbf{p}(1) = \arg\max_{1 \le i \le K_t} a_i$, $r_{1,1}^2 = \max_{1 \le i \le K_t} a_i$, $\mathbf{g}(1) = 1/r_{1,1}^2$, $\mathbf{Q}(:,1) = \mathbf{H}(:,\mathbf{p}(1)) / \left\| \mathbf{H}(:,\mathbf{p}(1)) \right\|$;

$\mathbf{I} = \mathbf{I} \setminus \mathbf{p}(1)$;

for $k = 2$ *to* N_t

update $a_i = a_i - \left\| \mathbf{H}(:,i)^H \mathbf{Q}(:,k-1) \right\|^2$, *for all* $i \in \mathbf{I}$;

$\mathbf{p}(k) = \arg\max_{i \in \mathbf{I}} a_i$;

$r_{k,k}^2 = \max_{i \in \mathbf{I}} a_i$;

$\mathbf{g}(k) = 1/r_{k,k}^2$;

$\mathbf{Q}(:,k) = \mathbf{H}(:,\mathbf{p}(k)) - \sum_{l=1}^{k-1} \mathbf{Q}(:,l)^H \mathbf{H}(:,\mathbf{p}(k)\mathbf{Q}(:,l)$;

$\mathbf{Q}(:,k) = \mathbf{Q}(:,k) / \left\| \mathbf{Q}(:,k) \right\|$;

$\mathbf{I} = \mathbf{I} - \mathbf{p}(k)$;

end

assume $\hat{\mathbf{m}}(N_t) = \arg\min_{\mathbf{m}(N_t)} \langle \mathbf{g}(N_t), \mathbf{m}(N_t) \rangle$;

return $\mathbf{p}(1:N_t)$ *and* $\hat{\mathbf{m}}(N_t)$;

Two points are noteworthy for the above algorithm. First, due to Lemma 4.a and 4.b, the antenna selection and link adaptation process is significantly expedited. Secondly, due to the recursive nature of the algorithm, searching an optimal N_t for a general link adaptation problem does not mean K_t times of effort (as calculation for N_t+1 is just one step further based on calculation for N_t), but rather the worst-case effort where all K_t transmit antennas must be deployed. Nonetheless, in case nearly all the K_t transmit antennas would be deployed, we provide a decremental selection rule for link adaptation, which is described in the following subsection.

3.3.2 Decremental Selection Rule with Link Adaptation

Table 3.2 Decremental antenna selection rule with link adaptation for uncorrelated MIMO

Set $\mathbf{H}^{(K_t)} = \mathbf{H}$ *;*

Using [30] to find square root of $\left((\mathbf{H}^{(K_t)})^H \mathbf{H}^{(K_t)}\right)^{-1}$*, assume* $\mathbf{P}^{(K_t)} = \left((\mathbf{H}^{(K_t)})^H \mathbf{H}^{(K_t)}\right)^{-1/2}$*;*

Assume $\hat{l} = \arg\min_{1\le l\le K_t} \left\|\mathbf{P}^{(K_t)}(l,:)\right\|$ ($\left\|\mathbf{P}(l,:)\right\|$ means the length of the *lth* row), *discard the* $\hat{l}$*th column of* $\mathbf{H}^{(K_t)}$*, assume the deflated matrix to be* $\mathbf{H}^{(K_t-1)}$ *;*

for $k=1$ *to* $K_t-(N_t+1)$

Using [30] to find square root of $(\mathbf{H}^{(K_t-k)^H}\mathbf{H}^{(K_t-k)})^{-1}$*, based on* $\mathbf{P}^{(K_t-k+1)}$ *;*

assume $\mathbf{P}^{(K_t-k)} = (\mathbf{H}^{(K_t-k)^H}\mathbf{H}^{(K_t-k)})^{-1/2}$ *;*

Discard the $\hat{l}$*th column of* $\mathbf{H}^{(K_t-k)}$*, where* $\hat{l} = \arg\min_{1\le l\le k} \left\|\mathbf{P}^{(k)}(l,:)\right\|$ *,*

assume the deflated matrix to be $\mathbf{H}^{(K_t-k-1)}$ *;*

end

Compute the sorted antenna gain vector $\mathbf{g}(N_t)$ *for* $\mathbf{H}^{(N_t)}$ *;*

assume $\hat{\mathbf{m}}(N_t) = \arg\min_{\mathbf{m}(N_t)} \langle \mathbf{g}(N_t), \mathbf{m}(N_t)\rangle$*;*

return $\mathbf{H}^{(N_t)}$ *and* $\hat{\mathbf{m}}(N_t)$*;*

The decremental selection rule with link adaptation for uncorrelated MIMO is summarized in Table 3.2. Our proposed decremental selection rule is related to the V-BLAST ordering rule first proposed in [61], which successively chooses the antenna (among those not already chosen) that maximizes post-detection SNR under the assumption of perfect feedback. Accordingly, we can successively discard the antenna (among those not already chosen) that minimizes the post-detection SNR under the assumption of perfect feedback. Usually repeated matrix inversion will be involved during the process of discarding, which may introduce much computation complexity and numerical instability. Thanks to the work in [30], we can avoid computing the inversion of the deflated channel matrix by means of a recursive square-root algorithm.

3.3.3 Simplified Link Adaptation for Uncorrelated Rayleigh MIMO Channels

In a general link adaptation problem where N_t is not fixed in advance, we need to test all possibilities $1 \le N_t \le K_t$ to find the optimal one using either the incremental or decremental selection rule. In this subsection, we propose a simplified selection rule based on the estimation of the optimal number of active transmit antennas to further reduce the complexity. With i.i.d. complex Gaussian channel matrix, $|r_{i,i}|^2$ in (3.7) is a c^2 distributed random variable with $2\times(N_r+1-i)$ degrees of freedom [8], where the probability density function of the c^2 distribution with v degrees of freedom is given as:

$$f(x \mid v) = \frac{x^{(v-2)/2} e^{-x/2}}{2^{v/2}\Gamma(v/2)}, \tag{3.8}$$

where $\Gamma(x)$ is the gamma function defined as $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$.

The expected value of $1/|r_{i,i}|^2$ can be obtained as

$$\int_0^{+\infty} \frac{1}{x} f(x \mid v)dx = \begin{cases} 1/(v-2) & for \quad v > 2 \\ +\infty & for \quad v = 2 \end{cases}. \tag{3.9}$$

Replacing $1/|r_{i,i}|^2$ in (3.7) with their expected values, we have

$$\min E\left(\langle \mathbf{g}(N_t), \mathbf{m}(N_t) \rangle\right) = \min \sum_{j=1}^{N_t} \frac{1}{2(N_r - j)} \times (M_j - 1), 1 \le N_t \le K_t \tag{3.10}$$

subject to

$$b_T = \log_2(M_1) + \log_2(M_2) + ... + \log_2(M_{N_t}) \text{ and } M_1 \ge M_2 \, \mathbf{L} \ge M_{N_t}. \tag{3.11}$$

Therefore, we can estimate the optimal number of active antennas in a pre-processing stage as follows: for all possible bit allocation vectors $\mathbf{m}(N_t)'s$ that satisfy (3.11), find the one that minimizes (3.10), denoted as $\hat{\mathbf{m}}(N_t)$; then find $\hat{N}_t = \arg\min_{1 \le N_t \le K_t} E\langle \mathbf{g}(N_t), \hat{\mathbf{m}}(N_t) \rangle$, which is our estimate of the optimal number of active antennas in i.i.d Rayleigh fading MIMO channels. Thus we can decide to use either the incremental or decremental selection rule for joint antenna selection and link adaptation for different system settings based on the value of $\hat{N}_t$. Furthermore, we can restrict ourselves to search optimal N_t only in the range around $\hat{N}_t$ to further reduce the computational complexity. Simulations results show that searching N_t in the range of $\left[\hat{N}_t - 1, \hat{N}_t + 1\right]$ and storing only *three* bit allocation vectors $\hat{\mathbf{m}}(N_t - 1)$, $\hat{\mathbf{m}}(N_t)$ and $\hat{\mathbf{m}}(N_t + 1)$ in the bit allocation lookup table incur little performance loss (See Section 3.5).

3.4 Joint Antenna Selection and Link Adaptation for Correlated MIMO Channels

3.4.1 Correlated MIMO Channels

In this section, we extend the study of joint antenna selection and link adaptation to correlated MIMO channels. We assume correlation only exists at the transmitter side, as described by the "one-ring" model in [51]. This model is feasible, e.g., for the outdoor macrocell situation where the transmitter at the base station is elevated high above the local scattering environment, while there are sufficient local scatters around the mobile receivers. For an $N_r \times K_t$ MIMO system, the channel can be modeled as $\mathbf{H} = \mathbf{H}_w \mathbf{A}_T^H$ with $\mathbf{A}_T * \mathbf{A}_T^H = \mathbf{R}_T$, where $\mathbf{H}_w$ is an $N_r \times K_t$ matrix containing i.i.d. complex Gaussian random variables and $\mathbf{R}_T$ is a $K_t \times K_t$ Hermitian semi-positive definite matrix representing the covariance matrix for each row of $\mathbf{H}$.

Again, we assume N_t out of K_t antennas are to be selected. As before, the channel matrix between the N_t transmit antennas and N_r receive antennas can be described as $\mathbf{H}(p) = \mathbf{H}_w(p)\mathbf{A}_T^H(p)$, where p contains the indices of the selected antennas, and $\mathbf{A}_T(p)\mathbf{A}_T^H(p) = \mathbf{R}_T(p)$ is the corresponding submatrix of $\mathbf{R}_T$.

We assume uniform linear arrays at both the transmitter and receiver, with antenna spacing Δ_T (relative to the carrier wavelength). We also assume there are L clusters of scatterers in the environment and the angle of departure for the $l-th$ path cluster is Gaussian distributed as $q_l \sim N(\bar{q}_l, s_l^2)$. Then the $(i,j)-th$ entry of the transmit covariance matrix contributed by the $l-th$ scattering cluster can be shown to be approximated as [48][47]-[4]:

$$\left[\mathbf{R}_{T,l}\right]_{i,j} \approx e^{-j2p(i-j)\Delta_T \cos(\bar{q}_l)} e^{-\frac{1}{2}(2p(i-j)\Delta_T \sin(\bar{q}_l)s_l)^2}. \quad (3.12)$$

For a narrowband system, the net correlation matrix can be obtained by summing the covariance matrices contributed by the L clusters weighted by the fraction of power in the corresponding cluster. As a counterpart to (3.1), the received signal in correlated MIMO can be written as:

$$\mathbf{y} = \mathbf{H}_w(p)\mathbf{A}_T^H(p)\mathbf{x} + \mathbf{n}. \quad (3.13)$$

Clearly, joint antenna selection and link adaptation algorithms described in the previous section can be readily applied to correlated MIMO channels and are expected to achieve more substantial gains. However it is noteworthy that for correlated MIMO, the elements of $\mathbf{A}_T^H(p)$ vary much more slowly than those of $\mathbf{H}_w(p)$, which is mainly determined by the local physical parameters, such as antenna spacing and angle spread. Since these parameters are relatively static and can be measured more accurately than instantaneous channel information, antenna selection and link adaptation based on $\mathbf{A}_T^H(p)$ is more attractive than that based on $\mathbf{H}_w(p)\mathbf{A}_T^H(p)$. Targeting on this goal, in the next subsection, we will describe a joint antenna selection and link adaptation algorithm for correlated MIMO only based on the channel correlation information.

3.4.2 Antenna Selection and Link Adaptation Only Based on Channel Correlation Information

By applying QR decomposition successively to the correlation matrix $\mathbf{A}_T^H(p) = \mathbf{Q}_1\mathbf{R}_1$ and $\mathbf{H}_w(p)\mathbf{Q}_1 = \mathbf{Q}_2\mathbf{R}_2$, (3.13) becomes

$$\mathbf{y} = \mathbf{Q}_2\mathbf{R}_2\mathbf{R}_1\mathbf{x} + \mathbf{n}. \quad (3.14)$$

Apply $\mathbf{Q}_2^H$ to the received vector, we have

$$\tilde{\mathbf{y}} = \mathbf{Q}_2^H \mathbf{y} = \mathbf{R}_2 \mathbf{R}_1 \mathbf{x} + \tilde{\mathbf{n}}. \tag{3.15}$$

The optimization goal for correlated MIMO channels is given as (cf. (3.2) and (3.7)):

$$\min \sum_{j=1}^{N_t} |\mathbf{R}_1(j,j)\mathbf{R}_2(j,j)|^{-2} \times (M_j - 1) = \min \langle \mathbf{g}'(N_t), \mathbf{m}(N_t) \rangle \tag{3.16}$$

$$\text{subject to } b_T = \log_2(M_1) + \log_2(M_2) + ... + \log_2(M_{N_t}),$$

with

$$\mathbf{g}'(N_t) = \left[\left| \mathbf{R}_1(1,1)\mathbf{R}_2(1,1) \right|^{-2} \quad ... \quad ... \quad \left| \mathbf{R}_1(N_t,N_t)\mathbf{R}_2(N_t,N_t) \right|^{-2} \right]^T \tag{3.17}$$

and

$$\mathbf{m}(N_t) = (M_1 - 1, M_2 - 1, ..., M_{N_t} - 1)^T$$

the corresponding antenna gain vector and bit allocation vector for correlated MIMO.

Since the distribution of $\mathbf{H}_w \mathbf{Q}_1$ is the same as $\mathbf{H}_w$, $|\mathbf{R}_2(j,j)|^2$ is still c^2 distributed with degree of freedom $2 \times (N_r + 1 - j)$. In order to derive an antenna selection and link adaptation rule only based on $\mathbf{R}_T$, we replace $|\mathbf{R}_2(j,j)|^{-2}$ in (3.17) with their expected values (see (3.9)) to get

$$\bar{\mathbf{g}}(N_t) = \left[\frac{\left| \mathbf{R}_1(1,1) \right|^{-2}}{2(N_r - 1)} \quad ... \quad ... \quad \frac{\left| \mathbf{R}_1(N_t,N_t) \right|^{-2}}{2(N_r - N_t)} \right]^T. \tag{3.18}$$

Hence (3.16) is turned into:

$$\min \sum_{j=1}^{N_t} \frac{|\mathbf{R}_1(j,j)|^{-2}}{2(N_r - j)} \times (M_j - 1) = \min \langle \bar{\mathbf{g}}'(N_t), \mathbf{m}(N_t) \rangle \tag{3.19}$$

subject to. $b_T = \log_2(M_1) + \log_2(M_2) + ... + \log_2(M_{N_t})$.

In recognition of $\mathbf{R}_1^H * \mathbf{R}_1 = \mathbf{A}_T(p)\mathbf{A}_T^H(p) = \mathbf{R}_T(p)$, for correlated MIMO, our goal is to find a submatrix of $\mathbf{R}_T$ whose Cholesky factor will provide a close-to-minimization result of (3.19).

Similar to Section 3.3.1, we decouple the antenna selection and link adaptation problems and present an incremental selection rule as follows. Starting from an empty set, in each step we would like to choose from the remaining components of $\mathbf{R}_T$ such that the next subchannel gain is maximized. This process is expedited by the following lemmas.

Lemma5.a Assume matrix $\mathbf{R}_T^{(k)}$ is Hermitian positive definite with size k, whose Cholesky decomposition is given by $\mathbf{R}_T^{(k)} = \mathbf{R}(k)^H \mathbf{R}(k)$, then for the enhanced matrix $\mathbf{R}_T^{(k+1)} = \begin{bmatrix} \mathbf{R}_T^{(k)} & \mathbf{v} \\ \mathbf{v}^H & 1 \end{bmatrix}$ with Cholesky decomposition $\mathbf{R}_T^{(k+1)} = \mathbf{R}(k+1)^H \mathbf{R}(k+1)$, the first k diagonal elements of $\mathbf{R}(k+1)$ $\{r_{i,i}\}_{i=1}^{k}$ keep the same with those of $\mathbf{R}(k)$, while the $(k+1)th$ one is given by $r_{k+1,k+1} = \sqrt{1 - \mathbf{v}^H \left(\mathbf{R}_T^{(k)}\right)^{-1} \mathbf{v}}$.

Proof: see the appendix B.

Based on Lemma 5.a, assume in step k, there are k selected transmit antennas, and $\mathbf{R}_T^{(k)}$ is the $k \times k$ covariance matrix for those k selected transmit antennas, which is guaranteed to be invertible according to our selection rule, then in step $k+1$, we will choose the antenna whose covariance vector $\mathbf{v}$ will maximize $r_{k+1,k+1} = \sqrt{1 - \mathbf{v}^H \left(\mathbf{R}_T^{(k)}\right)^{-1} \mathbf{v}}$. Note that the diagonal elements of the covariance matrix $\mathbf{R}_T$ are all 1's, thus $r_{1,1}$ is always 1 no matter which antenna is selected first. However, we can determine the first and second active

antennas jointly by means of maximizing $r_{2,2}$, i.e., choose the first two active antennas whose corresponding Cholesky decomposition will result in maximization of $r_{2,2}$. Also note that the condition number of $\mathbf{R}_T$ is high, hence we can set a positive threshold value C_0 in practice to discard those essentially zero-gain subchannels.

Similar to Lemma 4.b, the following lemma facilitates the optimization of (3.19).

Lemma5.b In the above incremental selection rule for correlated MIMO, $r_{1,1} \geq ... r_{k,k} \geq r_{k+1,k+1}$.

Proof: See the appendix B.

Table 3.3 Incremental antenna selection rule with link adaptation for correlated MIMO

Set $\mathbf{I} := \{1,2,3,..,K_t\}$ *and* $\mathbf{p} := \Phi$ *(empty set),* $\mathbf{T} := \Phi$ *(empty set)*

$\mathbf{p}(1:2) = \arg\max_{i,j}\left(1 - \mathbf{v}_1^H \mathbf{v}_1\right), \mathbf{T}(1) := 1, \mathbf{T}(2) := \max_{i,j}\left(1 - \mathbf{v}_1^H \mathbf{v}_1\right)$, *where* $\mathbf{v}_1 = R_T(i, j)$ *and* $i \leq j$ *;*

$\mathbf{I} := \mathbf{I} - \mathbf{p}(1) - \mathbf{p}(2)$ *;*

for $k = 3$ *to* N_t

$\mathbf{A}_{k-1} := \mathbf{R}_T(\mathbf{p}(1:k-1))$ *(Submatrix of* $\mathbf{R}_T$ *designated by* $\mathbf{p}(1:k-1)$ *);*

$\mathbf{T}(k) = \max_{j \in I}\left(1 - \mathbf{v}_{k-1}^H \mathbf{A}_{k-1}^{-1} \mathbf{v}_{k-1}\right)$*;*

$\mathbf{p}(k) = \arg\max_{j \in I}\left(1 - \mathbf{v}_{k-1}^H \mathbf{A}_{k-1}^{-1} \mathbf{v}_{k-1}\right)$, *where* $\mathbf{v}_{k-1} = \mathbf{R}_T(\mathbf{p}(1:k-1), \mathbf{p}(j))$ *(The column vector in* $\mathbf{p}(j)$ *column and* $\mathbf{p}(1:k-1)$ *rows of* $\mathbf{R}_T$ *);*

$\mathbf{I} := \mathbf{I} - \mathbf{p}(k)$ *;*

if $\frac{1}{N_r - k}\left\langle \frac{1}{\mathbf{T}(1:k)}, \hat{\mathbf{m}}(k) \right\rangle > \frac{1}{N_r - (k-1)}\left\langle \frac{1}{\mathbf{T}(1:k-1)}, \hat{\mathbf{m}}(k-1) \right\rangle$

return $\mathbf{p}(1:k-1)$ *and* $\hat{\mathbf{m}}(k-1)$

else if $\mathbf{T}(k) \leq C_0$ *(*C_0 *is a threshold value)*

break;

else if $k = N_t$

return $\mathbf{p}(1:N_t)$ *and* $\hat{\mathbf{m}}(N_t)$

end

end

Lemma5.b shows the elements in $\bar{\mathbf{g}}'(N_t)$ are already in an increasing order. Thus we only need to arrange the elements of candidate bit allocation vectors $\mathbf{m}(N_t)$ in a deceasing order according to Lemma3. In summary, the incremental selection rule with link adaptation for correlated MIMO is described in Table 3.3. For correlated MIMO, a decremental selection rule is usually not necessary, since the ill-conditioning of the channel matrix typically results in much fewer antennas being selected as opposed to the uncorrelated MIMO of the same size. Furthermore, antenna selection and link adaptation modes need to be updated only when the channel covariance matrix changes, which happens far less frequently compared to that based on the instantaneous channel fading.

Similarly, in a general link adaptation problem where N_t is unknown in advance, we can search over all possible $1 \le N_t \le K_t$ to find the optimal one.

3.5 Numerical Results

In this section, we evaluate the performance of our proposed joint antenna selection and link adaptation algorithms for both uncorrelated and correlated MIMO channels through several representative examples. Square QAM modulation is employed in all simulations with 256-QAM the largest constellation to be used.

Example 1. In this example, we demonstrate that both antenna selection gain and link adaptation gain are obtained through our algorithm. Consider a 3×6 MIMO ($N_r = 3, K_t = 6$) with i.i.d Rayleigh fading. The number of active transmit antennas to be chosen is $N_t = 3$ and the target throughput is 12 bits/s/Hz. For performance evaluation, we consider the following three systems: the first one is V-BLAST (i.e., equal power and rate allocation) with

random antenna selection; the second one is V-BLAST with a selected transmit antenna subset obtained through the incremental selection rule; and the last one is our proposed incremental antenna subset selection with link adaptation given in Table I. Link adaptation based on singular value decomposition (SVD) of the channel matrix [75] is also included, which can be viewed as a performance upper bound since the decomposed subchannels are truly interference free. From Figure 3.1, we can see that antenna selection gain is dominant, while link adaptation gain is also significant especially at high SNR (by observing the difference of the slope of the third curve (diversity gain) from that of the second one, and its similarity with that of the upper bound curve).

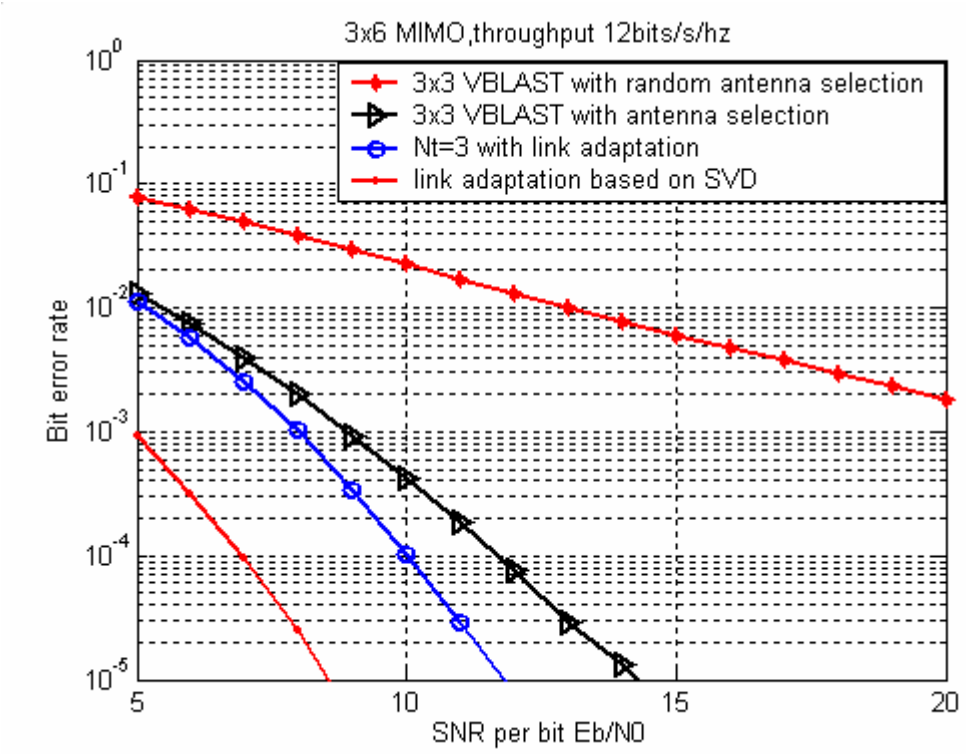

Figure 3.1 Antenna selection gain and link adaptation gain

In the next example, we compare the performances of the incremental and decremental selection rules for a general link adaptation problem where the number of active antennas is not given beforehand, based on both full-size lookup tables and reduced-size lookup tables (see Section 3.3.3).

Example 2. Here we consider a 6×6 MIMO with uncorrelated Rayleigh fading, and the target throughput is 12 bits/s/Hz.

The full-size lookup table is shown in (3.20) with each row representing the bit allocation for a certain number of active antennas. By Lemma 1 and Lemma 2.b, only the decreasing order of each possible combination is listed.

$$Bits\_table\_full = \begin{bmatrix} 8 & 4 & 0 & 0 & 0 & 0 \\ 6 & 6 & 0 & 0 & 0 & 0 \\ 8 & 2 & 2 & 0 & 0 & 0 \\ 6 & 4 & 2 & 0 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 & 0 \\ 6 & 2 & 2 & 2 & 0 & 0 \\ 4 & 4 & 2 & 2 & 0 & 0 \\ 4 & 2 & 2 & 2 & 2 & 0 \\ 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix}. \tag{3.20}$$

From (3.10), the estimated optimal number of active antennas is $\hat{N}_t = 4$, so we only need to store the *optimal* bit allocation vectors $\hat{\mathbf{m}}(N_t)$ for $N_t = \{3,4,5\}$ active antennas:

$$Bits\_table\_reduced = \begin{bmatrix} 4 & 4 & 4 & 0 & 0 & 0 \\ 4 & 4 & 2 & 2 & 0 & 0 \\ 4 & 2 & 2 & 2 & 2 & 0 \end{bmatrix}, \tag{3.21}$$

which incurs almost no performance loss as shown in Figure 3.2. It is also shown that the incremental and decremental selection rules achieve almost the same performance and

approach the SVD upper bound quite closely. (Note that the four curves for joint antenna selection and link adaptation are almost indistinguishable in Figure 3.2)

To verify the effectiveness of our approach in decoupling the antenna selection and link adaptation problems, exhaustive search (among all the possible combinations of the transmit antennas) is conducted to find the optimal pair of $(\mathbf{g}(N_t), \mathbf{m}(N_t))$ for (3.7). As observed in Figure 3.2, our algorithms incur negligible performance degradation.

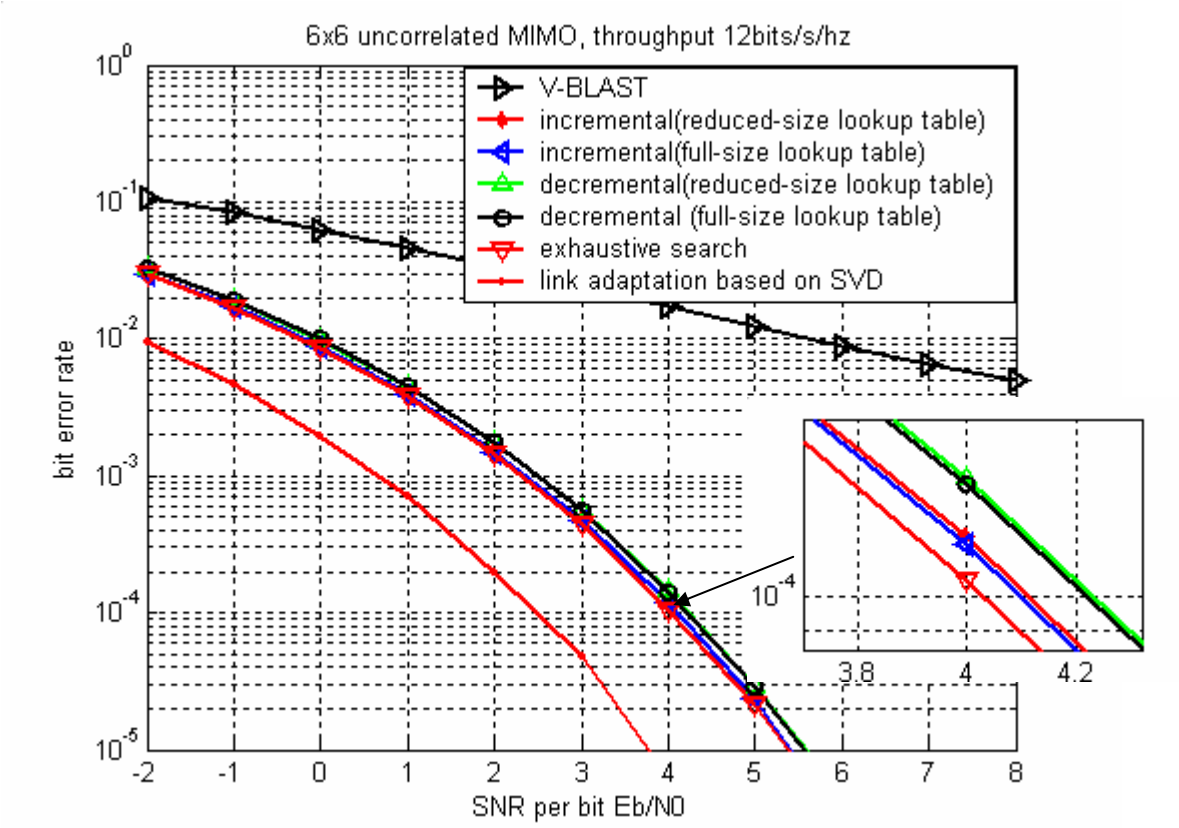

Figure 3.2 Performance comparisions of the proposed algorithms in 6x6 MIMO with throughput 12bits/s/hz

Example 3. This example demonstrates the application of our proposed algorithms on a large MIMO system, emphasizing the demand of complexity reduction. The MIMO system considered is of size 16×16, and the target throughput is 32 bits/s/Hz.

Clearly with this system, any exhaustive search will lead to tremendous computational complexity. Our proposed algorithms exhibit their simplicity advantage while closely approaching the SVD performance upper bound, as shown in Figure 3.3. By using the reduced-size lookup table, the computation complexity is significantly reduced further. From (3.10) the estimated optimal number of active antennas is $\hat{N}_t = 12$, hence in the lookup table, we only need to store the optimal bits pattern for $N_t = \{11,12,13\}$, which is shown below:

$$Bits\_table\_reduced = \begin{bmatrix} 4 & 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 4 & 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 0 & 0 & 0 \\ 4 & 4 & 4 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 0 & 0 & 0 \end{bmatrix}. \quad (3.22)$$

We do not provide the full size lookup table here because of its large size. From Figure 3.3, we can see that little performance loss is incurred when using the reduced-size lookup table. Note that even though matrix inversion is mostly eliminated in the decremental antenna selection rule by means of square root algorithm, it is still inevitable in the initial step, which incurs substantial computational complexity. Furthermore, while the antenna gain vector **g** is readily obtained through Lemma 4.a in the incremental antenna selection, explicit QR decomposition is required for the decremental selection. One can check the computational complexity in antenna selection for incremental method is $O\left(K_t N_r N_t\right)$, while for decremental method, the computation complexity is dominated by the initial matrix inversion step, which is $O\left(K_t N_r^2 + K_t^2 N_r\right)$. Therefore for MIMO systems of small to medium size (such as in Example 2), the incremental rule is preferred, while decremental rule is favored only for very large MIMO systems where the estimated optimal number $\hat{N}_t$ is very close to K_t (due to reduced complexity in link adaptation).

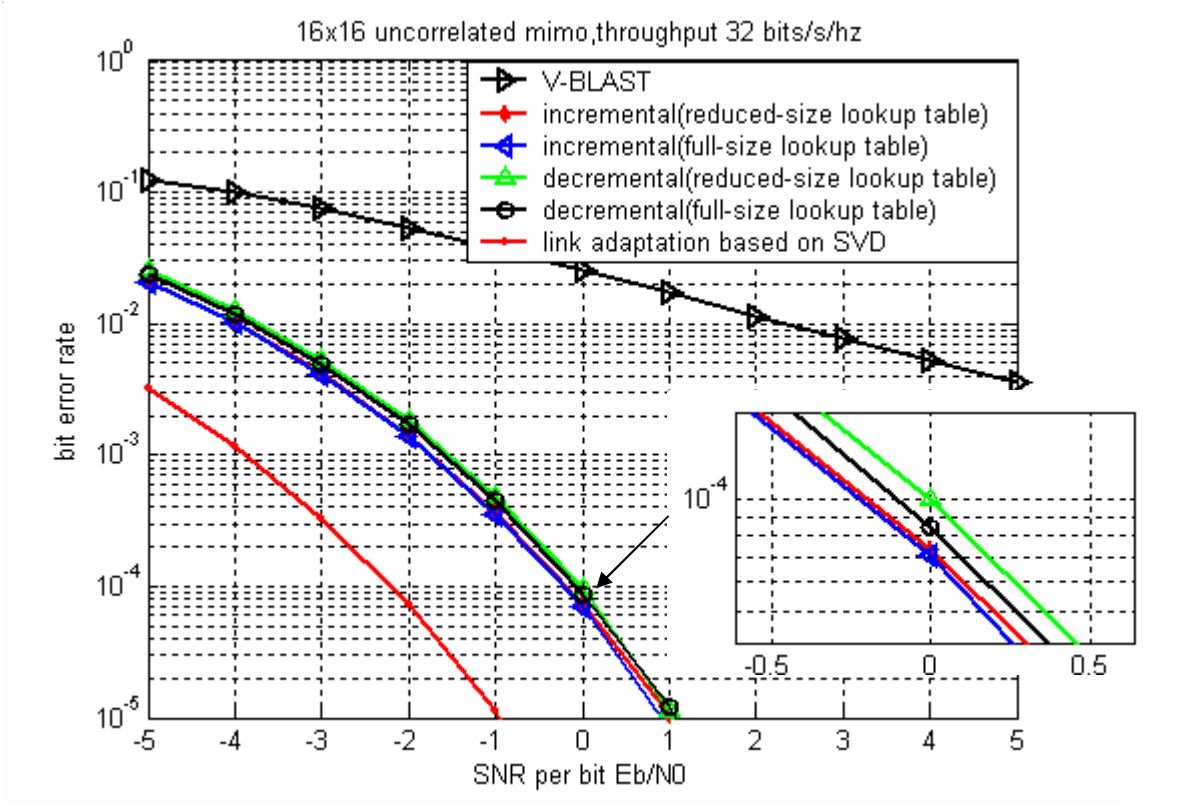

Figure 3.3 Performance comparisons of the proposed algorithms in 16x16 MIMO with throughput 32bits/s/hz

Example 4. In the above examples, we assume perfect CSI is available for antenna selection and link adaptation. In actual vehicular based communications, the channels may vary too fast to allow timely feedback. In this example, we re-evaluate our algorithms in fast fading channels with limited feedback, and explore the long range prediction (LRP) technique [29] as a remedy. LRP is essentially a linear prediction method based on autoregressive modeling. With this technique, one can measure and feedback the time-varying CSI to the transmitter at a much lower rate than the data rate. The transmitter will make compensations through prediction and interpolation, and then determine the active antenna set and modulation modes based on the predicted and interpolated CSI. The reader is referred to [29] for a detailed description of this technique.

Figure 3.4 shows the performance losses of our joint antenna selection and link adaptation algorithms with different feedback delays. Clearly there is a tradeoff between feedback delays and dedicated feedback channel bandwidth. For the LRP technique, the longer the feedback delays, the larger the prediction steps should be taken. In our simulation, we assume a Rayleigh fading channel with Jakes' model with the Doppler shift of f_d (thus the coherence time is $t_0 \approx 1/f_d$ seconds). The channel sampling rate is $8f_d$, while the date rate is $640f_d$, so the channel is measured and fed back once every 80 symbols. The prediction order of LRP is set as 50. It is observed that the proposed algorithms, in conjunction with the LRP technique, have a fairly graceful degradation in performance with increase of feedback delays.

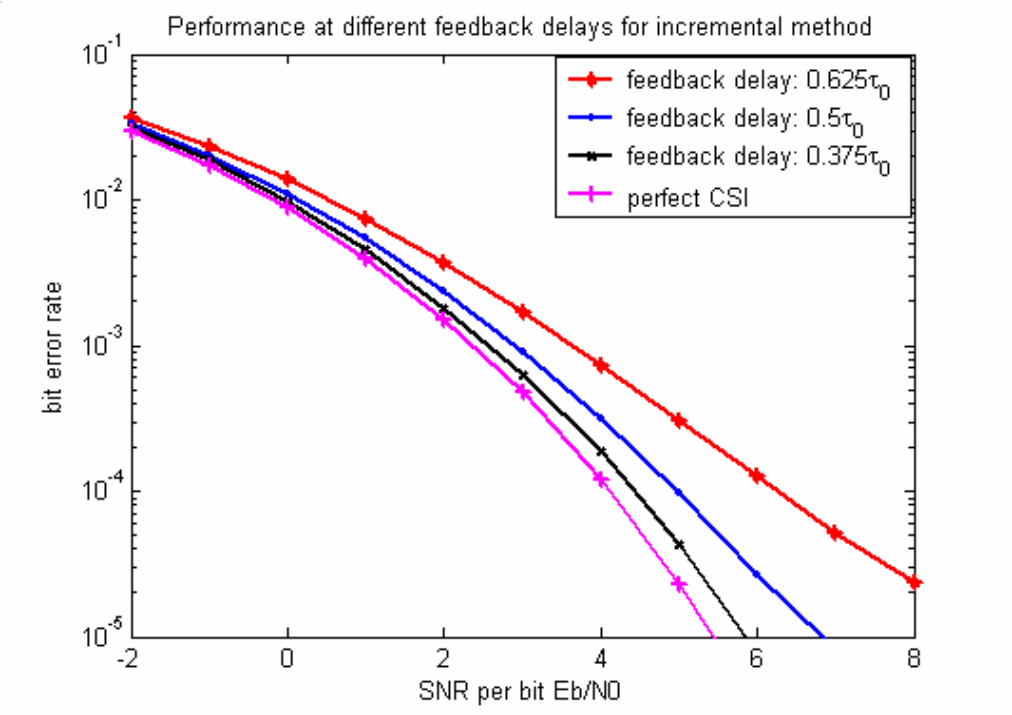

Figure 3.4-a Performance with different feedback delays for incremental methods

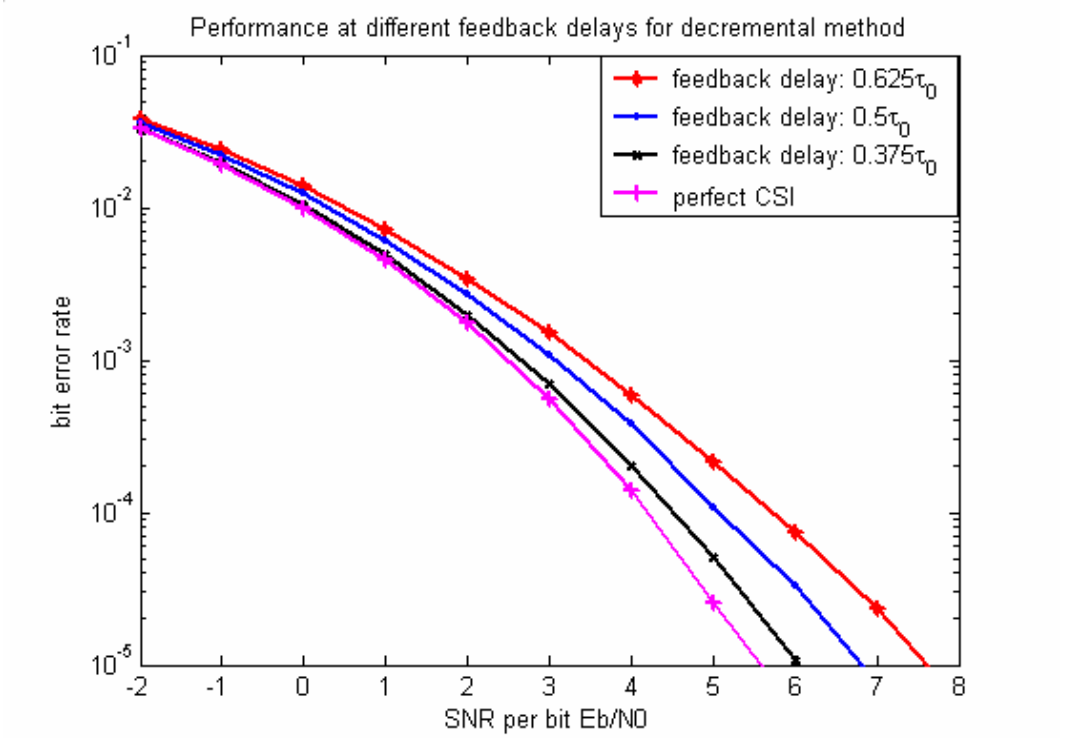

Figure 3.4-b Performance with different feedback delays for decremental methods

Example5. Finally, this numerical example demonstrates the performance of link adaptation only based on correlation information, compared with that based on the full channel information. Consider a 6×6 correlated MIMO with correlation matrix generated as in (3.12), and the target throughput is 12 bits/s/Hz. We consider three correlated fading scenarios as listed in Table 3.4 with an increasing order of fading correlation, all assuming there is only one transmit cluster in the communication environments.

Table 3.4 Fading correlation scenarios

Scenario 1.	$\bar{q} = \frac{p}{6}, s = \frac{p}{10}$
Scenario 2.	$\bar{q} = \frac{p}{6}, s = \frac{p}{15}$
Scenario 3.	$\bar{q} = \frac{p}{10}, s = \frac{p}{30}$

In the Figure 3.5a – Figure 3.7a, we compare the BER performance among link adaptation only based on the correlation information, link adaptation based on the full channel information and the conventional V-BLAST. From the simulation results, we can see the performance of the traditional V-BLAST degrades significantly in the correlated MIMO channels. On the other hand, antenna selection and link adaptation achieves more substantial gains for correlated MIMO than for uncorrelated MIMO, and the performance gap between link adaptation only based on channel correlated information and link adaptation based on the full channel information decreases as the degree of correlation increases.

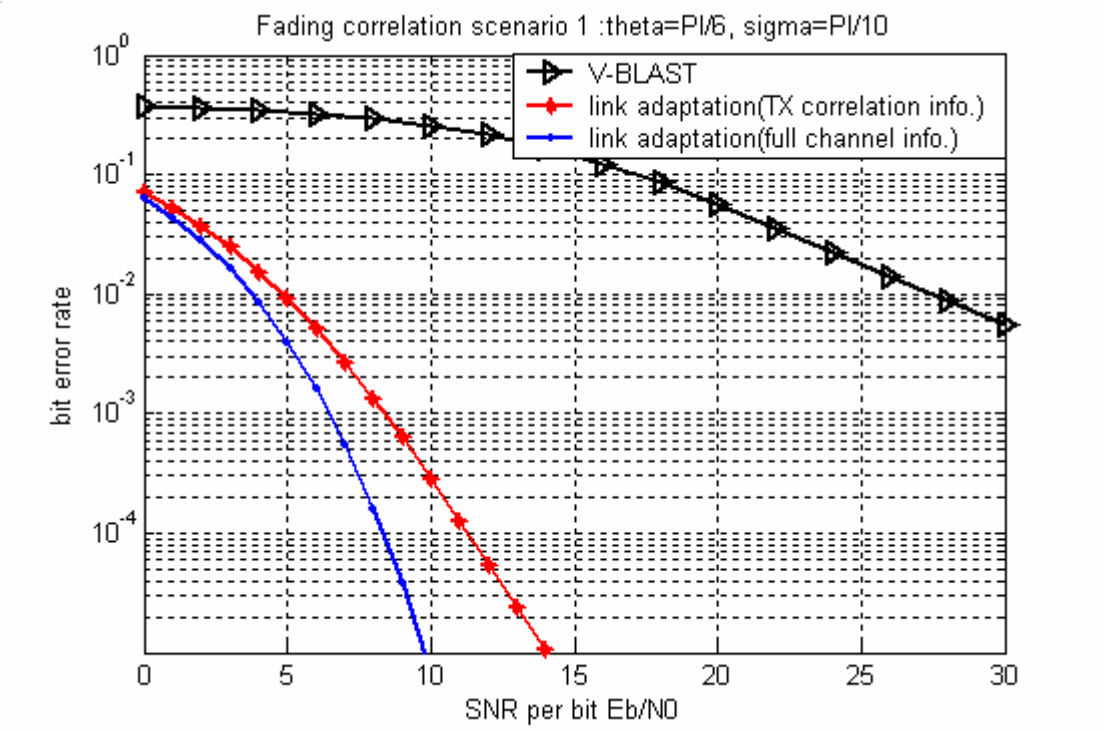

Figure 3.5-a Joint antenna selection and link adaptation for fading correlation scenario 1

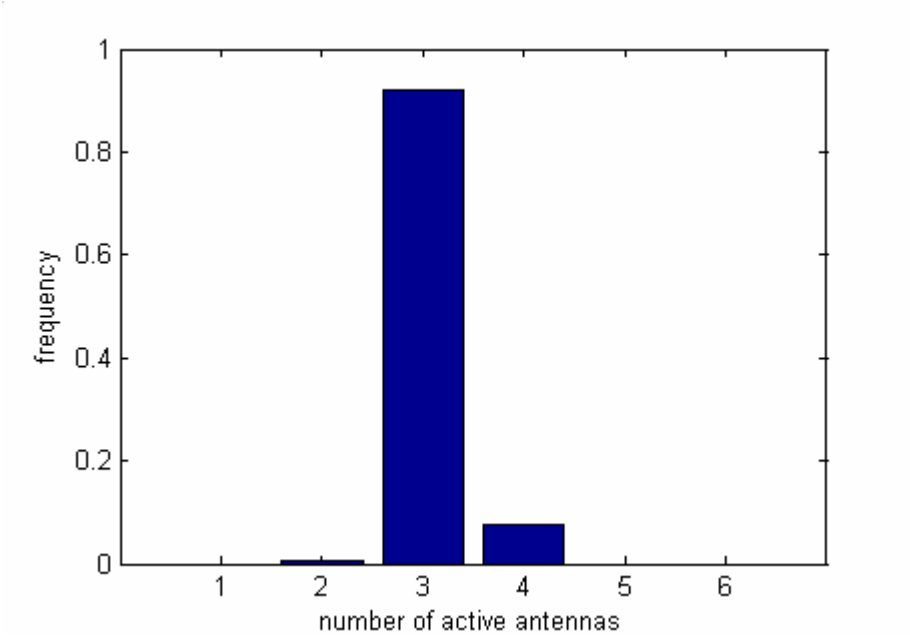

Figure 3.5-b Histogram of the number of active antennas for fading scenario 1

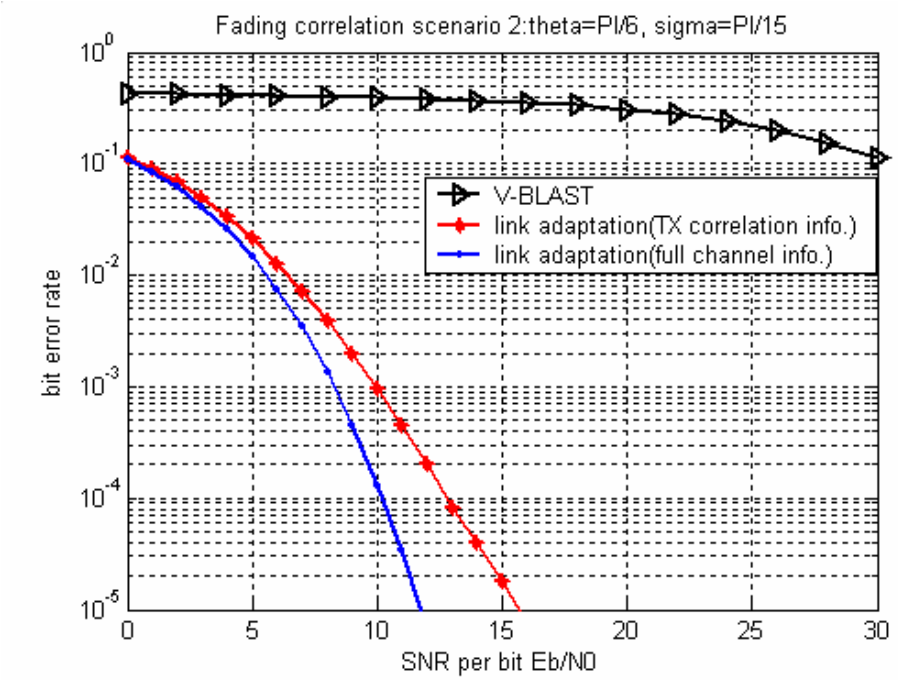

Figure 3.6-a Joint antenna selection and link adaptation for fading correlation scenario 2

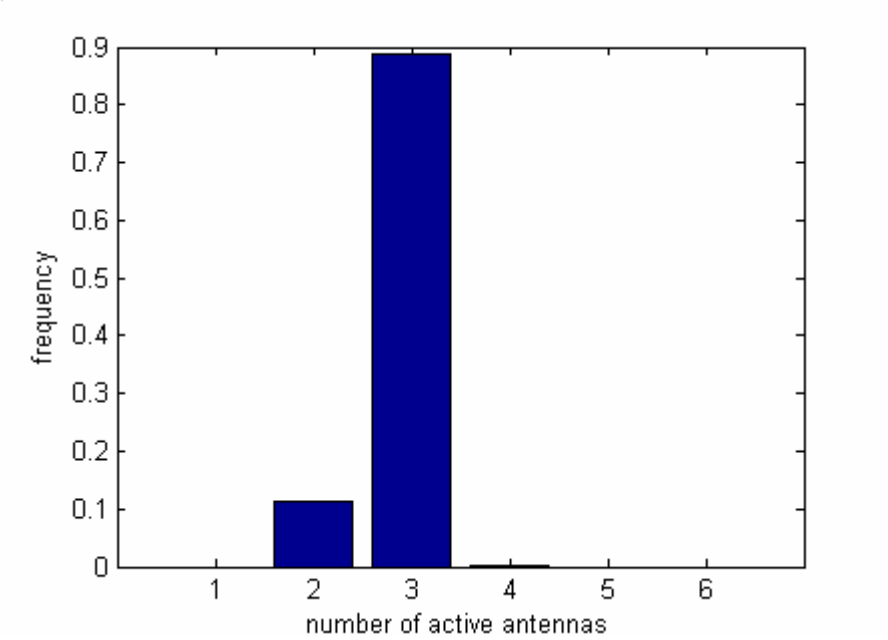

Figure 3.6-b Histogram of the number of active antennas for fading scenario 2

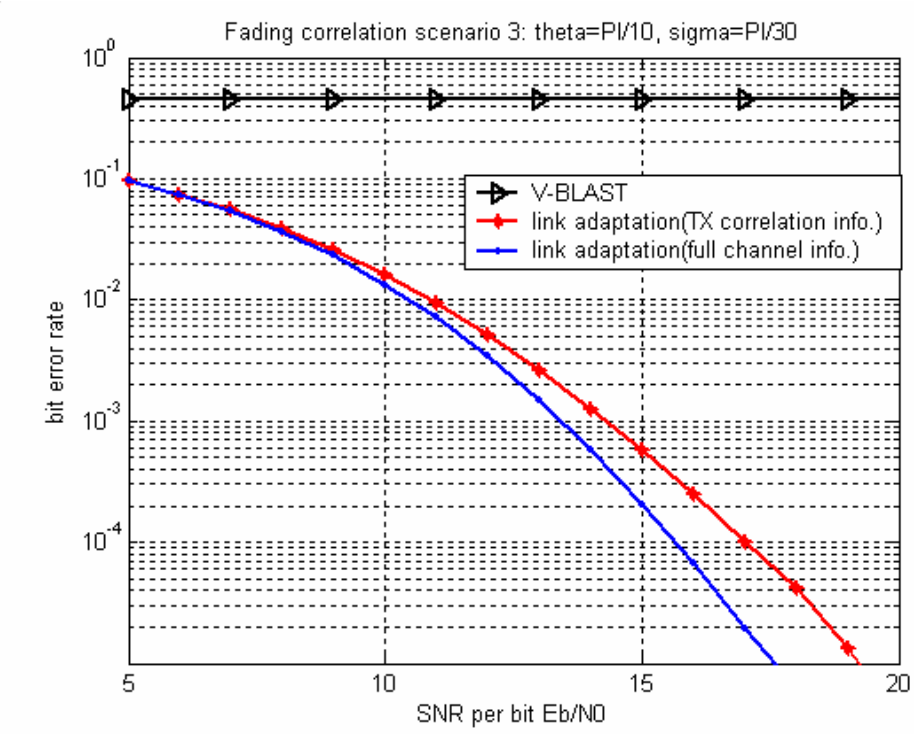

Figure 3.7-a Joint antenna selection and link adaptation for fading correlation scenario 3

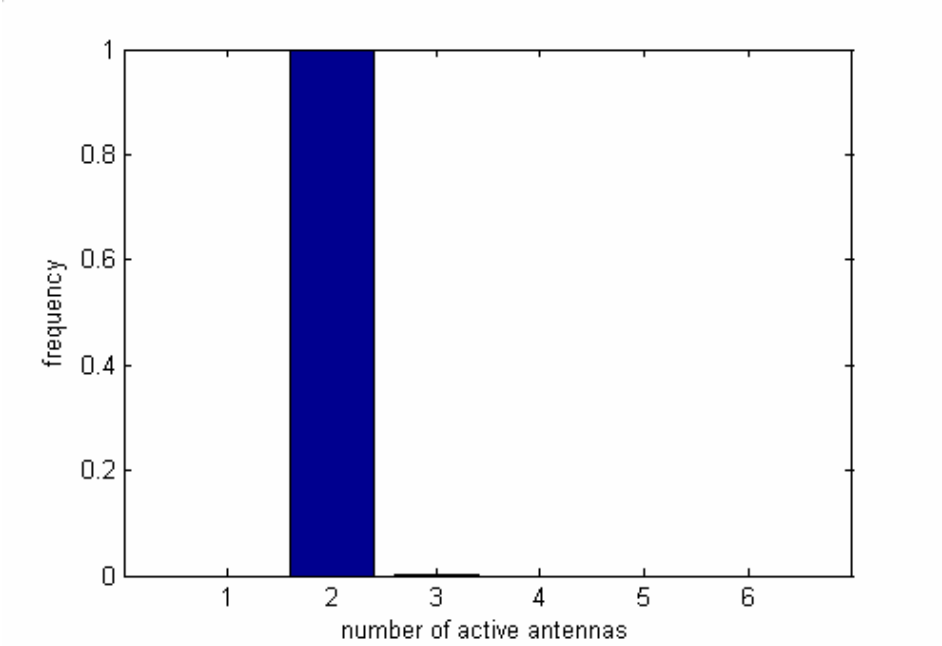

Figure 3.7-b Histogram of the number of active antennas for fading scenario 3

In the Figure 3.5b – Figure 3.7b (generated for the link adaptation based on the full channel information) we plot the histograms of the number of active antennas for the three fading correlation scenarios. From the histograms, we can see that the number of active antennas also decreases when the correlation increases.

Table 3.5 below illustrates the active transmit antenna index and the constellation carried by each active antenna using our link adaptation only based on channel correlation information. In contrast to that based on the full channel information, this configuration only depends on channel physical characteristics, and is invariant to instantaneous channel realizations. It's interesting to see that the first two active antennas are always antenna 1 and antenna 6, which accords with the practical situation: antenna 1 and antenna 6 have the largest distance, so their correlation is the smallest.

Table 3.5 Active antenna index and constellation carried by each active antenna

Scenario 1	Antenna 1(64QAM),	Antenna 6 (16QAM),	Antenna 3. (4QAM)
Scenario 2	Antenna 1 (16QAM),	Antenna 6 (16QAM),	Antenna 3 (16QAM)
Scenario 3	Antenna 1(256QAM),	Antenna 6 (4QAM)	

3.6 Summary

In this chapter, we propose joint antenna selection and link adaptation algorithms for both uncorrelated and correlated MIMO channels. Simulation results show that in most situations, significant performance gains are achieved compared with traditional equal power and equal rate V-BLAST. We also propose a simplified link adaptation algorithm based on the estimation of optimal number of active transmit antennas for Rayleigh i.i.d. MIMO channels. For correlated MIMO, we propose a link adaptation algorithm only based on channel correlation information, which is more practical in realization than that based on the instantaneous channel information, while approaching the latter in performance as the fading correlation increases. Finally, our antenna selection and link adaptation algorithms can be readily extended to other antenna selection applications, such as capacity maximization for both uncorrelated and correlated MIMO systems.

Chapter 4

Asymptotic Analysis on the Interaction between Spatial Diversity and Multiuser Diversity in Wireless Networks

4.1 Background

Diversity has long been established as key technology that enables reliable and high-data-rate wireless communications. While diversity can be achieved in many forms, two of them attract much research interest recently. One is spatial diversity realized through employing multiple antennas at either the transmitter or receiver end, or both, the idea of which is not new but interest on which is rekindled with the introduction of multi-input multi-output (MIMO) systems [22][21]. In a multiuser wireless network, there is another form of diversity called multiuser diversity [38][62], which reflects the fact of independent fluctuations of different users' channels. Multiuser diversity can be exploited to increase the system throughput, through intentionally transmitting to the user(s) with good channels at each instant (opportunistic scheduling). Spatial diversity techniques typically reside in the physical (PHY) layer, while multiuser diversity is obtained through user scheduling at the medium-access control (MAC) layer. It is therefore interesting to understand how these two diversity techniques combine to determine overall network performance and how they interact with each other.

There exist some work on joint spatial diversity and multiuser diversity systems. In particular, the capacity analysis for Rayleigh fading channels is given in [44], and in [9] for more general Nakagami fading channels. Some have suggested that spatial diversity can actually diminish the advantages of multiuser diversity [62][40][27][34]. Intuitively, this can

be explained by observing that multiuser diversity takes advantage of fading by "riding on the peak", which is unfortunately eliminated by spatial diversity. As noted in [40][9][39], however, this conclusion is valid only for open-loop but not closed-loop spatial diversity schemes, while user scheduling inherently requires feedback.

Our research is different from previous work in the following aspects. First, our study on the interaction between spatial diversity and multiuser diversity focuses on the asymptotic analysis, i.e., by allowing the number of antennas or users or both to go to infinity[4]. Besides mathematical tractability, asymptotic analysis also helps reveal some fundamental relationship of key system parameters, which may be concealed in the finite case by random fluctuations and other transient properties of channel matrices. Moreover in many scenarios (especially with respect to the number of antennas), convergence to the asymptotic limit is rather fast. Secondly, we put emphasis on the scheduling gain in capacity rather than the overall system signal-to-noise ratio (SNR) or capacity, which is the benefit we can really obtain through opportunistic scheduling over the traditional round robin scheduling. The impact of multiple antennas on multiuser wireless networks is increasingly drawing research interest very recently. This work will focus on spatial diversity systems; some pioneer study on spatial multiplexing systems can be found in [33][53].

This chapter is organized as follows. In Section 4.2, we give our system model with combined spatial diversity and multiuser diversity. Then we provide our asymptotic analysis corresponding to the above three scenarios in Section 4.3, 4.4 and 4.5, respectively, together with some numerical results for illustration purpose. Final conclusions are made in Section 4.6.

[4] Some asymptotic analysis with respect to the number of users is also pursued in [62] [34].

4.2 Joint Spatial Diversity and Multiuser Diversity System

We consider a homogeneous downlink multiuser MIMO communication scenario, which is envisioned to be of crucial importance for emerging wireless networks. Appropriate spatial diversity techniques are employed for each link. In this paper, we concretize our analysis with three spatial diversity schemes. The first employs well-known space-time block coding at the transmitter and maximum ratio combining at the receiver, coined as STBC/MRC, which does not require channel state information (CSI) at the transmitter end. As user scheduling inherently requires feedback, we further explore two closed-loop diversity schemes. One of them pursues joint maximum ratio transmission and maximum ratio combining (MRT/MRC), which provides the optimal performance reference for MIMO diversity techniques. The other exploits simple antenna selection on both ends (SC/SC), trading performance for complexity. MRT/MRC and SC/SC can be viewed as the two extremes for various hybrid selection combining schemes [45][65]. After diversity combining, the user with the best channel quality, in this case the highest effective link SNR, is chosen for communication in opportunistic scheduling. In contrast, the round robin scheduling simply selects the users in some deterministic order.

It is assumed that the base station has M antennas and each of the K users has N antennas. Throughout the chapter, when asymptotic analysis with respect to the size of antenna array is pursued, we allow both M and N to go to infinity, with their ratio $r = N/M$ fixed. The incorporation of the large M and fixed N scenario is relatively straightforward, and will be briefly discussed as well. We use $\mathbf{H}_k = \{h_{ij}^k\}$ $(1 \le k \le K)$ to denote the kth user's channel matrix. For simplicity, independent and identically distributed (i.i.d.) Rayleigh fading is considered for $\{\mathbf{H}_k\}_{k=1}^K$, but our analysis can be readily extended to

other fading scenarios. As will be seen, only the tail behaviors of the relevant probability distributions matter. The background noise is assumed to be white and Gaussian.

Throughout this paper we assume a block-flat fading scenario. Let $s_k(t)$ and $y_k(t)$ be the transmit and receive signal (after diversity combining) at time t for some selected user k respectively, then without loss of generality we have $y_k(t) = \sqrt{g_k}\, s_k(t) + n_k(t)$, where the noise $n_k(t)$ is assumed to have zero mean and unit variance, the average transmit SNR is $g_t = E\left|s_k(t)\right|^2$, and g_k is the channel gain obtained through diversity combining, which can be interpreted as normalized effective *link SNR*. Denote the probability distribution function (PDF) and cumulative distribution function (CDF) of g_k by $f_g(x)$ and $F_g(x)$ respectively, assumed the same for all users. In the opportunistic scheduling scheme, the base station chooses the user $k^* = \arg\max_k \left(g_k\right)_{k=1}^K$. Thus the resultant normalized *system SNR* seen by the base station is g_{k^*} with PDF

$$f_{g_{k^*}}(x) = K f_g(x) F_g^{K-1}(x). \tag{4.1}$$

Assuming that average transmit SNR is g_t, average system capacity obtained by opportunistic scheduling can be expressed as a function of K and M as

$$\bar{S}(K,M) = E\left(\log\left(1+g_t\left(\max_{1\le k\le K} g_k\right)\right)\right) = \int_0^{+\infty} \log(1+g_t x)\, f_{g_{k^*}}(x)dx. \tag{4.2}$$

We also define $\bar{R}(M)$, the corresponding average system capacity obtained by round-robin scheduling as a function of M as

$$\bar{R}(M) = E\left(\log(1+g_t g)\right) = \int_0^{+\infty} \log(1+g_t x)\, f_g(x)dx. \tag{4.3}$$

Finally, in order to measure the benefit brought by multiuser diversity, we define the scheduling gain $G(K,M)$ as the average capacity gain boosted by opportunistic scheduling from $\bar{R}(M)$:

$$G(K,M) = \bar{S}(K,M) - \bar{R}(M). \tag{4.4}$$

In the remainder of this chapter, we adopt the following notations for the limiting behaviors of two functions $f(x)$ and $g(x)$ with $\lim_{\substack{x\to\infty \\ or\ x\to 0}} \frac{g(x)}{f(x)} = c$: $g(x) = O(f(x))$ for $0 < |c| < \infty$; $g(x) \sim f(x)$ for $c = 1$; $g(x) = o(f(x))$ for $c = 0$; and $g(x) = w(f(x))$ for $c = \infty$. When convergence of a sequence of random variables is involved, shorthand notation "*D*" stands for in distribution, "*P*" for in probability, "*r*" for in *r*th mean, and "*a.s.*" for almost surely. The user index will be omitted from relevant notations when no ambiguity is incurred.

4.3 Asymptotic System Capacity and Scheduling Gain as K Goes to Infinity while M Keeps Fixed

In this section, we will examine $\lim_{K\to\infty} \bar{S}(K,M)$ with M fixed. As a motivation, we first summarize some relevant results in literature[12][20] as the following Lemma.

Lemma 6: Let $X_1, ..., X_K$ be i.i.d random variables with CDF $F(x)$, with $\Omega(F) = \sup\{x : F(x) < 1\}$. Suppose there is a real number x_1 such that for all $x_1 \le x < \Omega(F)$, $f(x) = F'(x)$ and $F''(x)$ exist and $f(x) \neq 0$. Define the growth function

$$g(x) = \frac{1 - F(x)}{f(x)}. \tag{4.5}$$

If $\lim_{x\to\Omega(F)} g(x) = c \ge 0$, the following standardized extreme converges in distribution as

$$\frac{\max_{1\le k\le K} X_k - b_K}{a_K} \xrightarrow{D} \Lambda(x) = \exp\left(-e^{-x}\right), \tag{4.6}$$

with

$$b_K = F^{-1}(1-1/K) \text{ and } a_K = (Kf(b_K))^{-1}. \tag{4.7}$$

Remark: From (4.5) and (4.7), we can find a connection between the growth function and a_K, i.e., $g(b_K) = \frac{1-F(b_K)}{f(b_K)} = \frac{1}{Kf(b_K)} = a_K$, therefore if $\lim_{x\to\Omega(F)} g(x) = c \geq 0$, we can obtain $\lim_{K\to\infty} a_K = c \geq 0$. Thus the result of (4.6) indicates that $\max_{1\leq k\leq K} X_k$ "grows like" b_K in a coarse sense [34][62], and is widely used in the study of opportunistic communications involving extreme values and order statistics. This result can actually be strengthened from existing literature [20]: if $c = 0$ (or $\lim_{K\to\infty} a_K = 0$), $\max_{1\leq k\leq K} X_k - b_K \xrightarrow{P} 0$, otherwise $\max_{1\leq k\leq K} X_k / b_K \xrightarrow{P} 1$. Nonetheless, our desired outcomes, which are concerned with the convergence of the expected values of functions of $\max_{1\leq k\leq K} g_k$, require a yet stronger result as stated below. Our main contributions in this section lie in providing sufficient conditions for this stronger result to hold, and an explicit expression for the corresponding system capacity (and scheduling gain) that is general enough to include many practical scenarios of interest (see (4.9)).

Theorem 1: Let $g_1,...,g_K$ be i.i.d. positive random variables with absolutely continuous CDF $F_g(x)$ and PDF $f_g(x)$, as given in Lemma 6 with $\Omega(F_g) = +\infty$. Define $g_g(x) = \frac{1-F_g(x)}{f_g(x)}$. If $\lim_{x\to\infty} g_g(x) = c \geq 0$, whose derivative $g_g'(x) = O(1/x^{d_1})$ with $d_1 > 0$, and $b_K = F_g^{-1}(1-1/K) = O\left((\log K)^{d_2}\right)$ with $0 < d_2 \leq 1$, then

$$\lim_{K\to\infty} \left\{ \bar{S}(K,M) - \log\left(1+g_t b_K\right) \right\} = 0. \tag{4.8}$$

Proof: See Appendix C.

According to (4.8), the system capacity (and scheduling gain) is asymptotically equivalent to $\log\left(1+g_t b_K\right)$ when given conditions are fulfilled. Note that all these conditions involve only the tail behaviors of the distributions of individual link SNR. In the following, we examine a form of special interest, which is general enough to cover common fading models and spatial diversity schemes.

Corollary 1: If $f_g(x) \sim a\, x^p e^{-qx^v}$ as $x \to \infty$ with $a > 0$, $q > 0$, $v \geq 1$ and any p, i.e., $f(x)$ is tail equivalent to $a\, x^p e^{-qx^v}$, then $\lim_{K\to\infty}\left\{\overline{S}\left(K,M\right) - \log\left(1+g_t b_K\right)\right\} = 0$, where (up to the second-order approximation[5])

$$b_K = \left(\frac{1}{q}\log t\,K\right)^{1/v} + \frac{p+1-v}{qv^2}\frac{\log\dfrac{\log t\,K}{q}}{\left(\dfrac{1}{q}\log t\,K\right)^{(v-1)/v}}, \tag{4.9}$$

with $t = \dfrac{a}{qv}$.

Proof: See Appendix C.

Remark: The parameter a only appears in t, which is typically not important in large K analysis. In general, a smaller n and q indicate a better system performance, as seen from the first term of (4.9). A larger p also helps, though only at the second-order sense.

In the remaining part of this section, we demonstrate the applications of our results, Theorem 1 and Corollary 1, through some representative systems jointly exploiting spatial diversity and multiuser diversity. As mentioned before, we assume Raleigh fading for simplicity. The key step lies in examining the tail behavior of the PDF of the corresponding effective link SNR. Once it is verified to take the form given in Corollary 1, we can readily

[5] We define the first order approximation when truncated at $\log K$, and the second order approximation when truncated at $\log\log K$.

conclude that the corresponding asymptotic system capacity (and scheduling gain) is given by $\log(1+g_t b_K)$, with b_K given by (4.9).

STBC/MRC

As seen in (2.6), the PDF of the normalized effective link SNR for a generic user of STBC/MRC is

$$f_g^{STBC/MRC}(x) = \frac{M^{MN}}{(MN-1)!} x^{MN-1} e^{-Mx}, x \geq 0. \quad (4.10)$$

Clearly Corollary 1 holds with

$$a = \frac{M^{MN}}{(MN-1)!}, p = MN-1, q = M, \text{and } n = 1.$$

So the corresponding asymptotic system capacity (and scheduling gain) is given by $\log\left(1+g_t b_K^{STBC/MRC}\right)$ with

$$b_K^{STBC/MRC} = \frac{1}{M}\left(\log\frac{M^{MN-1}}{(MN-1)!}K + (MN-1)\log\log\frac{M^{MN-1}}{(MN-1)!}K\right) + O(\log\log\log K). \quad (4.11)$$

SC/SC

From (2.8), the PDF of the normalized effective link SNR for a generic user in this spatial diversity mode is

$$f_g^{SC/SC}(x) = MNe^{-x}(1-e^{-x})^{MN-1}, x \geq 0. \quad (4.12)$$

The PDF in (4.12) is tail equivalent to MNe^{-x}. Again Corollary 1 holds with

$$a = MN, p = 0, q = 1, \text{and } n = 1.$$

So the corresponding asymptotic system capacity (and scheduling gain) is given by $\log\left(1+g_t b_K^{SC/SC}\right)$ with[6]

$$b_K^{SC/SC} = \log(MNK). \tag{4.13}$$

MRT/MRC

From (2.3), we know the PDF of the normalized effective link SNR for a generic user of MRT/MRC is

$$f_g^{MRT/MRC}(x) = F_g^{MRT/MRC}(x) tr(\boldsymbol{\Psi}_c^{-1}(x)\boldsymbol{\Phi}_c(x)), x \in (0,+\infty), \tag{4.14}$$

Though the PDF (4.14) is rather involved, fortunately we are only concerned with its tail behavior, as dictated by the following lemma.

Lemma 8: $f_g^{MRT/MRC}(x)$ is tail equivalent to $\frac{1}{(M-1)!(N-1)!}e^{-x}x^{M+N-2}$.

Proof: See Appendix C.

Therefore Corollary 1 holds with

$$a = \frac{1}{(M-1)!(N-1)!}, p = M+N-2, q = 1, \text{and } n = 1.$$

So the corresponding asymptotic system capacity (and scheduling gain) is given by $\log\left(1+g_t b_K^{MRT/MRC}\right)$ with

$$\begin{aligned} b_K^{MRT/MRC} &= \log\frac{1}{(M-1)!(N-1)!}K \\ &+ (M+N-2)\log\log\frac{1}{(M-1)!(N-1)!}K + O(\log\log\log K). \end{aligned} \tag{4.15}$$

Some interesting observations are readily in order. For all the above three schemes we have $n = 1$, which simplifies the expressions. From (4.11), we can observe a tradeoff

[6] This is one rare accurate expression. Note that in this case, the growth in transmit and receive antennas can be equivalently seen as an increase in the number of users (due to the i.i.d. assumptions).

between spatial diversity and multiuser diversity for an open-loop spatial diversity system (the factor of $1/M$ has a negative role in $b_K^{STBC/MRC}$, which directly determines the asymptotic system capacity $\bar{S}(K,M)$). Here we give a more rigorous proof and reveal how the ultimate capacity is related to M and N. For example, our result does show the positive role of the number of receive antennas N, though in a second-order sense, which is not clear from previous results in literature. It is also observed that the detrimental effect of multiple transmit antennas can be avoided with the closed-loop spatial diversity schemes, as seen in (4.13) and (4.15)[7]. And also from (4.13) and (4.15), we can infer that for the general hybrid selection combining schemes, the scaling laws should only have differences in the second order approximations. Numerical results in Fig. 1 verify that $\log(1+g_t b_K)$ is a good

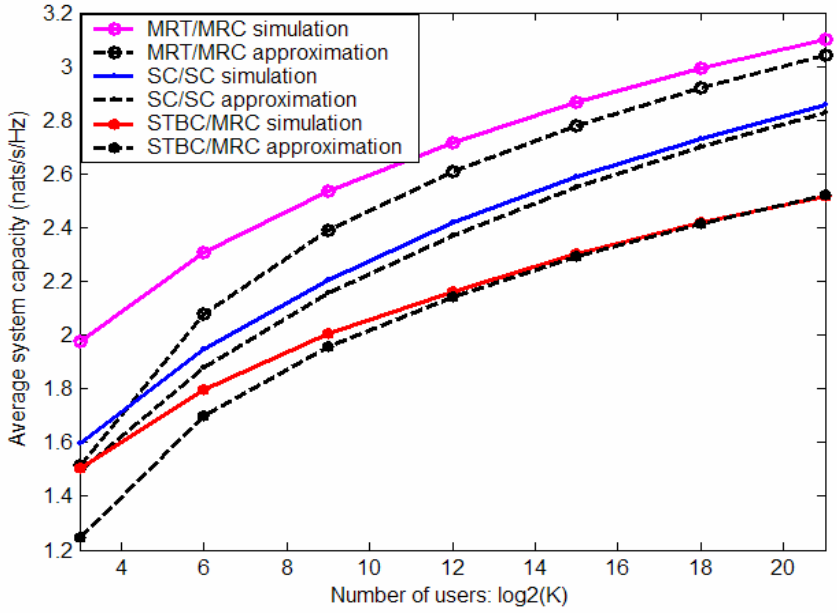

Figure 4.1 Average system capacity of opportunistic scheduling ($g_t = 0\ dB,\ M = N = 2$)

[7] Note that the coefficients of K inside the log functions are not important when K becomes large. In this sense, multiple antennas even help the MRT/MRC scheme.

approximation for the average capacity of the STBC/MRC, SC/SC and MRT/MRC systems using the opportunistic scheduler.

Note that Theorem 1 and Corollary 1 can potentially address a larger class of problems with respect to system and channel characteristics than what are presented here. Clearly the Nakagami-m distribution takes the form given in Corollary 1. For the Log-normal distribution, the transformation $y = \log g$ results in a normal distributed random variable. Ricean fading admits the following distribution

$$f_X(x) = \frac{1}{2s^2}\exp\left(-\frac{x+s^2}{2s^2}\right) I_0\left(\frac{\sqrt{x}s}{s^2}\right) \sim \frac{1}{2s}\frac{1}{\sqrt{2p}}(s^2x)^{-1/4}\exp\left(\frac{-\left(\sqrt{x}-s\right)^2}{2s^2}\right), \qquad (4.16)$$

where $s > 0$ is the amplitude of the line-of-sight component, $I_n(x)$ is the *n*th-order modified Bessel function of the first kind, and the tail-equivalence is due to the fact that for fixed *n*, $I_n(x) \sim \frac{e^x}{\sqrt{2p\,x}}$. As an example, we can show that for i.i.d. Ricean fading, SC/SC admits

$$b_{K,Ricean}^{SC/SC} = \left(s + \sqrt{2s^2 \log\left(\frac{KMNs}{\sqrt{2p}\,s}\right)} - \frac{s^2}{4}\frac{\log\log\frac{KMNs}{\sqrt{2p}\,s}}{\sqrt{2s^2\log\left(\frac{KMN}{\sqrt{2p}\,s}\right)}} \right)^2. \qquad (4.17)$$

As seen from above, scheduling gain $G(K,M)$ is an asymptotically increasing function of K. In the next section, we will show that $G(K,M)$ asymptotically decreases with M, no matter for open-loop or closed-loop spatial diversity systems.

4.4 Asymptotic Scheduling Gain as M Goes to Infinity while K Keeps Fixed

In this section, we will examine $\lim_{M\to\infty} G(K,M) = \lim_{M\to\infty}\left(\bar{S}(K,M) - \bar{R}(M)\right)$ with K fixed, which complements the study in Section 4.3. The scenario when both M and K go to

infinity will be studied in Section 4.5. We focus on the scenario that $N \to \infty$ as well, with $r = N/M$ fixed. However, all results apply to the case of fixed N as well, with r taken as 0 when applicable. The following theorem summarizes the main result in this scenario.

Theorem 2: Let m_M and s_M be the mean and standard deviation of the normalized effective SNR g_M for each individual link[8]. If $\lim_{M\to\infty} \frac{s_M}{m_M} = 0$, then $\lim_{M\to\infty} \left\{ \frac{\bar{R}(M)}{\log(1+g_t m_M)} \right\} = 1$ and $\lim_{M\to\infty} \frac{G(K,M)}{\bar{R}(M)} = 0$ when K keeps fixed. If we further have $\left\{ \log \frac{g_M}{m_M} \mathbf{I}_{(0,1)} \left(\frac{g_M}{m_M} \right) \right\}$ uniformly integrable[9], then $\lim_{M\to\infty} \left\{ \bar{R}(M) - \log(1+g_t m_M) \right\} = 0$, and $\lim_{M\to\infty} G(K,M) = 0$ when K keeps fixed.

Proof: See Appendix C.

Remark: As shown in the proof, $\lim_{M\to\infty} s_M / m_M = 0$ leads to the conclusion that g_M / m_M converges to 1 in 2nd mean (mean square). It is relatively straightforward to show that $\log(1+g_M/m_M) \xrightarrow{2} \log 2$ as g_M/m_M is always positive and $\log(1+x)$ grows slower than x in $(0,\infty)$. The difficulty with $\log g_M/m_M$ occurs when the argument falls on $(0,1)$, which necessitates the condition of uniform integrability. Also, according to Appendix C, we can see that s_M/m_M can be roughly used as a parameter to measure the scheduling gain.

In the following, we apply Theorem 2 on three representative systems. The key steps lie in showing $\lim_{M\to\infty} \frac{s_M}{m_M} = 0$ and the uniform integrability of $\left\{ \log \frac{g_M}{m_M} \mathbf{I}_{(0,1)} \left(\frac{g_M}{m_M} \right) \right\}$. Similar to

[8] We use subscript M to explicitly denote the dependence of corresponding quantities on M.

[9] $\mathbf{I}_A(\mathbf{g})$ is the indicator function on the set A, i.e. $\mathbf{I}_A(w) = \begin{cases} 1 & if \quad w \in A \\ 0 & if \quad w \notin A. \end{cases}$

Section 4.3, the cases of STBC/MRC and SC/SC are relatively easy, while things become much more involved with MRT/MRC.

STBC/MRC

In this case, the PDF of $g_M^{STBC/MRC}$ is given in (4.10), from which it's straightforward to obtain $m_M^{STBC/MRC} = N = rM$ and $s_M^{STBC/MRC} = \sqrt{N/M}$. Clearly we have

$$\lim_{M\to\infty} \frac{s_M^{STBC/MRC}}{m_M^{STBC/MRC}} = \lim_{M\to\infty} \frac{1}{\sqrt{rM}} = 0. \tag{4.18}$$

The PDF of $X_M^{STBC/MRC} = g_M^{STBC/MRC} \big/ m_M^{STBC/MRC}$ is given by

$$f_M^{STBC/MRC}(x) = \frac{MN^{MN}}{(MN-1)!} x^{MN-1} e^{-MNx} \le Cx, \text{ when } x<1, \tag{4.19}$$

where we can bound the coefficient due to Stiring's formula. Clearly, $\left\{\log X_M^{STBC/MRC} \mathbf{I}_{(0,1)}\left(X_M^{STBC/MRC}\right)\right\}$ is uniformly integrable. So

$$\lim_{M\to\infty} \left\{\bar{R}^{STBC/MRC}(M) - \log\left(1+g_t rM\right)\right\} = 0, \text{ and } \lim_{M\to\infty} G^{STBC/MRC}(K,M) = 0, \tag{4.20}$$

with K fixed.

SC/SC

In this case it's easy to obtain $m_M^{SC/SC} = \sum_{i=1}^{MN} \frac{1}{i} \to \log(MN) + C_0$ and $s_M^{SC/SC} = \sqrt{\sum_{i=1}^{MN} \frac{1}{i^2}} \to \sqrt{\frac{p^2}{6}}$ as $M \to \infty$, where C_0 is the Euler's constant (see page 298 of [12]). Clearly we have

$$\lim_{M\to\infty} \frac{s_M^{SC/SC}}{m_M^{SC/C}} = \lim_{M\to\infty} \frac{1}{\log(MN)} = 0. \tag{4.21}$$

When $M \to \infty$, the single-link SC/SC is equivalent to the corresponding multiuser scheduling scenario when $K \to \infty$. Results in Section 4.3 can be directly applied to get for fixed K

$$\lim_{M\to\infty}\left\{\bar{R}^{SC/SC}(M)-\log\left(1+g_t\log(MN)\right)\right\}=0, \text{ and } \lim_{M\to\infty} G^{SC/SC}(K,M)=0. \quad (4.22)$$

MRT/MRC

The calculation for the MRT/MRC case is more difficult. The closed-form expressions for the $m_M^{MRT/MRC}$ and $s_M^{MRT/MRC}$ are unknown. In the asymptotic scenario, it is known that [69]

$$\frac{1}{M} l_{\max} \xrightarrow{a.s.} (1+\sqrt{r})^2. \quad (4.23)$$

But surprisingly, $\lim_{M\to\infty}\frac{1}{M} m_M^{MRT/MRC} = \lim_{M\to\infty} E\left(\frac{1}{M} l_{\max}\right)$ remains open in literature, which is solved here through the following lemma.

Lemma 9: Let $\mathbf{H}$ be an $N\times M$ matrix with i.i.d. complex entries with $E(h_{ij})=0$, $E(|h_{ij}|^2)=1$, and $E(|h_{ij}|^4)<\infty$. Define $l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^H\right)=\left\|\frac{1}{\sqrt{M}}\mathbf{H}\right\|^2$ [10]. Then

$$\lim_{M\to\infty} E\left(l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^H\right)\right)=\left(1+\sqrt{r}\right)^2 \text{ and } \lim_{M\to\infty} s\left(l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^H\right)\right)=0,\ r=\lim_{M,N\to\infty}\frac{N}{M}. \quad (4.24)$$

Proof: See Appendix C.

Remark: A more important conclusion from Lemma 9 is that

$$\lim_{M\to\infty}\frac{s_M^{MRT/MRC}}{m_M^{MRT/MRC}}=0. \quad (4.25)$$

[10] Here $\|\cdot\|$ is the induced spectral norm on matrix, denoting the largest singular value.

The uniform integrability of $\left\{\log X_M^{MRT/MRC}\mathbf{I}_{(0,1)}\left(X_M^{MRT/MRC}\right)\right\}$ is already verified in Proposition 4.2 of [15]. Therefore, for fixed K,

$$\lim_{M\to\infty}\left\{\bar{R}^{MRT/MRC}(M)-\log\left(1+g_t(1+\sqrt{r})^2M\right)\right\}=0\text{, and }\lim_{M\to\infty}G^{MRT/MRC}(K,M)=0. \quad (4.26)$$

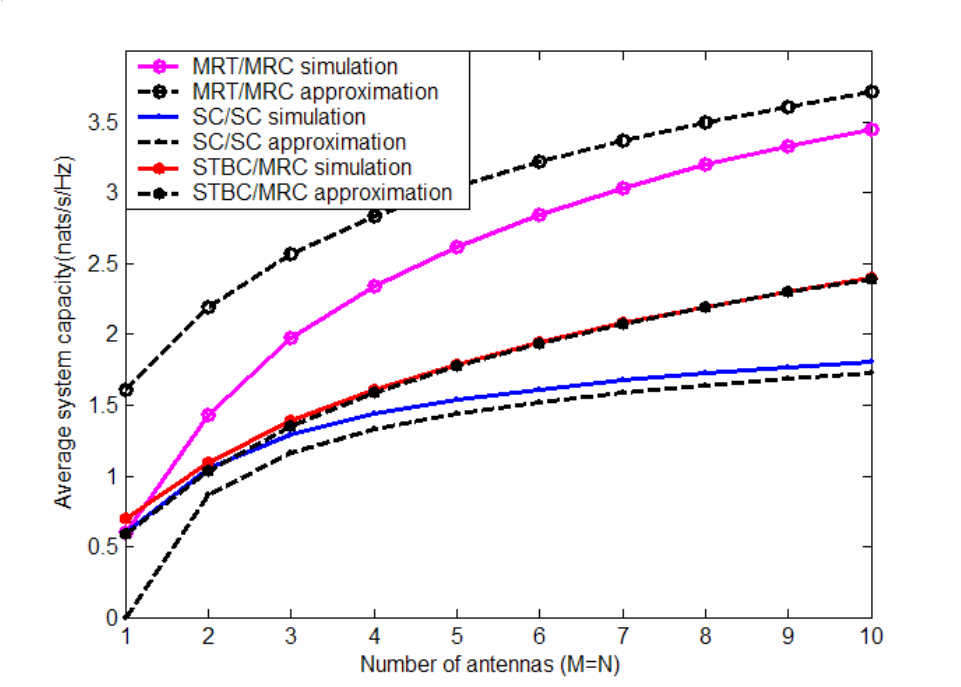

Figure 4.2 Average throughput of round robin scheduling (SNR=0dB)

Theorem 2 and the above examples indicate that, given the number of users, the scheduling gain will diminish when the number of antennas goes to infinity, if the mean of the link SNR grows at a higher-order rate than its variance. Intuitively, the mean corresponds to what we obtain through round-robin scheduling, while the variance really contributes to the scheduling gain. This is reminiscent of the multiple-antenna channel hardening effect studied in [33]. Therefore, multiuser scheduling is not worthwhile in an antenna-dominant environment. It is also interesting to see the difference in $\lim_{M\to\infty}\bar{R}(M)$ for different diversity techniques. For STBC/MRC, it achieves a constant unless the number of receive antennas N

also grows with M [11]. For SC/SC, it grows like $\log\log M$, but less impressive than that achieved by MRT/MRC, $\log M$. Numerical results in Figure 4.2 indicate a good match between the simulation results $\bar{R}(M)$ and the approximation results $\log(1+g_t m_M)$ for the above three cases as M grows. Figure 4.3 verifies through simulations that the scheduling gain will diminish as the number of antennas grows for both the open-loop and closed-loop spatial diversity schemes. Furthermore, based on the above discussion[12] we roughly have $\frac{s_M^{SC/SC}}{m_M^{SC/C}} \sim \frac{1}{\log M}$, $\frac{s_M^{STBC/MRC}}{m_M^{STBC/MRC}} \sim \frac{1}{M}$, and $\frac{s_M^{MRT/MRC}}{m_M^{MRT/MRC}} \sim \frac{1}{M^{2/3}}$, which intuitively explains the different decay rates shown in Figure 4.3. The analysis on the mean and variance will also be useful for the discussion in the following section.

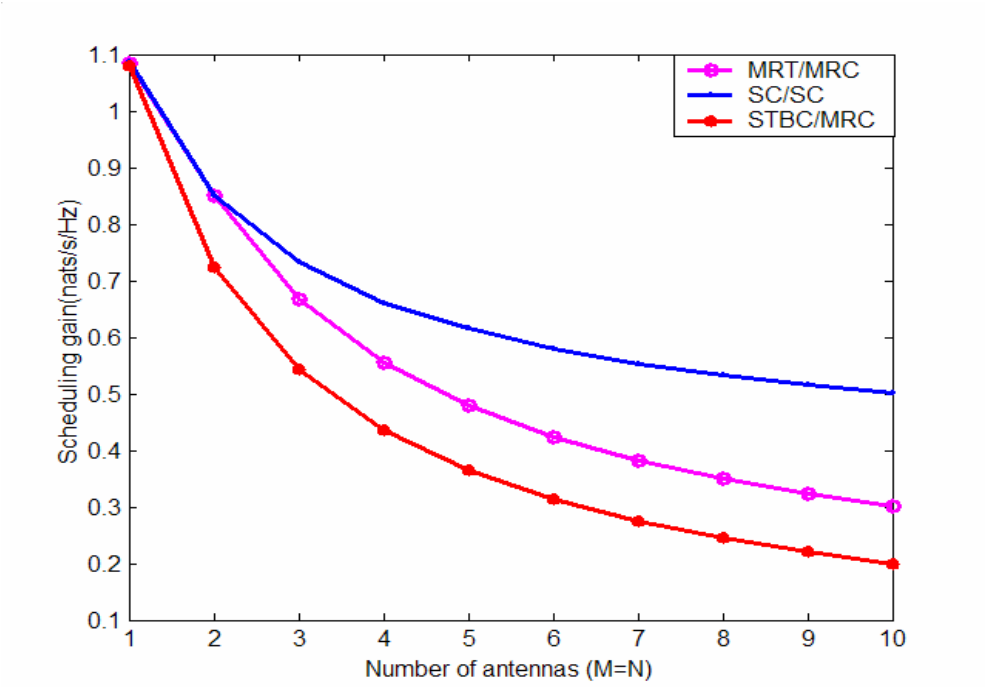

Figure 4.3 Scheduling gain as the number of antennas grows (SNR=0dB, K=50)

[11] As we mentioned, the results in (4.20), (4.22) and (4.26) hold for large M and fixed N as well. In this case, rM in (4.20) should be replaced with N, and r in (4.26) taken as 0.

[12] For MRT/MRC, as will be seen in (4.29) and (4.30), we can roughly consider $s_M^{MRT/MRC} \sim M^{1/3}$.

4.5 Scheduling Gain when both M and K Goes to Infinity

In the previous two sections we have given rigorous asymptotic results when either M or K grows. An interesting question naturally arises: when both the number of users and antennas are allowed to grow simultaneously, how will $G(K,M)$ behave? Intuitively, this depends on the relative growth rate of M and K [13]. Our goal is to find a critical point $K = O(f(M))$, only beyond which multiuser scheduling is meaningful.

We again facilitate the study through asymptotic analysis. Since the number of antennas also grows to infinity, the results derived in Section 4.3 do not apply. We therefore take the following approach. Let $g(k,M)$ be the effective link SNR for the *kth* user. First we find two norming constants p_M and q_M to form $\frac{g(k,M)-q_M}{p_M}$, whose distribution is asymptotically independent of M as $M \to \infty$, i.e., $\frac{g(k,M)-q_M}{p_M} \xrightarrow{D} w_k$ for some random variable w_k. Desirably, the PDF of w_k takes the form given in Corollary 1. Then we can apply the results in Section III to obtain the scaling law for $\max\{w_k\}_{k=1}^{K}$, denoted as b_K, which leads us to approximate the scaling law for $\max\{g(k,M)\}_{k=1}^{K}$ by $q_M + p_M b_K$ when both M and K grow. Finally Combing this result with Theorem 2, we can approximate the asymptotic scheduling gain as $\log\left(\frac{q_M + p_M b_K}{m_M}\right)$. As will be seen, most often $q_M = m_M$; our approach thus nicely combine the effect of multiple antennas (p_M and q_M (or m_M)) and multiple users (b_K) for convenience of analysis. Now the problem boils down to the determination of the dominant factor between the two. Note that such approach was also

[13] A similar study is conducted in [53] for spatial multiplexing systems to guarantee that the system throughput can still scale linearly with M.

taken in the relevant study of [33], where MIMO capacity is shown to be asymptotically Gaussian when the number of antennas grows.

In the following, we demonstrate this approach through examining the asymptotic scheduling gain for STBC/MRC, SC/SC and MRT/MRC when both the number of antennas and users go to infinity, with $r = N/M$ fixed. The case of fixed N follows the same line and will also be briefly discussed.

STBC/MRC

Choose $p_M = \sqrt{N/M} = \sqrt{r}$ and $q_M = m_M = N = rM$. By Central Limit Theorem, we can get $\frac{g(k,M) - q_M}{p_M} \xrightarrow{D} w_k$, whose PDF is the standard normal distribution function $\frac{1}{\sqrt{2p}} e^{-\frac{x^2}{2}}$. From Corollary 1, we can obtain $b_K \sim \sqrt{2\log K}$. According to our approach the asymptotic scheduling gain is given by $\log\left(1 + \frac{\sqrt{2r\log K}}{rM}\right)$, which admits

$$G^{STBC/MRC}(K,M) \to \begin{cases} 0, & when \quad \log K = o\left(M^2\right) \\ c, & when \quad \log K = O\left(M^2\right) \\ +\infty, & when \quad \log K = w\left(M^2\right). \end{cases} \tag{4.27}$$

SC/SC

According to the results in Section 4.3, we know if we choose $q_M = \log(MN) \sim m_M$, $p_M = 1$, then $g(k,M) - q_M \xrightarrow{D} w_k$, whose CDF is $\exp\left(-e^{-x}\right)$. We can then obtain $b_K \sim \log K$ through Corollary 1 and the asymptotic scheduling gain is given by $\log\left(1 + \frac{\log K}{\log\left(rM^2\right)}\right)$. Therefore

$$G^{SC/SC}(K,M) \to \begin{cases} 0, & when \quad \log K = o(\log M) \\ c, & when \quad \log K = O(\log M) \\ +\infty, & when \quad \log K = w(\log M). \end{cases} \tag{4.28}$$

MRT/MRC

In this scenario, the norming constants p_M and q_M have already been obtained in [35] as

$$q_M = \left(\sqrt{M} + \sqrt{N}\right)^2 = M\left(1+\sqrt{r}\right)^2 \sim m_M, \tag{4.29}$$

$$p_M = (\sqrt{M} + \sqrt{N})(\frac{1}{\sqrt{M}} + \frac{1}{\sqrt{N}})^{1/3} = M^{1/3}\frac{\left(1+\sqrt{r}\right)^{4/3}}{\sqrt{r}}. \tag{4.30}$$

Further

$$\frac{g\left(k,M\right) - q_M}{p_M} \xrightarrow{D} w_k, \tag{4.31}$$

where w_k is a random variable whose distribution follows the Tracy-Widom law of order 2. This distribution is defined by

$$F_2(s) = \exp\left\{\int_s^\infty (x-s)q^2(x)dx\right\}, s \in \Re, \tag{4.32}$$

where $q(x)$ solves the nonlinear Painleve II differential equation

$$q''(x) = xq(x) + 2q^3(x), \tag{4.33}$$

and $q(x) \sim 2^{-1}p^{-1/2}s^{-1/4}e^{\frac{-2}{3}s^{3/2}}, as \quad x \to \infty$. From (4.32), we can obtain

$$f_2(s) = dF_2(s)/ds = F_2(s)\cdot\int_s^\infty q^2(x)dx \sim \int_s^\infty q^2(x)dx \sim (8p\,s)^{-1}e^{\frac{-4}{3}s^{3/2}}. \tag{4.34}$$

We can check that $f_2(x)$ given in (4.34) satisfies the necessary conditions given in Corollary 1, which leads to $b_K \sim (\log K)^{2/3}$. The asymptotic scheduling gain is thus given by $\log\left(1+\left(\frac{\log K}{M}\right)^{2/3}\right)$ and

$$G^{MRT/MRC}(K,M) \to \begin{cases} 0, & when \quad \log K = o(M) \\ c, & when \quad \log K = O(M) \\ +\infty, & when \quad \log K = w(M). \end{cases} \tag{4.35}$$

It is also interesting to extend our current analysis to the case of fixed N and large M. For STBC/MRC, the same norming constants can be used but with different interpretation; note that in this scenario, the shrinking of the link standard deviation with respect to the link mean occurs at a slower rate ($1/\sqrt{M}$ rather than $1/M$). As a consequence, we have

$$G_N^{STBC/MRC}(K,M) \to \begin{cases} 0, & when \quad \log K = o(M) \\ c, & when \quad \log K = O(M) \\ +\infty, & when \quad \log K = w(M). \end{cases} \tag{4.36}$$

But this is achieved at the price of a saturated link capacity even when $M \to \infty$ (see Section 4.4). The analysis (4.28) for SC/SC remains unchanged for fixed N, i.e.,

$$G_N^{SC/SC}(K,M) \sim G^{SC/SC}(K,M). \tag{4.37}$$

For MRT/MRC, the fixed N analysis deviates from above in that w_k formed as in (4.31) does not follow the Tracy-Widom law. Alternatively, by law of large numbers, it can be shown that $\max\{g(k,M)\}_{k=1}^{K}$ approaches the largest element among KN i.i.d. random

variables of the form $\sum_{i=1}^{M}\left|h_{i,1}\right|^2$. Following the same approach as above, we can obtain the asymptotic scheduling gain as $\log\left(1+\sqrt{\frac{2\log(KN)}{M}}\right)$, which still leads to

$$G_N^{MRT/MRC}(K,M) \sim G^{MRT/MRC}(K,M). \tag{4.38}$$

To verify the above results (when both M and N are large), we consider two relative growth rate between K and M, with $N/M=1$. First, we assume $K=M$ in Figure 4.4, which shows that the scheduling gain for SC/SC has a tendency towards saturation; while for MRT/MRC and STBC/MRC, the scheduling gain asymptotically decreases. Then in Figure 4.5, we assume $K=e^M$, which shows the scheduling gain for MRT/MRC almost saturates as M grows, while scheduling gain asymptotically increases for SC/SC and decreases for STBC/MRC. Due to the computation constraint, it's difficult to simulate the scenario $K=e^{M^2}$, but our results already show that for STBC/MRC, much more users are required to make opportunistic scheduling beneficial.

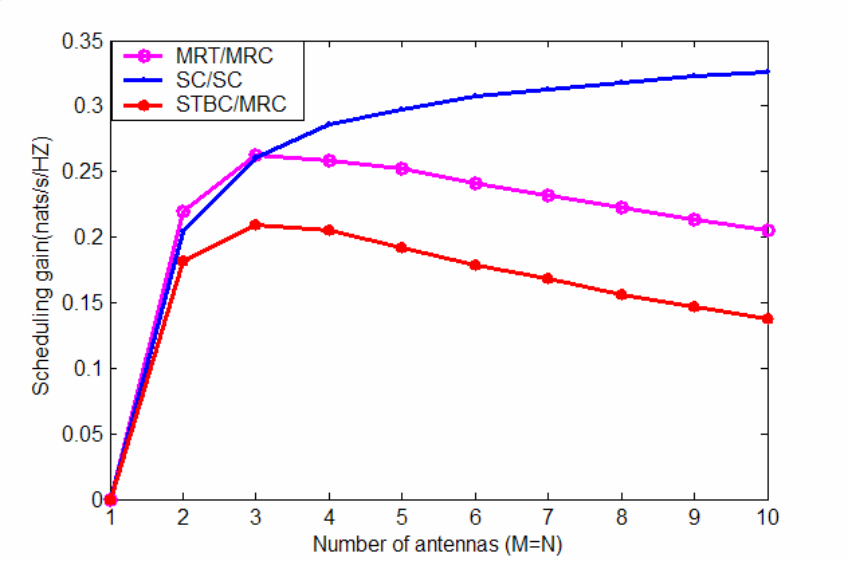

Figure 4.4 Scheduling gain as both the number of antennas and users grow (SNR=0dB,K=M)

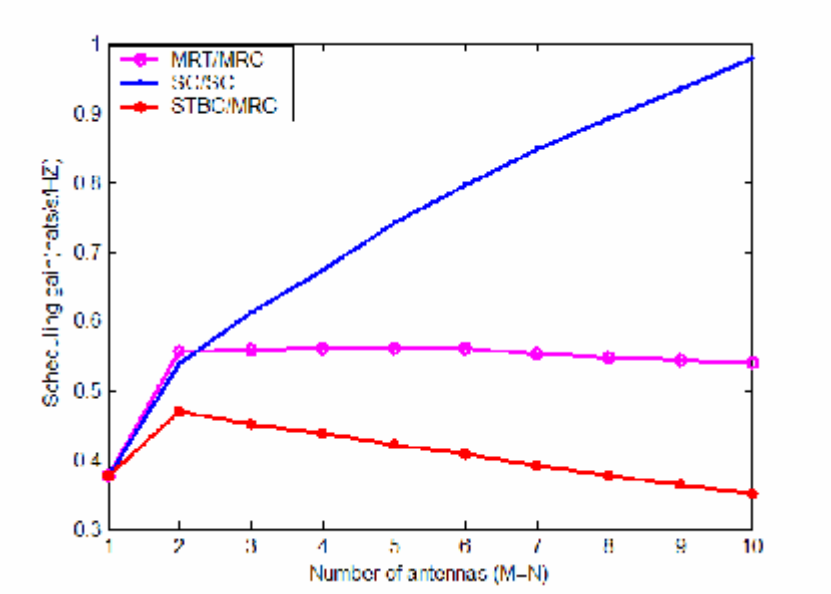

Figure 4.5 Scheduling gain as the number of antennas and users grow (SNR=0dB, K=exp(M))

One of the interesting observations from this section is that, when the number of antennas grows, generally we need even greater (sometimes exponentially greater) users to maintain the scheduling gain.

4.6 Summary

In this chapter, we present asymptotic analysis on the interaction between spatial diversity and multiuser diversity in wireless networks. Rigorous proofs and necessarily stronger results in terms of convergence are provided for some intuitions in this area. Equally important, explicit expressions of scheduling gain and average system capacity in various scenarios that reveal inter-connections and fundamental tradeoffs among key system parameters are given, which afford us some insights in real system design. The results of this chapter shows that in a multiuser MIMO wireless network, to increase the instantaneous throughput, scheduling only one best user is not a wise scheme. In the next chapter, we will propose a multiuser scheduling scheme which will greatly increase the system throughput compared with the one that schedules only one best user.

Chapter 5

Joint Tomlinson-Harashima Precoding and Scheduling for Multiuser MIMO with Imperfect Feedback

5.1 Background

Crosslayer study has become one of the burgeoning research fields in these few years, based on the realization that the commonly adopted layered protocol architecture, while facilitating the development within each layer, has hindered the optimization of the overall system, especially those utilizing wireless links [52]. At current stage, however, no formal systematic approaches have been built and applications are vastly different. In this paper, we focus on joint physical (PHY) and medium-access control (MAC) layer considerations, which are most relevant to harsh and unstructured wireless medium, and whose design has least interactions against other layers. Our target application is the downlink multiuser MIMO communications, which is envisioned to be of crucial importance to future wireless networks, and is believed to benefit significantly from a crosslayer design. Crosslayer study on multiuser MIMO systems has begun to attract attention only very recently, and very few of them ever explicitly address the details at the PHY layer [31] [59][7][1].

In multiuser communication scenario, multiuser diversity can be exploited through making judicious selections among users with independently faded channels [62]. While choosing the best user is optimal for single antenna systems and is simple to implement, it is decisively suboptimal for multiuser MIMO systems [64] [54] [70]. To the best of our knowledge, no rules of the thumb are yet known for multiuser MIMO scheduling, especially when perfect instantaneous channel information is not available. In the literature, multiuser

scheduling has been considered in the context of channel allocation for a SDMA/TDMA network (e.g., [57][73][56]), but mainly with the uplink and the assumptions that users are equipped with only one antenna or transmit only one data stream. In most of this work, "spatially compatible" users are grouped together in the same time or frequency slot, which is usually measured by channel correlation among users. This approach raises two potential concerns. First, a globally optimal allocation requires a thorough search of all possible choices, and suboptimal or heuristic alternatives induce complexity versus performance tradeoffs. Second, the physical layer details are largely neglected: either (1) the compatibility metric depends solely on the channel and is independent of the underlying transceiver structures; or (2) a conservative view is taken that treats multiuser interference as background noise. Such designs clearly fail to fully exploit the design opportunities at the physical layer. As an alternative, we propose to explore advanced yet feasible signal processing techniques at the physical layer in order to reduce the burden at the MAC layer and enhance overall network performance.

This chapter is organized as follows. A joint PHY and MAC design for multiuser MIMO downlink is described in Section 5.2, where perfect feedback from the users is assumed. In Section 5.3, we exploit the long-range prediction (LRP) technique to effectively reduce feedback and quantify the system throughput loss under imperfect feedback. Simulation results are given and analyzed in Section 5.4. Finally conclusions are presented in Section 5.5.

5.2 THP and Multiuser Scheduling for downlink MIMO

Downlink multiuser MIMO forms a vector broadcast channel, whose capacity region is resolved only recently with the dirty paper coding (DPC) approach [64]. As the realization

of DPC is rather involved [72][16], we exploit a more feasible PHY approach, Tomlinson-Harashima Precoding [17], which can be viewed as a suboptimal one-dimensional implementation of DPC. As a counterpart of the decision-feedback multiuser detection technique, THP has been employed in DSL systems and more recently in DS-CDMA and multiple-antenna systems to combat ISI and MAI [71][66][41]. In this work, we first propose a THP design for the multiuser MIMO downlink (especially, each user employs multiple antennas and receive multiple data streams), which does not seem to be completely addressed before. Furthermore, such a precoding structure results in interference-free parallel single-user MIMO channels, which greatly reduces the scheduling complexity while simultaneously improves overall system performance.

5.2.1 THP for multiuser MIMO

We consider a multiuser MIMO system with M antennas at the base station and N_i antennas at the ith user (each receive antenna is associated with a data stream), $1 \le i \le K$. A block diagram for the proposed THP scheme is given in Figure 5.1, and briefly illustrated here. At the transmitter, we use $\mathbf{x}_i$ to denote an $N_i \times 1$ symbol vector to be transmitted by user i. The backward signal vectors $\mathbf{b}_i$'s from the feedback filter bank $\mathbf{B}$ with size $\sum_{i=1}^{K} N_i \times \sum_{i=1}^{K} N_i$ are added to the intended transmitted vectors $\mathbf{x}_i$'s to pre-eliminate the interference from previous users (users from 1 to $i-1$), and the resultant signals are fed to modulo-operators, which serve to limit the transmit power. The output signals of modulo-operators are then passed through a power control unit before being transformed by feedforward filters $\mathbf{W}_i$'s to further remove the interference from future users

(users from $i+1$ to K). Finally, the signals are launched into the MIMO channels. As all interference is taken care of at the transmitter side, the receivers at the mobile users are left with some simple operations including power scaling (which is realized through automatic gain control (AGC) in Figure 5.1), reverse modulo-operation, and single user detection. The interested reader is referred to [17] for details.

In the following, we focus on the design of feedforward matrices $\{\mathbf{W}_i\}_{i=1}^K$ with size $M \times N_i$ and the feedback filter bank $\mathbf{B}$. We assume user channels $\{\mathbf{H}_i\}_{i=1}^K$ are quasi-static, whose information is perfectly known at the base station for the moment. The received signals at the receivers are then given as

$$\mathbf{y}_i = \mathbf{H}_i(\mathbf{W}_1\mathbf{s}_1 + \mathbf{W}_2\mathbf{s}_2 + ... + \mathbf{W}_K\mathbf{s}_K) + \mathbf{n}_i, \quad 1 \le i \le K, \tag{5.1}$$

where $\mathbf{s}_i$ with size $N_i \times 1$ denotes the output signal vector of the modulo operator for user *i*, and $\mathbf{n}_i$ is the circularly symmetric complex Gaussian noise vector with covariance matrix $s^2\mathbf{I}$. Define the overall system channel transfer function as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1\mathbf{W}_1 & \mathbf{H}_1\mathbf{W}_2 & .. & \mathbf{H}_1\mathbf{W}_K \\ \mathbf{H}_2\mathbf{W}_1 & \mathbf{H}_2\mathbf{W}_2 & .. & \mathbf{H}_2\mathbf{W}_K \\ : & : & : & : \\ \mathbf{H}_K\mathbf{W}_1 & \mathbf{H}_K\mathbf{W}_2 & .. & \mathbf{H}_K\mathbf{W}_K \end{bmatrix}. \tag{5.2}$$

The design target for feedforward matrices $\{\mathbf{W}_i\}_{i=1}^K$ is to make $\mathbf{H}$ a lower triangular one. This can be carried out in two steps. First, we enforce $\mathbf{H}$ to be block lower triangular and next we further enforce each block on the diagonal to be lower triangular. Let $\mathbf{\Gamma}_{i+1}$ denote the matrix of coefficients of orthonormal basis spanning $null(\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_i)$. The matrix $\mathbf{H}$

would be block lower triangular if we let $\mathbf{W}_i = \mathbf{\Gamma}_i \mathbf{A}_i$, where $\mathbf{A}_i$ is yet to be determined. The matrix $\mathbf{\Gamma}_{i+1}$ can be determined through singular value decomposition (SVD) as follows.

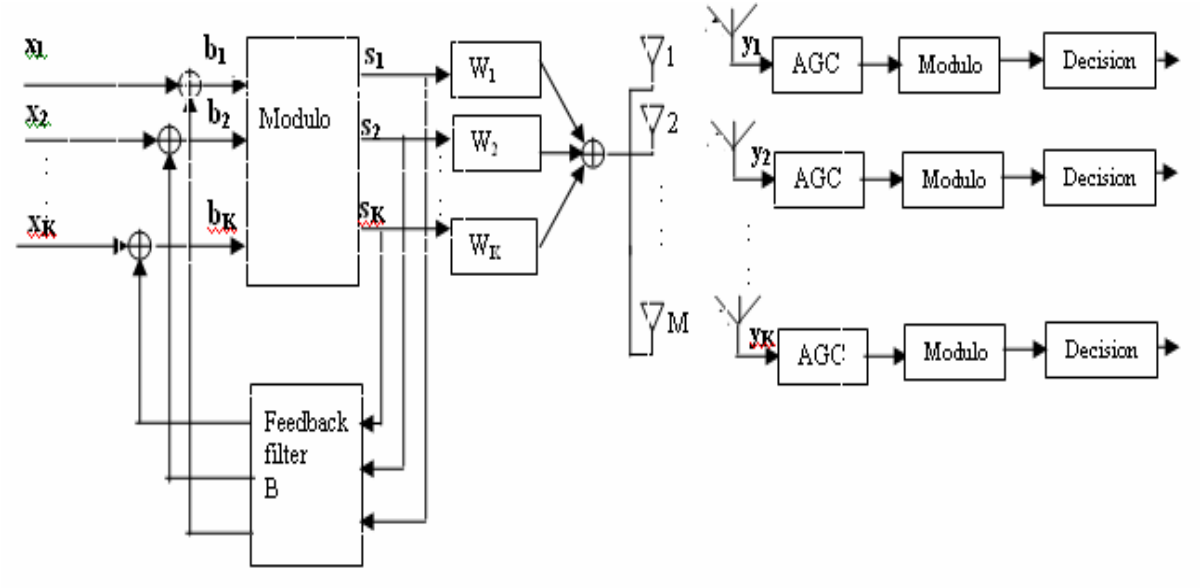

Figure 5.1 Block diagram of the proposed THP for multi-user MIMO downlink

Assume

$$[\mathbf{U}_i, \mathbf{\Sigma}_i, \mathbf{V}_i^*] = svd\left(\begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \mathbf{M} \\ \mathbf{H}_i \end{bmatrix}\right), \quad 1 \le i \le K-1, \tag{5.3}$$

then the rightmost $M - \sum_{k=1}^{i} N_k$ columns of $\mathbf{V}_i$ compose the joint null space of $\mathbf{H}_1, \mathbf{H}_2, ..., \mathbf{H}_i$, denoted as $\mathbf{\Gamma}_{i+1}$. Now we further enforce $\mathbf{H}_i \mathbf{\Gamma}_i \mathbf{A}_i$ to be lower triangular, which is realized through QR decomposition $(\mathbf{H}_i \mathbf{\Gamma}_i)^H = \mathbf{Q}_i \mathbf{R}_i$, and let $\mathbf{A}_i = \mathbf{Q}_i(:, 1: N_i)$ (the first N_i columns of $\mathbf{Q}_i$)

Once the feedforward matrices $\{\mathbf{W}_i\}_{i=1}^K$ are determined as above, the design of the feedback filter bank $\mathbf{B}$ endeavors to eliminate the residual inter-stream and inter-user interference. Decompose $\mathbf{H} = \mathbf{GL}$, where $\mathbf{G}$ is a diagonal matrix that extracts the diagonal elements of $\mathbf{H}$ and $\mathbf{L}$ is a lower triangular matrix with unit diagonal elements, then

$$\mathbf{B} = \mathbf{I} - \mathbf{L}. \tag{5.4}$$

is what we desire.

5.2.2 THP-Aided Scheduling for Multi-user MIMO Downlink

A direct consequence of the THP design is that, each user sees an interference-free MIMO channel, as evidenced from the following relationship (cf. Figure 5.1)

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n} = \mathbf{G}(\mathbf{Ls} + \mathbf{G}^{-1}\mathbf{n}) = \mathbf{G}\left[\mathbf{x} + (\mathbf{s} - \mathbf{b})\right] + \mathbf{n}, \tag{5.5}$$

where $\mathbf{y} = \left[\mathbf{y}_1^T \quad \mathbf{y}_2^T \quad ... \quad \mathbf{y}_K^T\right]^T$, $\mathbf{s} = \left[\mathbf{s}_1^T \quad \mathbf{s}_2^T \quad ... \quad \mathbf{s}_K^T\right]^T$ and $\mathbf{n} = \left[\mathbf{n}_1^T \quad \mathbf{n}_2^T \quad ... \quad \mathbf{n}_K^T\right]^T$ are given as in (5.1); $\mathbf{x} = \left[\mathbf{x}_1^T \quad \mathbf{x}_2^T \quad ... \quad \mathbf{x}_K^T\right]^T$ represents the transmitted data symbols with $\mathrm{tr}(E[\mathbf{x}_i\mathbf{x}_i^H]) \le P_i$ and $\sum_{i=1}^K P_i \le P$; and $\mathbf{b} = \left[\mathbf{b}_1^T \quad \mathbf{b}_2^T \quad ... \quad \mathbf{b}_K^T\right]^T$ collects the input vectors to the modulo-operators for all users At the receiver side, a power scaling with $\mathbf{G}_{ii}^{-1}$ $(1 \le i \le K)$ followed by the modulo-operator and decision device suffices to recover $\mathbf{x}_i$. Due to the THP design, $\mathbf{H}_i\mathbf{W}_i$ is the equivalent single-user channel matrix for user *i*. In this chapter, we follow the common practice in literature and adopt the information-theoretic spectral efficiency (assuming a Gaussian codebook and equal power allocation)

$$R_i = \log_2\left(\left|\mathbf{I} + \frac{P_i}{N_i s^2}(\mathbf{H}_i\mathbf{W}_i)(\mathbf{H}_i\mathbf{W}_i)^H\right|\right), 1 \le i \le K, \tag{5.6}$$

as the metric[14] for the user channel quality, which can be readily modified to accommodate actual modulation and coding schemes.

Based on the above observations, a simple multiuser scheduling scheme is given below.

Step1. Schedule the first user whose index is $\arg\max_{1\le i\le K} \log_2\left(\left|\mathbf{I}+\frac{P_i}{N_i s^2}\mathbf{H}_i\mathbf{H}_i^H\right|\right)$.

Step2. Based on the selected user(s), compute the feedforward matrix $\mathbf{W}_i$, then select among the remaining users whose equivalent channel matrix $\mathbf{H}_i\mathbf{W}_i$ will result in the largest contribution (as given in (5.6)) in sum rate).[15]

Step3. Repeat Step 2 until a given number of users has been selected, or no more users can be added due to channel rank deficiency.

This THP-aided multiuser scheduling scheme will be compared with various sub-optimal and optimal approaches in Section 5.4 to demonstrate its advantages.

5.3 Analysis of Imperfect Feedback

In previous discussion, perfect CSI is assumed available at the transmit side for PHY and MAC design. In practice, especially for FDD systems, such information is typically measured at the receiver side and fed back to the transmitter with some dedicated channels. The cost of perfect feedback in multiuser MIMO, if ever possible, grows quickly with the number of antennas, users, and system bandwidth, while in real systems the dedicated feedback channels are typically of low rate and prone to errors. In this section, we exploit the

[14] Rigorously speaking, (5.6) is not the achievable rate of the THP, however, it is shown in [7] that the sum rate achievable using THP can converge to the actual sum rate capacity at high SNR.

[15] Note that existing designs need not be changed, as newly added users are invisible to already selected users.

LRP technique to effectively reduce feedback and further quantify the system throughput loss under imperfect feedback.

5.3.1 Channel Prediction via LRP

LRP is a linear prediction method based on autoregressive modeling. With this technique, one can measure and feedback the time-varying CSI at a much lower rate than the data rate. Assume the complex fading process $h(t)$ is sampled at a rate $f_s = 1/T_s$, which is at least twice the maximum Doppler shift f_{dm} but can be much slower than the data rate. The sampled data is represented by $h_n = h(nT_s)$. Then the linear MMSE prediction of the future CSI sample $\hat{h}_n$ based on p previously observed CSI samples[16] $h_{n-1}, h_{n-2}, ..., h_{n-p}$ is

$$\hat{h}_n = \sum_{j=1}^{p} d_j h_{n-j}, \tag{5.7}$$

where $\{d_j\}$ are the coefficients of the linear prediction filter and p is the prediction order. This approach can also be extended to predict $t > 1$ samples ahead. Clearly, multi-step prediction can tolerate more delay in the CSI feedback with some loss in performance. The predicted samples can be interpolated to forecast the fading samples at the same rate as data rate. The reader is referred to [29] for a detailed description of this technique.

5.3.2 Actual Achievable Rate under Imperfect CSI

With imperfect feedback, THP design will not be able to completely eliminate all the interference, which necessarily degrades the system performance. Assume the estimated CSI for the ith user is $\mathbf{H}_i'$ ($1 \le i \le K$), and the designed feedforward matrix based on $\mathbf{H}_i'$

[16] We assume the samples of the fading channels are perfectly obtained through trainings.

$(1 \le i \le K)$ is $\hat{\mathbf{W}}_i$. In the precoding stage, the interference from the previous users for user i seen by the transmitter is $\hat{\mathbf{H}}_i \hat{\mathbf{W}}_j \mathbf{s}_j$ $(1 \le j \le i-1)$, while the actual interference is $\mathbf{H}_i \hat{\mathbf{W}}_j \mathbf{s}_j$. Therefore the residual interference from the previous users for the ith user is $(\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_j \mathbf{s}_j$ $(1 \le j \le i-1)$. Similarly the interference from the future users is $\mathbf{H}_i \hat{\mathbf{W}}_j \mathbf{s}_j$ $(i+1 \le j \le K)$. As for the self-interference due to the imperfect channel estimation, we can decompose $\mathbf{H}_i \hat{\mathbf{W}}_i$ into three non-overlapping parts representing its lower triangular, diagonal, and upper triangular components as $\mathbf{H}_i \hat{\mathbf{W}}_i = (\mathbf{H}_i \hat{\mathbf{W}}_i)_l + (\mathbf{H}_i \hat{\mathbf{W}}_i)_d + (\mathbf{H}_i \hat{\mathbf{W}}_i)_u$. Then the interference from user i itself is $((\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_i)_l + (\mathbf{H}_i \hat{\mathbf{W}}_i)_u$. In summary for the ith user, the actual received signal is given by

$$\mathbf{y}_i = \underbrace{(\mathbf{H}_i \hat{\mathbf{W}}_i)_d \mathbf{s}_i}_{\text{desired signal}} + \underbrace{\left(((\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_i)_l + (\mathbf{H}_i \hat{\mathbf{W}}_i)_u\right)\mathbf{s}_i}_{\text{self-interference}} + \underbrace{\sum_{j=1}^{i-1} (\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_j \mathbf{s}_j}_{\text{interference from previous users}} + \underbrace{\sum_{j=i+1}^{K} \mathbf{H}_i \hat{\mathbf{W}}_j \mathbf{s}_j}_{\text{interference from future users}} + \mathbf{n}_i, 1 \le i \le K. \tag{5.8}$$

With $\mathbf{H}_{i,d} \equiv (\mathbf{H}_i \hat{\mathbf{W}}_i)_d$, $\mathbf{H}_{i,s} \equiv ((\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_i)_l + (\mathbf{H}_i \hat{\mathbf{W}}_i)_u$, $\mathbf{H}_{i,j<i} \equiv (\mathbf{H}_i - \hat{\mathbf{H}}_i)\hat{\mathbf{W}}_j$, and $\mathbf{H}_{i,j>i} = \mathbf{H}_i \hat{\mathbf{W}}_j$, the achievable rate for the kth substream of user i (assuming Gaussian coding and equal power allocation) is given as

$$R_{i,k} = \log_2\left(1 + \frac{P_i}{N_i \sigma_I^2} \mathbf{H}_{i,d}(k,k)\mathbf{H}_{i,d}(k,k)^*\right). \tag{5.9}$$

with

$$s_I^2 = s^2 + \frac{P_i}{N_i}\mathbf{H}_{i,s}(k,:)\mathbf{H}_{i,s}(k,:)^* + \sum_{j=1}^{i-1}\frac{P_j}{N_j}\mathbf{H}_{i,j<i}(k,:)\mathbf{H}_{i,j<i}(k,:)^* + \sum_{j=i+1}^{K}\frac{P_j}{N_j}\mathbf{H}_{i,j>i}(k,:)\mathbf{H}_{i,j>i}(k,:)^*. \quad (5.10)$$

where (k,k) denotes the kth diagonal element of a matrix and $(k,:)$ represents the kth row of a matrix. The achievable rate for user i with imperfect feedback is then given by $R_i = \sum_{k=1}^{N_i} R_{i,k}$.

It is observed that the noise plus interference matrix $s^2\mathbf{I} + \frac{P_i}{N_i}\mathbf{H}_{i,s}\mathbf{H}_{i,s}^H + \sum_{j=1}^{i-1}\frac{P_j}{N_j}\mathbf{H}_{i,j<i}\mathbf{H}_{i,j<i}^H + \sum_{j=i+1}^{K}\frac{P_j}{N_j}\mathbf{H}_{i,j>i}\mathbf{H}_{i,j>i}^H$ is typically a diagonally-dominant matrix (a matrix whose diagonal elements are much larger than the off-diagonal ones), therefore the actual achievable rate of user i can be approximated by

$$R_i \approx \log_2\left[\left|\mathbf{I} + \frac{P_i}{N_i}\mathbf{H}_{i,d}\mathbf{H}_{i,d}^* \left(s^2\mathbf{I} + \frac{P_i}{N_i}\mathbf{H}_{i,s}\mathbf{H}_{i,s}^* + \sum_{j=1}^{i-1}\frac{P_j}{N_j}\mathbf{H}_{i,j<i}\mathbf{H}_{i,j<i}^* + \sum_{j=i+1}^{K}\frac{P_j}{N_j}\mathbf{H}_{i,j>i}\mathbf{H}_{i,j>i}^*\right)^{-1}\right|\right] \quad (5.11)$$

5.4 Numerical Results

To illustrate the effectiveness of our proposed THP based multiuser MIMO downlink scheduling scheme, the total average throughput of our scheduling scheme is compared with that of several well-known schemes in literature, assuming that L users will be scheduled simultaneously and each selected user receive as many streams as its antennas. The HDR approach [62], i.e., serving only the best user at each time, is used as a reference. In contrast, a thorough search based on DPC readily serves as a performance upper bound. Very recently, some preliminary study on feasible multiuser MIMO scheduling schemes is presented in [1],

which in some sense bears the same spirit as the "spatially compatible" scheduling mentioned before with a conservative view on the physical layer. In this work, the transmit antennas are partitioned among active users. Each user feedbacks the indices of the N_i antennas it desires together with those of the $\sum_{k \neq i} N_k$ antennas it wants the other users to be assigned to such that his instantaneous channel capacity is maximized. While this scheme is relatively easy to implement and has some throughput advantages over HDR (particularly for large K and moderate SNR), it suffers from two drawbacks as indicated above. One is computational complexity, since optimal scheduling requires a search of $\binom{M}{N_i}\binom{M-N_i}{\sum_{k \neq i} N_k}$ possibilities at *each user*. Second, as shown in Figure 5.2, it is still interference limited at high SNR.

In Figure 5.2, we compare the system throughputs of different scheduling schemes as a function of total transmit SNR. We consider a symmetric multiuser MIMO downlink system with $M = 4$ antennas at the BS and $N_i = 2$ antennas at each of $K = 20$ mobiles, and schedule at most $L = 2$ users in each time instance. It can be seen that our joint THP and scheduling scheme performs much better than the HDR and the MMUD-L scheduling scheme in [1], and approaches the performance of dirty paper coding (a thorough search for *two* best users), while with much affordable computational complexity.

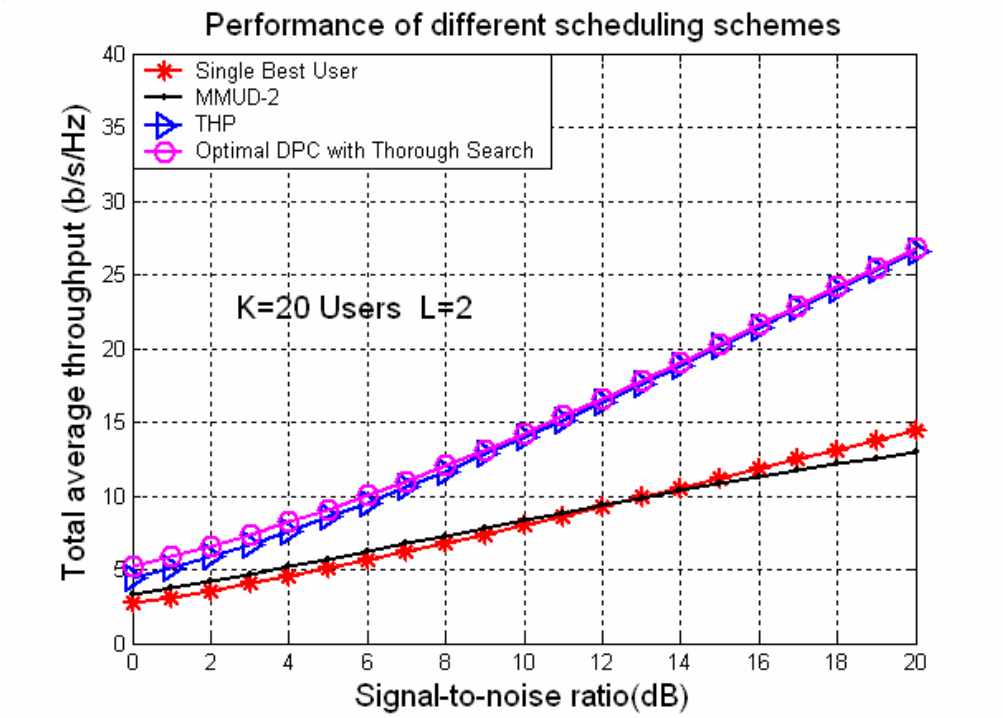

Figure 5.2 Performance comparison of different scheduling scheme

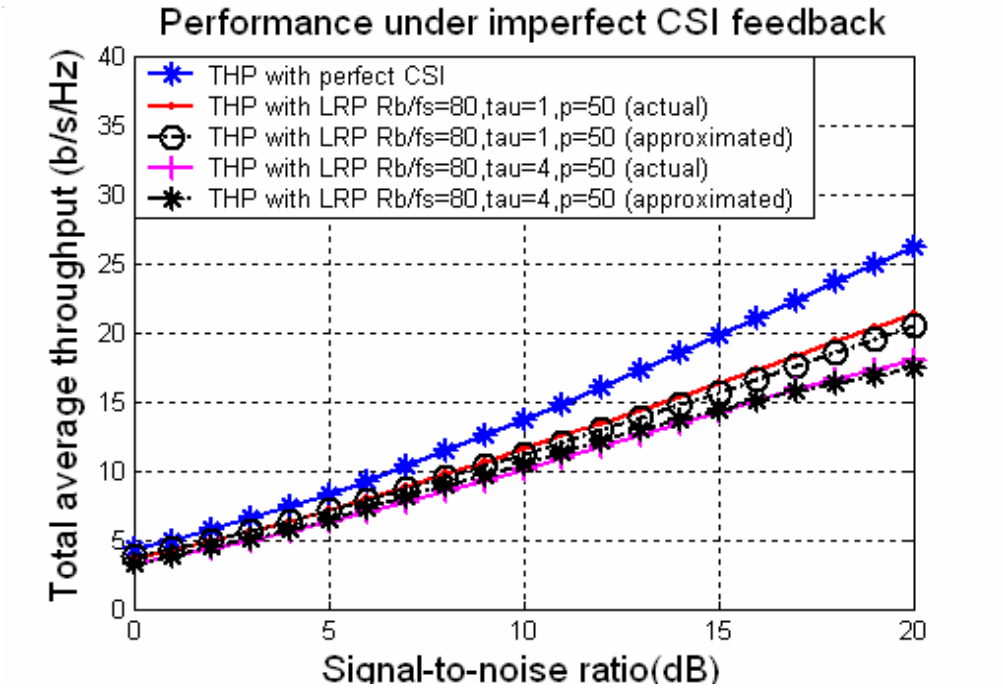

Figure 5.3 Performance under imperfect CSI feedback

In Figure 5.3, we re-evaluate our scheduling scheme with imperfect CSI feedback. A flat Rayleigh fading channel with Jakes' model with a Doppler frequency of 200 Hz (corresponding to a vehicular speed of 65 mi/h at a carrier frequency of 2 GHz) is considered. We use a channel sampling frequency of $f_s = 1600\text{Hz}$ and the prediction order in (5.7) is taken as $p = 50$, while the transmit data rate of each user is $R_b = 128\text{k}$ bps. So the channel is measured and fed back once every 80 symbols. We choose two different prediction steps for evaluation $t = 1$ and $t = 4$ (which corresponds to a prediction of 320 data symbols ahead). From Figure 5.3, we can see that even with imperfect CSI and slower feedback rate through LRP, our scheme still has a significant gain over the MMUD-L scheme and single best user scheduling scheme. Our simulation also indicates the good match between the true value and the approximate value of the total average rate under imperfect feedback.

5.5 Summary

In this chapter, we proposed a joint THP and scheduling scheme for multi-user MIMO downlink. Compared to some existing scheduling schemes, the proposed scheme greatly reduces the scheduling complexity while simultaneously improves overall system performance. Strictly speaking, the proposed THP-aided multiuser scheduling is not necessarily globally optimal, but we conjecture that the loss in optimality is negligible (when the number of scheduled users is given) and plan to undertake a comprehensive analysis of this approach.

Chapter 6

Conclusions and Future Work

In this dissertation, we studied several interesting topics with related to MIMO wireless communications, the main contributions and future work of the dissertation are summarized below:

Chapter 2: By deriving the approximation of the PDF of $s_{\max}^2$ at the origin, we obtain the asymptotic average SER (in terms of SNR) for MIMO MRT/MRC systems, based on which we verify the two observations made in [13]. We also compare the MIMO MRT/MRC system with two other widely deployed MIMO diversity schemes: one is joint space-time block coding and maximum ratio combining (STBC/MRC), and the other is selection combining at both ends (SC/SC). This comparison enables better understanding of MIMO diversity systems.

Our future work will include the analysis of the diversity scheme that employs hybrid selection scheme at both ends.

Chapter 3: Both incremental and decremental antenna selection rules with link adaptation are proposed for uncorrelated MIMO systems. Both rules are realized with recursive algorithms, thus greatly reducing the computational complexities and feasible for practical implementation. Rigorously speaking, neither rule provides the optimum solution, but the performance loss is negligible. For uncorrelated MIMO channels with independent and identically distributed (i.i.d) Rayleigh fading, we propose an antenna selection rule based on the expectation of the optimal number of active antennas. Based on this rule, the computational complexities can be further reduced while little performance degradation

would be incurred. Such computation reduction is especially prominent for large MIMO systems. For correlated MIMO channels, we propose an incremental antenna selection rule with link adaptation based on the slowly varying channel covariance information, which is also implemented in a recursive way to avoid the computational complexity of exhaustive search.

Our future work includes more detailed computational complexity analysis of the "incremental" and "decremental" selection rules. The theoretical study on the performance gap between out proposed methods and the optimal one is also of our interest.

Chapter 4: We derive explicit expressions for average (ergodic) capacity of joint spatial diversity and multiuser diversity systems when the number of users goes to infinity while the number of antennas keeps fixed. As expected, the average system capacity and scheduling gain grow with the number of users; and we contribute by providing a rather general asymptotic expression that builds an explicit connection with key system parameters and reveals their interactions, and by providing a strict proof of convergence that is in a stronger sense than what is assumed in previous study. As an application, we confirm that in this scenario, there is a tradeoff between spatial diversity and multiuser diversity for an open-loop spatial diversity system, but the detrimental effect of multiple transmit antennas can be avoided with the closed-loop schemes. We also show that all closed-loop schemes perform similarly in this scenario, in the sense that their differences only occur at the second-order (i.e., $\log\log K$). We show rigorously that the scheduling gain nonetheless diminishes to zero as the size of antenna arrays grows while the number of users keeps fixed, no matter for open-loop or closed-loop spatial diversity systems, through asymptotic study on the mean

and variance of the effective link SNR. In this sense, multiuser scheduling is not worthwhile in an antenna-dominant environment. On the other hand, different spatial diversity schemes do make significant difference with respect to system capacity for round robin scheduling. Since the scheduling gain asymptotically decreases with the number of antennas and increases with the number of users, it's interesting to study the asymptotic trend when both are allowed to grow. We reveal how the scheduling gain behaves depending on the relative growth rate between the two. In particular, we determine a critical point, only beyond which multiuser scheduling is meaningful.

Our future work includes study of the interaction between spatial diversity and multiuser diversity in a correlated fading scenario, and extension to the situations when users' channels are heterogeneous, together with the associated fairness issues. The interaction of multiuser diversity and the diversity-multiplexing tradeoff in MIMO systems also deserves further study.

Chapter 5: We provide a THP structure for multiuser MIMO, and based on this structure, we propose a multiuser scheduling law. The proposed scheme greatly reduces the scheduling complexity while simultaneously improves overall system performance.

In our work, there is an inherent limitation on the maximum number of users that can be simultaneously supported for our proposed scheme, due to the geometric structure revealed in Section 5.2. Therefore whether there exists an optimal number for the scheduled users, in terms of system throughput averaged over all channel realizations, remains open and constitutes our future work.

BIBLIOGRAPHY

[1] D. Aktas and H. El Gamal, "Multiuser scheduling for MIMO wireless systems," *Proc. 2003 IEEE 58th Vehicular Technology Conference*, vol. 3, pp. 1743-1747, Orlando, Oct. 2003.

[2] S.M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp.1451-1458, Oct. 1998.

[3] M. S. Alouini and A. Goldsmith, "Adaptive M-QAM modulation over Nakagami fading channels," in *Proc. Communication Theory Mini-Conf. (CTMC-VI) in conjuction with IEEE Global Communication Conf. GLOBECOM'97*, Phoenix, AZ, Nov. 1997, pp.218-223.

[4] D. Asztely, "On antenna arrays in mobile communication systems: fast fading and GSM base station receiver algorithms," *Tech. Rep. IR-S3-SB-9611*, Royal Inst. Technol., Stockholm, Sweden, Mar. 2002.

[5] Z. D. Bai, and Jack W. Silverstein, "CLT for linear spectral statistics of large-dimensional sample covariance matrices," *The Annals of Probability*, vol. 32, no. 1A, pp.553-605, 2004.

[6] H. Bolcskei, M. Borgmann and A.J. Paulraj, "Impact of the propagation environment on the performance of space-frequency coded MIMO-OFDM," *IEEE Journal on Selected Areas in Communications*, vol.21, no.3, pp. 427-439, April, 2003.

[7] G. Caire, and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no.7, pp. 1691-1706, June 2003.

[8] S. Catreux, V. Erceg, D. Gesbert and R. W. Heath, Jr., "Adaptive modulation and MIMO coding for broadband wireless data networks," *IEEE Communication Magazine*, vol. 40, no. 6, pp.108-115, June 2002.

[9] C.-J. Chen and L.-C. Wang, "A unified capacity analysis for wireless systems with joint antenna and multiuser diversity in Nakagami fading channels," *in Proc. International Conference on Communications*, vol.6, Paris, June 2004, pp.3523-3527.

[10] C.-N. Chuah, D. N. C. Tse, J. M. Kahn and R.A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Transactions on Information theory*, vol. 48, no.3, pp.637-650, Mar. 2002.

[11] J.M.Cioffi,"A multicarrier primer," *http://www.stanford.edu/group/cioffi/papers.html*.

[12] H. A. David and H. N. Nagaraja, *Order Statistics*, 3rd edition, New York: Wiley, 2003.

[13] P. A. Dighe, R. K. Mallik, and S. S. Jamuar, "Analysis of transmit-receive diversity in Rayleigh fading," *IEEE Transactions on Communications*, vol. 51, no. 4 , pp. 694 - 703, Apr. 2003.

[14] K. Young-Doo, K. Inhyoung, C. Jihoon, A. Jae-Young and Y.H. Lee, "Adaptive modulation for MIMO systems with V-BLAST detection," in *Proc. 57th IEEE Semiannual Vehicular Technology Conference*, Jeju, Korea, April 2003, vol. 2, pp. 22-25.

[15] A. Edelman, "Eigenvalues and condition numbers of random matrices," *Siam J. Matrix Anal. Appl.*, vol.9, no. 4, pp. 543-560, Oct. 1988.

[16] U. Erez and S. Brink, "Approaching the dirty paper limit for canceling known interference," *Proc. 41th Ann. Allerton Conf. On Commun., Control and Computing*, Monticello, IL, Oct. 1-3, 2003.

[17] G. D. Forney and M. V. Eyuboglu, "Combined equalization and coding using precoding," *IEEE Communications Magazine*, vol. 29 no. 12, pp. 25 -34, Dec. 1991.

[18] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, pp.41-59, Autumn 1996.

[19] 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, 1998.

[20] J. Galambos, *the Asymptotic Theory of Extreme Order Statistics*, 2nd edition, New York Wiley, 1978.

[21] D. Gesbert et al., "From theory to practice: An overview of MIMO space-time coded wireless systems", *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281-302, Apr. 2003.

[22] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels", *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684-702, June 2003.

[23] G. Golub and C. F. Van Loan, *Matrix computation,* John Hopkins University Press, Baltimore, MD, 3rd edition,1996.

[24] D. Gore, R. W. Heath, Jr. and Arogyaswami J. Paulraj, "Transmit selection in Spatial Mulplexing Systems," *IEEE Communications Letters*, vol. 6, no. 11, pp. 491-493, Nov. 2002.

[25] D. Gore, R. Heath and A. Paulraj, "Statistical antenna selection for spatial multiplexing systems," in *Proc. 2002 IEEE International Conference on Communications*, New York City, April, 2002, vol. 1, pp. 450-454.

[26] A. Gorokhov, D. A. Gore and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms," *IEEE Transactions on Signal Processing*, vol.51, no.11, Nov. 2003.

[27] R. Gozali, R. M. Buehrer and B. D. Woerner, "The impact of multiuser diversity on space-time block coding," *IEEE Communications Letters*, vol. 7, no. 5, pp. 213-215, May 2003.

[28] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd edition, Oxford University, 2001.

[29] A. Duel-Hallen, H. Shengquan, and H. Hallen, "Long-range prediction of fading signals," *IEEE Signal Processing Magazine*, vol. 17, no. 3, pp.62-75, May 2000.

[30] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '00)*, Istanbul, Turkey, June, 2000, vol. 2, pp. II737-II740.

[31] R. W. Heath, M. Airy, and A. J. Paulraj, "Multiuser diversity for MIMO wireless systems with linear receivers," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, pp. 1194–1199, Pacific Grove, CA, Nov. 2001.

[32] R. W. Heath Jr. and A.Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE International Conference on Communications (ICC'2001)*, Helsinki, Finland, June, 2001, vol. 7, pp. 2276-2280.

[33] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multi-antenna channel-hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inform. Theory*, vol. 50, pp. 1893-1909, Sept. 2004.

[34] J. Jiang, R. M. Buehrer and W. H. Tranter, "Antenna diversity in multiuser data networks," *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 490-497, Mar. 2004.

[35] I. M. Johnstone, "On the distribution of the largest eigenvalue in principal components analysis," *The Annals of Statistics*, vol. 29. no. 2, pp. 295-327, 2001.

[36] M. Kang, and M. S. Alouini, "Largest eigenvalue of complex Wishart matrices and performance analysis of MIMO MRC systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 418-426, Apr. 2003.

[37] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Stat.*, vol. 35, pp.1807-1810, Dec. 1964.

[38] R. Knopp and P. Humblet, "Information capacity and power control in single user cell multiuser communications," *in Proc. International Conference on Communications*, vol. 1, Seattle, WA, June 1995, pp. 331-335.

[39] E. G. Larsson, "On the combination of spatial diversity and multiuser diversity," *IEEE Communications Letters*, vol.8, no. 8, pp. 517-519, Aug. 2004.

[40] V. K. N. Lau, Y. Liu and T. A. Chen, "The role of transmit Diversity on wireless communications – Reverse link analysis with partial feedback," *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 2082-2090, Dec. 2002.

[41] J. Liu and A. Duel-Hallen, "Tomlinson-Harashima transmitter precoding for synchronous multiuser communications," the 37th Conference in Information Sciences and Systems, the John Hopkins Univ., Mar. 2003.

[42] T. K. Y. Lo, "Maximum ratio transmission," *IEEE. Trans. Commun.*, vol. 47, no. 10, pp. 1458-1461, Oct. 1999.

[43] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, 1982.

[44] C. Mun, et al, "Exact capacity analysis of multiuser diversity combined with transmit diversity," *Electronic Letters*, vol. 40, no. 22, pp. 1423-1424, Oct. 2004.

[45] A. F. Molisch, "MIMO systems with antenna selection-an overview", *IEEE Microwave Magazine*, vol. 5, no. 1, pp 46-56, Mar. 2004.

[46] A. Molisch, M. Win, and J. Winter, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE International Conference on Communications*, (Helsinki), June 2001, pp. 570-574.

[47] R. Nabar, H. Bolcskei, and A. Paulraj, "Transmit optimization for spatial multiplexing in the presence of spatial fading correlation," in *Proc. IEEE Globecom*, San Antonio, TX, Nov. 2001, pp.131-135.

[48] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2829- 2838, Nov. 2003.

[49] K-Wai Ng, R.S. Cheng and R. D. Murch, "A simplified bit allocation for V-BLAST based OFDM MIMO systems in frequency selective fading channels," in *Proc. 2002 IEEE International Conference on Communications*, New York City, NY, April 2002, vol. 1, pp. 411-415.

[50] A. J. Paulraj, D. A. Gore, R. U. Nabar and H. Bolcskei, "An overview of MIMO communications-A key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, no. 2, pp.198-218, Feb.2004.

[51] D. Shiu, G. J. Foschini, M. J. Gans and J. M. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, vol. 48, no.3, pp. 502-513, March 2000.

[52] S. Shakkottai, T. S. Rappaport and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications Magazine*, pp. 74-80, Oct. 2003.

[53] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Transactions on Information Theory*, vol.51, no.2, pp.506-522, Feb. 2005.

[54] M. Sharif and B. Hassibi, "A comparision of time-sharing, DPC, and beamforming for MIMO broadcast channels with many users," submitted to *IEEE Transactions on Communications.*

[55] M. K. Simon and M. S. Alouini, *Digital Communications over Fading Channels*, New York: John Wiley, 2000.

[56] Q. H. Spencer, and A. L. Swindlehurst, "Channel allocation in multi-user MIMO wireless communications systems," *Proc. 2004 IEEE International Conference on Communications*, vol. 5, pp. 3035-3039, Paris, June 2004.

[57] B. Suard, X. Guanghan, H. Liu, and T. Kailath, "Uplink channel capacity of space-division-multiple access schemes," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1468-1476, July 1998.

[58] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT&T Bell Labs Internal Tech. Memo.*, 1995.

[59] Z. Tu, and R. S. Blum, "Multiuser diversity for a dirty paper approach," *IEEE Communications letters*, vol. 7, no. 8, pp. 370-372, Aug. 2003.

[60] N. T. Uzgoren, "The asymptotic development of the distribution of the extreme values of a sample," *Studies in Mathematics and Mechanics Presented to Richard von Mises*. New York: Academic, 1954, pp. 346-353.

[61] M. K. Varanasi, "Group detection for synchronous Gaussian code-division multiple-access channels," *IEEE Transactions on Information Theory*, vol. 41, no. 4, pp.1083-1096, July 1995.

[62] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Transactions on Information Theory*, vol. 48, no. 6 , pp. 1277 -1294, June 2002.

[63] Z. Wang, and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels", *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1389-1397, Aug. 2003.

[64] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian MIMO broadcast channel," in Conference on Information Sciences and Systems (CISS), 2004.

[65] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximum ratio combining in Rayleigh fading," *IEEE Transactions on Communications*, vol. 47, no. 12, pp. 1773-1776, Dec. 1999.

[66] C. Windpassinger, et al., "Precoding in multiantenna and multiuser communications," *IEEE Transactions on Wireless Communications*, vol.3, no.4, pp. 1305-1316, July 2004.

[67] J. H. Winter, "On the capacity of radio communications systems with diversity in Rayleigh fading environments," *IEEE Journal of Selected Areas in Communications*, vol. 5, no. 5, pp. 871-878, June 1987.

[68] P. W. Wolniansky, G.J. Foschini, G.D. Golden and R. A. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. 1998 URSI International Symposium on Signals, systems, and Electronics* (ISSSE 98), Pisa, Italy, Oct.1998, pp. 295-300.

[69] Y. Q. Yin, Z. D. Bai, and P.R. Krishnaiah, "On the limit of the largest eigenvalue of the large dimensional sample covariance matrix," *Probability Theory and Related Fields*, vol. 78, pp. 509-521, 1988.

[70] W. Yu and W. Rhee, "Degrees of freedom in multiuser spatial multiplex systems with multiple antennas," submitted to *IEEE Transactions on Communications*.

[71] W. Yu, D. P. Varodayan and J. M. Cioffi, "Trellis and convolutional precoding for transmitter-based interference pre-subtraction," submitted to *IEEE Transactions on Communications*.

[72] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1250-1276, June 2002.

[73] R. Zhang, "Scheduling for maximum capacity in SDMA/TDMA systems," *Proc. 2002 ICASSP*, pp. III. 2141-III. 2144, 2002.

[74] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: a fundamental trade-off in multiple-antenna channels," IEEE Transactions on Information Theory, vol. 49, no. 5, pp.1073-1096, May 2003.

[75] Q. Zhou, H. Zhang and H. Dai, "Adaptive spatial multiplexing techniques for distributed MIMO systems, " in *Proc. 2004 Conference on Information Sciences and Systems* (CISS), Princeton University, Princeton, NJ, March 2004, pp.674-677.

[76] H. Zhuang, L. Dai, S. Zhou and Y. Yao, "Low complexity per-antenna rate and power control approach for closed-loop V-BLAST," *IEEE Transactions on Communications*, vol. 51, pp.1783-1787, Nov. 2003.

APPENDIX

APPENDIX A: Some proofs of chapter 2

Proof of Lemma 1: By Maclaurin Series expansion

$$\{\boldsymbol{\Psi}_c(x)\}_{i,j} = g(t-s+i+j-1,x) = \frac{1}{t-s+i+j-1}x^{t-s+i+j-1} + o(x^{t-s+i+j-1}), \quad \text{(A.1)}$$

we can obtain the approximation of $|\boldsymbol{\Psi}_c(x)|$ at $x = 0^+$ after some manipulation as

$$|\boldsymbol{\Psi}_c(x)| = |\boldsymbol{\Phi}| x^{MN} + o(x^{MN}), \quad \text{(A.2)}$$

with $\boldsymbol{\Phi}_{i,j} = 1/(t-s+i+j-1)$, for $i, j = 1,2,...,s$. The determinant of $\boldsymbol{\Phi}$ can be obtained in a similar fashion as that of a Hilbert matrix. After some algebra we get

$$|\boldsymbol{\Phi}| = \frac{\prod_{k=0}^{s-1}(k!)^2((l+k)!)^2}{\prod_{k=0}^{2s-1}(l+k)!}, F_g^{MRT/MRC}(x) = \frac{\prod_{k=0}^{s-1}k!}{\prod_{k=0}^{s-1}(t+k)!}x^{MN} + o(x^{MN}). \quad \text{(A.3)}$$

■

***Proof of Lemma 2*:** From $s_1 + t_1 > s_2 + t_2$, we can obtain $s_1 < s_2 < t_2 < t_1$. As

$$\begin{aligned} a^{(MRT/MRC)}(s_1,t_1) &= \frac{\prod_{k=0}^{s_1-1}k!}{\prod_{k=0}^{s_1-1}(t_1+k)!} \\ &= \frac{1}{1\times 2\times ...\times t_1}\mathbf{g}\frac{1}{2\times 3\times ...\times (t_1+1)}\mathbf{g}...\frac{1}{s_1\times ...\times (s_1+t_1-1)}, \end{aligned} \quad \text{(A.4)}$$

$$\begin{aligned} a^{(MRT/MRC)}(s_2,t_2) &= \frac{\prod_{k=0}^{s_2-1}k!}{\prod_{k=0}^{s_2-1}(t_2+k)!} \\ &= \frac{1}{1\times 2\times ...\times t_2}\mathbf{g}\frac{1}{2\times 3\times ...\times (t_2+1)}\mathbf{g}...\frac{1}{s_2\times ...\times (s_2+t_2-1)}. \end{aligned} \quad \text{(A.5)}$$

The proof of Lemma 2 is equivalent to show that

$$(1\times ...\times t_1)\times ...\times (s_1\times ...\times (s_1+t_1-1)) > (1\times ...\times t_2)\times ...\times (s_2\times ...\times (s_2+t_2-1)). \quad \text{(A.6)}$$

The left hand side of (A.6) can be rewritten as

$$1^{f(1)} \times 2^{f(2)} \times \ldots \times (s_1 + t_1 - 1)^{f(s_1+t_1-1)}, \tag{A.7}$$

with

$$f(i) = \begin{cases} i, & 1 \le i \le s_1 \\ s_1, & s_1 + 1 \le i \le t_1 \\ s_1 + t_1 - i, & t_1 + 1 \le i \le s_1 + t_1 - 1. \end{cases} \tag{A.8}$$

Similarly the right hand side of (A.6) is given by

$$1^{g(1)} \times 2^{g(2)} \times \ldots \times (s_2 + t_2 - 1)^{g(s_2+t_2-1)}, \tag{A.9}$$

with

$$g(i) = \begin{cases} i, & 1 \le i \le s_2 \\ s_2, & s_2 + 1 \le i \le t_2 \\ s_2 + t_2 - i, & t_2 + 1 \le i \le s_2 + t_2 - 1. \end{cases} \tag{A.10}$$

It is not difficult to get $\sum_{i=1}^{s_1+t_1-1} f(i) = s_1 \times t_1 = s_2 \times t_2 = \sum_{i=1}^{s_2+t_2-1} g(i)$. Therefore, after canceling out the same factors in (A.7) and (A.9), we can observe (A.7) is definitely larger than (A.9).

APPENDIX B: Some proofs of chapter 3

Lemma3. For two ordered real sequence $\{a_i\}_{i=1}^{i=n}$ and $\{b_i\}_{i=1}^{i=n}$ such that $a_1 \le a_2 \le a_3 \le .... \le a_n$ and $b_1 \le b_2 \le b_3 \le .... \le b_n$, if $c_1, c_2, ..., c_n$ is any permutation of $b_1, b_2, ..., b_n$, then $\sum_i a_i b_i \ge \sum_i a_i c_i \ge \sum_i a_i b_{n-i+1}$.

Proof: for any ordered multiplication, if $i > j$, $c_i \ge c_j$, consider

$$T = a_1 c_1 + a_2 c_2 + ... + a_j c_i + ... + a_i c_j + ... + a_n c_n ,$$

$$S = a_1 c_1 + a_2 c_2 + ... + a_j c_j + ... + a_i c_i + ... + a_n c_n ,$$

$$S - T = a_i c_i + a_j c_j - a_i c_j - a_j c_i = (a_i - a_j)(c_i - c_j) \ge 0 .$$

By induction, we can see if we sort b_i in an ascending order, the corresponding summation will be maximized.

Proof of Lemma4.a: From [23], we can see that matrix QR decomposition is related to Gram-Schmidt orthogonalization, hence it is not difficult to see that $\mathbf{R}(k+1)$ shares the same first k diagonal elements with $\mathbf{R}(k)$ and $\mathbf{Q}(k+1)$ shares the same first k columns with $\mathbf{Q}(k)$ while the $(k+1)-th$ one given by $\mathbf{Q}(:, k+1) = \mathbf{h} - \sum_{l=1}^{k} \mathbf{Q}(:, l)^H \mathbf{h} \mathbf{Q}(:, l)$.

Assume the QR decomposition of $\mathbf{H}^{(k)}$ is $\mathbf{Q}(k)\mathbf{R}(k)$, then

$$\det((\mathbf{H}^{(k)})^H \mathbf{H}^{(k)}) = \det(\mathbf{R}(k)^H \mathbf{R}(k)) = \prod_{i=1}^{k} |r_{i,i}|^2 . \qquad \text{(B.1)}$$

In the $(k+1)-th$ step, assume the QR decomposition of $\mathbf{H}^{(k+1)}$ is $\mathbf{Q}(k+1)\mathbf{R}(k+1)$, then

$$
\begin{aligned}
&\det((\mathbf{H}^{(k+1)})^H\mathbf{H}^{(k+1)}) = \det\left(\begin{bmatrix} (\mathbf{H}^{(k)})^H\mathbf{H}^{(k)} & \mathbf{H}^{(k)}\mathbf{h} \\ \mathbf{h}^H\mathbf{H}^{(k)} & \mathbf{h}^H\mathbf{h} \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} (\mathbf{H}^{(k)})^H\mathbf{H}^{(k)} & \mathbf{H}^{(k)}\mathbf{h} \\ \mathbf{0} & \mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{H}^{(k)}((\mathbf{H}^{(k)})^H\mathbf{H}^{(k)})^{-1}(\mathbf{H}^{(k)})^H\mathbf{h} \end{bmatrix}\right) \\
&= (\mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{H}^{(k)}((\mathbf{H}^{(k)})^H\mathbf{H}^{(k)})^{-1}(\mathbf{H}^{(k)})^H\mathbf{h})\det((\mathbf{H}^{(k)})^H\mathbf{H}^{(k)}) \\
&== \det(\mathbf{R}(k+1)^H\mathbf{R}(k+1))
\end{aligned}
\tag{B.2}
$$

hence

$$
\begin{aligned}
&\frac{\det(\mathbf{R}(k+1)^H\mathbf{R}(k+1))}{\det(\mathbf{R}(k)^H\mathbf{R}(k))} \\
&= \mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{H}^{(k)}((\mathbf{H}^{(k)})^H\mathbf{H}^{(k)})^{-1}(\mathbf{H}^{(k)})^H\mathbf{h} = \mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{Q}(k)\mathbf{Q}(k)^H\mathbf{h}.
\end{aligned}
\tag{B.3}
$$

Hence the amplitude of the $(k+1)-th$ diagonal element of $\mathbf{R}(k+1)$ is given by (B.3), which is

$$
\left|r_{k+1,k+1}\right| = \sqrt{\mathbf{h}^H\mathbf{h} - \mathbf{h}^H\mathbf{Q}(k)\mathbf{Q}(k)^H\mathbf{h}}. \tag{B.4}
$$

Proof of Lemma4.b: In the $(k+1)-th$ step, assume column $\mathbf{h}(k+1)$ is selected. Partition $\mathbf{Q}(k)$ as $\mathbf{Q}(k) = [\mathbf{Q}(k-1) \quad \mathbf{q}(k)]$, where $\mathbf{q}(k)$ is the rightmost column of $\mathbf{Q}(k)$. From (B.4), we get

$$
\begin{aligned}
\left|r_{k+1,k+1}\right|^2 &= \mathbf{h}^H(k+1)\mathbf{h}(k+1) - \mathbf{h}^H(k+1)\mathbf{Q}(k)\mathbf{Q}^H(k)\mathbf{h}(k+1) \\
&= \mathbf{h}^H(k+1)\mathbf{h}(k+1) - \mathbf{h}^H(k+1)\mathbf{Q}(k-1)\mathbf{Q}^H(k-1)\mathbf{h}(k+1) - |\mathbf{h}^H(k+1)\mathbf{q}(k)|^2 \\
&\le \mathbf{h}^H(k+1)\mathbf{h}(k+1) - \mathbf{h}^H(k+1)\mathbf{Q}(k-1)\mathbf{Q}^H(k-1)\mathbf{h}(k+1).
\end{aligned}
\tag{B.5}
$$

According to our selection criterion, we know in the $k-th$ step, $\mathbf{h}(k)$ is the selected column vector in the remaining of $\mathbf{H}$ such that $f(\mathbf{x}) = \mathbf{x}^H\mathbf{x} - \mathbf{x}^H\mathbf{Q}(k-1)\mathbf{Q}^H(k-1)\mathbf{x}$ is maximized. Thus,

$$\begin{aligned}\left|r_{k,k}\right|^2 &= \mathbf{h}^H(k)\mathbf{h}(k) - \mathbf{h}^H(k)\mathbf{Q}(k-1)\mathbf{Q}^H(k-1)\mathbf{h}(k) = \max f(\mathbf{h}(k)) \geq f(\mathbf{h}(k+1)) \\ &= \mathbf{h}^H(k+1)\mathbf{h}(k+1) - \mathbf{h}^H(k+1)\mathbf{Q}(k-1)\mathbf{Q}^H(k-1)\mathbf{h}(k+1) \geq \left|r_{k+1,k+1}\right|^2.\end{aligned} \tag{B.6}$$

Proof of Lemma 5.a: The Cholesky factor of $\mathbf{R}_T^{(k+1)} = \begin{bmatrix} \mathbf{R}_T^{(k)} & \mathbf{v} \\ \mathbf{v}^H & 1 \end{bmatrix}$ can be partitioned as [23]

$$\mathbf{R}(k+1) = \begin{bmatrix} \mathbf{R}(k) & \boldsymbol{\eta} \\ \mathbf{0}^H & r_{k+1,k+1} \end{bmatrix}. \tag{B.7}$$

where $\mathbf{R}(k)$ is the Cholesky factor of $\mathbf{R}_T^{(k)}$, $\boldsymbol{\eta}$ is a $k \times 1$ vector, $\mathbf{0}$ is a $k \times 1$ vector comprising all zeros elements, and $r_{k+1,k+1}$ is the scalar we are interested in. Writing out the Cholesky decomposition of $\mathbf{R}_T^{(k+1)}$:

$$\mathbf{R}_T^{(k+1)} = \begin{bmatrix} \mathbf{R}(k)^H\mathbf{R}(k) & \mathbf{R}(k)^H\boldsymbol{\eta} \\ \boldsymbol{\eta}^H\mathbf{R}(k) & r_{k+1,k+1}^2 + \boldsymbol{\eta}^H\boldsymbol{\eta} \end{bmatrix}, \tag{B.8}$$

we have

$$\begin{cases} \mathbf{R}(k)^H\boldsymbol{\eta} = \mathbf{v} \\ r_{k+1,k+1}^2 + \boldsymbol{\eta}^H\boldsymbol{\eta} = 1 \end{cases}. \tag{B.9}$$

From (B.9), we can get

$$r_{k+1,k+1} = \sqrt{1 - \boldsymbol{\eta}^H\boldsymbol{\eta}} = \sqrt{1 - \mathbf{v}^H(\mathbf{R}_T^{(k)})^{-1}\mathbf{v}}. \tag{B.10}$$

Proof of Lemma5.b: Assume $\mathbf{R}_T^{(k-1)}$, $\mathbf{R}_T^{(k)}$ and $\mathbf{R}_T^{(k+1)}$ are the covariance matrix for the $k-1$, k and $k+1$ selected antennas respectively. Also we assume $\mathbf{R}(k-1)$ and $\mathbf{R}(k)$ are the Cholesky factors of $\mathbf{R}_T^{(k-1)}$, $\mathbf{R}_T^{(k)}$ respectively. According to our recursive selection rule,

$\mathbf{R}_T^{(k+1)}$ is an "enhanced matrix" based on $\mathbf{R}_T^{(k)}$ and $\mathbf{R}_T^{(k)}$ is an "enhanced matrix" based on $\mathbf{R}_T^{(k-1)}$.

Assume $\mathbf{R}_T^{(k)} = \begin{bmatrix} \mathbf{R}_T^{(k-1)} & \mathbf{v}_{k-1} \\ \mathbf{v}_{k-1}^H & 1 \end{bmatrix}$, where $\mathbf{v}_{k-1}$ is the $(k-1)\times 1$ covariance vector between the $k-th$ selected antenna and the previous $(k-1)$ selected antennas. Then according to (B.10),

$$r_{k,k} = \sqrt{1-\mathbf{v}_{k-1}^H(\mathbf{R}_T^{(k-1)})^{-1}\mathbf{v}_{k-1}} = \sqrt{1-\mathbf{v}_{k-1}^H\mathbf{R}(k-1)^{-1}(\mathbf{R}(k-1)^{-1})^H\mathbf{v}_{k-1}}. \qquad \text{(B.11)}$$

Similarly,

$$r_{k+1,k+1} = \sqrt{1-\mathbf{v}_k^H\mathbf{R}(k)^{-1}(\mathbf{R}(k)^{-1})^H\mathbf{v}_k}. \qquad \text{(B.12)}$$

Note $\mathbf{R}(k)^{-1} = \begin{bmatrix} \mathbf{R}(k-1)^{-1} & \times \\ \mathbf{0}^H & \times \end{bmatrix}$, where "×" denotes irrelevant entries. Assuming the rightmost column of $\mathbf{R}(k)^{-1}$ is $\mathbf{b}_k$, then

$$\begin{aligned} r_{k+1,k+1} &= \sqrt{1-\mathbf{v}_k^H\mathbf{R}(k)^{-1}(\mathbf{R}(k)^{-1})^H\mathbf{v}_k} \\ &= \sqrt{1-\mathbf{v}_k^H\begin{bmatrix}\mathbf{R}(k-1)^{-1}\\ \mathbf{0}^H\end{bmatrix}\begin{bmatrix}(\mathbf{R}(k-1)^{-1})^H & \mathbf{0}\end{bmatrix}\mathbf{v}_k - \left|\mathbf{v}_k^H\mathbf{b}_k\right|^2} \\ &\le \sqrt{1-\mathbf{v}_k^H\begin{bmatrix}\mathbf{R}(k-1)^{-1}\\ \mathbf{0}^H\end{bmatrix}\begin{bmatrix}(\mathbf{R}(k-1)^{-1})^H & \mathbf{0}\end{bmatrix}\mathbf{v}_k}. \end{aligned} \qquad \text{(B.13)}$$

Assuming $\tilde{\mathbf{v}}_k$ is a $(k-1)\times 1$ vector truncated from $\mathbf{v}_k$ by discarding the last element of $\mathbf{v}_k$.

Thus (B.13) becomes $r_{k+1,k+1} \le \sqrt{1-\tilde{\mathbf{v}}_k^H(\mathbf{R}_T^{(k-1)})^{-1}\tilde{\mathbf{v}}_k}$. According to our selection rule, in the $k-th$ step, we choose the $(k-1)\times 1$ vector $\mathbf{v}_{k-1}$ such that $r_{k,k} = \sqrt{1-\mathbf{v}_{k-1}^H(\mathbf{R}_T^{(k-1)})^{-1}\mathbf{v}_{k-1}}$ is maximized, hence $r_{k+1,k+1} \le \sqrt{1-\tilde{\mathbf{v}}_k^H((\mathbf{R}_T^{(k-1)})^{-1}\tilde{\mathbf{v}}_k} \le r_{k,k}$.

APPENDIX C: Some proofs of chapter 4

Proof of Theorem 1:

1). Preliminary:

In order to prove Theorem 1, we first provide some preliminary results in [60] through the following lemma. Based on this lemma, a corollary follows, which is key to deriving a tight lower bound for $\bar{S}(K,M)$.

Lemma 7: Let $X_1,...,X_K$ be i.i.d. random variables as given in Lemma 1. If $\lim_{x\to\infty} g(x) = c \geq 0$, then the asymptotic expansion of $\log\left[-\log F^K\left(b_K + xg(b_K)\right)\right]$ at b_K is given by

$$\log\left[-\log F^K(b_K + xg(b_K))\right] = -x + \frac{x^2}{2!}g'(b_K) + \frac{x^3}{3!}\left[g(b_K)g^{(2)}(b_K) - 2\left(g'(b_K)\right)^2\right] ... + ... \\ + \frac{e^{-x} + ...}{2K} + \frac{5e^{-2x+...}}{24K^2} + ... - \frac{1}{8K^3}e^{-3x} + ... + ..., \tag{C.1}$$

where $b_K = F^{-1}(1-1/K)$. Furthermore if $\lim_{K\to\infty}[K \cdot g'(b_K)] = \infty$, the terms in the last group of (C.1) starting with the term $e^{-x}/2K$ are negligibly small compared to the terms in the first group.

Corollary 2: Let $X_1,...,X_K$ be i.i.d. random variables as given in Lemma 1. If $\lim_{x\to\infty} g(x) = c \geq 0$, $g'(x) = O(1/x^{d_1})$ with $d_1 > 0$, and $b_K = O\left((\log K)^{d_2}\right)$ with $0 < d_2 \leq 1$, then there exists a $k > 0$, such that

$$\mathrm{P}\left\{-k\log\log K \leq \left(\max_{1\leq k\leq K} X_k\right) - b_K \leq k\log\log K\right\} \geq 1 - O\left(\frac{1}{\log K}\right). \tag{C.2}$$

Remark: The proof follows readily from Lemma 4 by choosing $x = \log\log K$. This result has been demonstrated in [53] (see (A5)) with conditions $c > 0$ (in this case we can let $k = c$), $g^{(m)}(x) = O(1/x^m)$ and $b_K = O(\log K)$.

2). Tight Lower Bound for $\bar{S}(K,M)$:

Apply an extension of the Markov's inequality, we have (for sufficiently large K)

$$\begin{aligned}\bar{S}(K,M) &= E\left(\log\left(1+g_t\left(\max_{1\le k\le K} g_k\right)\right)\right)\\ &\ge P\left(\max_{1\le k\le K} g_k \ge b_K - k\log\log K\right)\times\log\left(1+g_t\left(b_K - k\log\log K\right)\right)\\ &\ge \left(1-O\left(\frac{1}{\log K}\right)\right)\times\log\left(1+g_t\left(b_K - k\log\log K\right)\right)\\ &\ge \log\left(1+g_t b_K\right) - o(1),\end{aligned} \tag{C.3}$$

where the second inequality follows from Corollary 2, and the last one follows from $b_K = O\left((\log K)^{d_2}\right)$ with $0 < d_2 \le 1$.

3). Tight Upper Bound for $\bar{S}(K,M)$:

As a final step, we provide an upper bounded for $\bar{S}(K,M)$, which coincides with the lower bound asymptotically. Let $S(K,M) = \log\left(1+g_t\left(\max_{1\le k\le K} g_k\right)\right)$, which is positive with probability 1, then

$$\begin{aligned}\bar{S}(K,M) &= \int_0^{\infty} P(S(K,M) > x)dx\\ &= \int_0^{\log(1+g_t b_K)} P(S(K,M) > x)dx + \int_{\log(1+g_t b_K)}^{+\infty} P(S(K,M) > x)dx\\ &\le \log\left(1+g_t b_K\right) + \int_{\log(1+g_t b_K)}^{+\infty} P(S(K,M) > x)dx.\end{aligned} \tag{C.4}$$

In the following, we show that the second term above diminishes as $K \to \infty$. First note

$$P(S(K,M) > x) = 1 - P(S(K,M) \le x) = 1 - F_g^K\left(\frac{e^x - 1}{g_t}\right). \tag{C.5}$$

Since $\lim_{x\to\infty}\frac{1-F_g(x)}{f_g(x)} = c \ge 0$, we can find positive constants c_2 and x_0, such that $1 - F_g(x) < c_2 f_g(x)$, when $x > x_0$. Thus for sufficiently large x we have

$$1-F_g^K\left(\frac{e^x-1}{g_t}\right)=\left(1-F_g\left(\frac{e^x-1}{g_t}\right)\right)\left(1+F_g\left(\frac{e^x-1}{g_t}\right)+\ldots+F_g^{K-1}\left(\frac{e^x-1}{g_t}\right)\right)$$
$$\le Kc_2 f_g\left(\frac{e^x-1}{g_t}\right). \tag{C.6}$$

Therefore for sufficiently large K and $b_K=O\left((\log K)^{d_2}\right)$,

$$\begin{aligned}&\int_{\log(1+g_t b_K)}^{+\infty} P(S(K,M)\ge x)dx \le \int_{\log(1+g_t b_K)}^{+\infty} Kc_2 f\left(\frac{e^x-1}{g_t}\right)dx\\ &=\int_{b_K}^{+\infty} Kc_2 f(x)\frac{g_t}{1+xg_t}dx \le K\frac{c_2 g_t}{1+g_t b_K}\int_{b_K}^{+\infty} f(x)dx\\ &=\frac{Kc_2 g_t}{1+g_t b_K}\left(1-F\left(b_K\right)\right)=O\left(\frac{1}{(\log K)^{d_2}}\right),\end{aligned} \tag{C.7}$$

where the second to last equality uses the fact $\left(1-F\left(b_K\right)\right)=1/K$.

Based on (C.3), (C.4) and (C.7) we can conclude

$$\lim_{K\to\infty}\left\{E\left(\log\left(1+g_t\left(\max_{1\le k\le K} g_K\right)\right)\right)-\log\left(1+g_t b_K\right)\right\}=0. \tag{C.8}$$

■

Proof of Corollary 1: First we can check that

$$\lim_{x\to\infty}\left(qvx^{v-1}\frac{1-F_g(x)}{f_g(x)}\right)=1, \tag{C.9}$$

when $f_g(x)\sim a\,x^p e^{-qx^v}$. This leads to the conclusion that $\lim_{x\to\infty} g(x)=c\ge 0$ ($c=0$ when $n>1$), and $g'(x)=O\left(\frac{1}{x^v}\right)$, therefore by Theorem 1, we are left to verify (4.9).

It can also be referred from (C.9) that $\lim_{x\to\infty}\frac{1-F(x)}{j\,(x)}=1$, where $j\,(x)=a\,x^{p+1-v}e^{-qx^{v}}/qv$.

Therefore, we only need to solve $j\,(b_K)=1/K$, i.e.,

$$b_K^v=\frac{1}{q}\log\frac{Ka}{qv}+\frac{p+1-v}{q}\log b_K. \tag{C.10}$$

The first order approximation for b_K is readily given by $\left(\frac{1}{q}\log Kt\right)^{1/v}$. To obtain the second order approximation, we just replace b_K on the right hand side of (C.10) with $\left(\frac{1}{q}\log Kt\right)^{1/v}$ to get

$$b_K^v=\frac{1}{q}\log\frac{Ka}{qv}+\frac{p+1-v}{vq}\log\log\frac{Ka}{qv}, \tag{C.11}$$

which leads to

$$\begin{aligned} b_K&=\left(\frac{1}{q}\log\frac{Ka}{qv}\right)^{1/v}\left(1+\frac{p+1-v}{qv}\frac{\log\log\frac{Ka}{qv}}{\frac{1}{q}\log\frac{Ka}{qv}}\right)^{\frac{1}{v}}=\left(\frac{1}{q}\log\frac{Ka}{qv}\right)^{1/v}\left(1+\frac{p+1-v}{qv^2}\frac{\log\log\frac{Ka}{qv}}{\frac{1}{q}\log\frac{Ka}{qv}}\right)\\ &=\left(\frac{1}{q}\log\frac{Ka}{qv}\right)^{1/v}+\frac{p+1-v}{qv^2}\frac{\log\log\frac{Ka}{qv}}{\left(\frac{1}{q}\log\frac{Ka}{qv}\right)^{(v-1)/v}}. \end{aligned} \tag{C.12}$$

■

Proof of Lemma 8: When $x\to+\infty$, $F_g^{MRT/MRC}(x)\to 1$, therefore

$$f_g^{MRT/MRC}(x)\sim\lim_{x\to\infty}tr(\Psi_c^{-1}(x)\Phi_c(x)). \tag{C.13}$$

As $\lim_{x\to\infty}\{\Psi_c(x)\}_{i,j}=\lim_{x\to\infty}g\,(t-s+i+j-1,x)=(t-s+i+j-2)!$, letting $l=t-s$, we have

$$\Psi_c(+\infty)=\begin{bmatrix} l! & (l+1)! & .. & (l+s-1)! \\ (l+1)! & & & \\ : & : & : & : \\ (l+s-1)! & .. & .. & (l+2s-2)! \end{bmatrix} \tag{C.14}$$

Meanwhile

$$\Phi_c(x)=\begin{bmatrix} x^l e^{-x} & x^{l+1}e^{-x} & .. & x^{l+s-1}e^{-x} \\ x^{l+1}e^{-x} & : & : & : \\ : & : & : & : \\ x^{l+s+1}e^{-x} & : & : & x^{l+2s-2}e^{-x} \end{bmatrix}=\begin{bmatrix} 1 & x & .. & x^{s-1} \\ x & : & : & : \\ : & : & : & : \\ x^{s-1} & : & : & x^{2s-2} \end{bmatrix} x^l e^{-x}. \tag{C.15}$$

Therefore

$$\begin{aligned} tr(\Psi_c^{-1}(+\infty)\Phi_c(x)) &= e^{-x}x^l[a_1x^{2s-2}+a_2x^{2s-3}+...+a_{2s-2}x+a_{2s-1}] \\ &= e^{-x}x^{l+2s-2}[a+O(1/x)], \end{aligned} \tag{C.16}$$

where the coefficients $\{a_i\}$ come from linear combinations of elements in $\Psi_c^{-1}(+\infty)$. We are only concerned with the dominant term, whose coefficient can be obtained as follows.

$$\begin{aligned} a = a_1 &= \frac{\left|\begin{pmatrix} l! & (l+1)! & ... & (l+s-2)! \\ (l+1)! & .. & .. & (l+s-1)! \\ : & : & : & : \\ (l+s-2)! & .. & .. & (l+2s-4)! \end{pmatrix}\right|}{\left|\Psi_c(+\infty)\right|} \\ &= \frac{\prod_{k=1}^{s-1}(t-k-1)!(s-k-1)!}{\prod_{k=1}^{s}(t-k)!(s-k)!} = \frac{1}{(t-1)!(s-1)!} = \frac{1}{(M-1)!(N-1)!}. \end{aligned} \tag{C.17}$$

■

Proof of Theorem 2: Let $X_M = \frac{g_M}{m_M}$, $\lim_{M\to\infty}\frac{s_M}{m_M}=0$ indicates that $X_M \xrightarrow{2} 1$ as $M\to\infty$.

First we show the weaker conclusion that the relative scheduling gain diminishes as $M\to\infty$. Fix $d>0$. By Markov's inequality, we have

$$E\left(\log(1+g_t g_M)\right) \geq P\left(X_M \geq (1-d)\right)\log\left(1+g_t(1-d)m_M\right), \tag{C.18}$$

which together with $X_M \xrightarrow{2} 1$ leads to

$$\liminf_{M\to\infty} \frac{E\left(\log(1+g_t g_M)\right)}{\log\left(1+g_t(1-d)m_M\right)} \geq \lim_{M\to\infty} P\left(X_M \geq (1-d)\right) = 1. \tag{C.19}$$

Now we let $d \to 0$ to get

$$\liminf_{M\to\infty} \frac{E\left(\log(1+g_t g_M)\right)}{\log\left(1+g_t m_M\right)} \geq 1. \tag{C.20}$$

On the other hand, by Jensen's Inequality $\limsup_{M\to\infty} \frac{E\left(\log(1+g_t g_M)\right)}{\log\left(1+g_t m_M\right)} \leq 1$, which completes the proof of the weaker conclusion.

It is known [12] that $E\left[g_{k^*}\right] \leq m_M + \frac{(K-1)s_M}{\sqrt{2K-1}}$. Using Jenson's inequality, we have

$$\bar{S}(K,M) \leq \log\left(1+g_t\left(m_M + \frac{(K-1)s_M}{\sqrt{2K-1}}\right)\right). \tag{C.21}$$

So

$$0 \leq \lim_{M\to\infty} \frac{G(K,M)}{\bar{R}(M)} \leq \lim_{M\to\infty} \frac{\log\left(1+g_t\left(m_M + \frac{(K-1)s_M}{\sqrt{2K-1}}\right)\right)}{\log\left(1+g_t m_M\right)} - 1 = 0. \tag{C.22}$$

We now turn to the proof of the convergence of the absolute scheduling gain. We can write

$$\log X_M = \log X_M \mathbf{I}_{(0,1)}\left(X_M\right) + \log X_M \mathbf{I}_{[1,\infty)}\left(X_M\right) = Y_M^{(1)} + Y_M^{(2)}. \tag{C.23}$$

First we have

$$0 \leq Y_M^{(2)} \leq (X_M - 1)\mathbf{I}_{[1,\infty)}\left(X_M\right). \tag{C.24}$$

Therefore $E\left(Y_M^{(2)}\right) \to 0$ since $E\left((X_M-1)\mathbf{I}_{[1,\infty)}\left(X_M\right)\right) \leq E\left(\left|X_M - 1\right|\right) \to 0$ as $M \to \infty$. In order to show that $E\left(Y_M^{(1)}\right) = E\left(\log X_M \mathbf{I}_{(0,1)}\left(X_M\right)\right) \to 0$, we make the following claim.

Claim 1: If a random variable $X_n \xrightarrow{P} 0$ as $n \to \infty$, $X_n I_E(X_n) \xrightarrow{P} 0$ for any event E.

This claim is easy to verify as $\forall e$, $P\left(\left|X_n I_E(X_n)\right| > e\right) \le P\left(\left|X_n\right| > e\right) \to 0$ as $n \to \infty$.

Now that $X_M \xrightarrow{2} 1$ we have $\log X_M \xrightarrow{P} 0$, as X_M is positive and the logarithm function is continuous. By claim 1 we in turn have $\log X_M \mathbf{I}_{(0,1)}\left(X_M\right) \xrightarrow{P} 0$. This together with the uniform integrability of $\left\{\log X_M \mathbf{I}_{(0,1)}\left(X_M\right)\right\}$ results in $Y_M^{(1)} \xrightarrow{1} 0$ [28], and it follows that $\lim_{M\to\infty} \bar{R}(M) = \log\left(1+g_t m_M\right)$. Hence

$$\begin{aligned} 0 \le \lim_{M\to\infty} G(K,M) &= \lim_{M\to\infty} \left(\bar{S}(K,M) - \bar{R}(M)\right) \\ &\le \lim_{M\to\infty} \left(\log\left(1+g_t\left(m_M + \frac{(K-1)s_M}{\sqrt{2K-1}}\right)\right) - \log\left(1+g_t m_M\right)\right) = 0. \end{aligned} \tag{C.25}$$

■

Proof of Lemma 9: Define $\mathbf{B} = \frac{1}{M}\mathbf{H}\mathbf{H}^H$, assume $b > \left(1+\sqrt{r}\right)^2$ and $l_{\max}(\mathbf{B}) = \left\|\frac{1}{\sqrt{M}}\mathbf{H}\right\|^2$.

Thus

$$E\left(l_{\max}(\mathbf{B})\right) = E\left(l_{\max}(\mathbf{B})\mathbf{I}_{[0,b]}\left(l_{\max}(\mathbf{B})\right)\right) + E\left(l_{\max}(\mathbf{B})\mathbf{I}_{(b,\infty)}\left(l_{\max}(\mathbf{B})\right)\right). \tag{C.26}$$

Using dominated convergence theorem together with Theorem 3.1 of [69], we can obtain

$$\lim_{M\to\infty} E\left(l_{\max}(\mathbf{B})\mathbf{I}_{[0,b]}\left(l_{\max}(\mathbf{B})\right)\right) = \left(1+\sqrt{r}\right)^2. \tag{C.27}$$

The remaining task is to show

$$\lim_{M\to\infty} E\left(l_{\max}(\mathbf{B})\mathbf{I}_{(b,\infty)}\left(l_{\max}(\mathbf{B})\right)\right) = 0. \tag{C.28}$$

We define two new $N \times M$ matrices $\mathbf{Y}$ and $\mathbf{Z}$ based on $\mathbf{H}$, with each entry of $\mathbf{Y}$ being $y_{i,j} = h_{i,j}\mathbf{I}_{\left[0,d\sqrt{M}\right]}\left(\left|h_{i,j}\right|\right)$ and each entry of $\mathbf{Z}$ being $z_{i,j} = h_{i,j}\mathbf{I}_{\left(d\sqrt{M},\infty\right)}\left(\left|h_{i,j}\right|\right)$, where the detailed

definition of $d = d_M \to 0$ can be found in [69] (see the proof Lemma 2.2) and [5] (see the discussion below 1.8). What we need in this proof is following two results from [5][69]:

$$\frac{1}{d^2} E\left(\left| h_{1,1} \right|^4 \mathrm{I}_{\left(d\sqrt{M}, \infty \right)} \left(\left| h_{1,1} \right| \right) \right) \to 0, \text{ as } M \to 0, \tag{C.29}$$

and

$$\sum_{M=1}^{\infty} E\left(\frac{l_{\max}\left(\mathbf{B}^{'} \right)}{b^{'}} \right)^k < \infty, \tag{C.30}$$

where $\mathbf{B}^{'} = \frac{1}{M}\left(\mathbf{Y} - E\left(y_{1,1} \right) \mathbf{1}_N \mathbf{1}_M^T \right)\left(\mathbf{Y} - E\left(y_{1,1} \right) \mathbf{1}_N \mathbf{1}_M^T \right)^H$ (where $\mathbf{1}_s$ is an $s \times 1$ column vector with all one entries), $b^{'}$ be a real number such that $\left(1 + \sqrt{r} \right)^2 < b^{'} < b$, and $k = k_M$ satisfies (4.3) and (4.4) of [69]. Further assume $\mathbf{B}^{''} = \frac{1}{M} \mathbf{Z}\mathbf{Z}^H$, With these definitions,

$$\begin{aligned}
& E\left(l_{\max}\left(\mathbf{B} \right) \mathbf{I}_{(b,\infty)} \left(l_{\max}\left(\mathbf{B} \right) \right) \right) \\
& = E\left(\left\| \frac{1}{\sqrt{M}} \left(\mathbf{Y} - E\left(y_{1,1} \right) \mathbf{1}_N \mathbf{1}_M^T \right) + \frac{1}{\sqrt{M}} E\left(y_{1,1} \right) \mathbf{1}_N \mathbf{1}_M^T + \frac{1}{\sqrt{M}} \mathbf{Z} \right\|^2 \mathbf{I}_{(b,\infty)} \left(l_{\max}\left(\mathbf{B} \right) \right) \right) \\
& \le 3\left[E\left(l_{\max}\left(\mathbf{B}^{'} \right) \mathbf{I}_{(b,\infty)} \left(l_{\max}\left(\mathbf{B} \right) \right) \right) + \frac{1}{M} \left| E\left(y_{1,1} \right) \right|^2 \left\| \mathbf{1}_N \mathbf{1}_M^T \right\|^2 + E\left(l_{\max}\left(\mathbf{B}^{''} \right) \right) \right].
\end{aligned} \tag{C.31}$$

The second term above admits

$$\begin{aligned}
& \left| E\left(y_{1,1} \right) \right|^2 \frac{1}{M} \left\| \mathbf{1}_N \mathbf{1}_M^T \right\|^2 = N \left| E\left(h_{1,1} \mathbf{I}_{\left[0, d\sqrt{M} \right]} \left(\left| h_{1,1} \right| \right) \right) \right|^2 = N \left| -E\left(h_{1,1} \mathbf{I}_{\left(d\sqrt{M}, \infty \right)} \left(\left| h_{1,1} \right| \right) \right) \right|^2 \\
& = \frac{N}{M} \frac{1}{d^2} \left| d\sqrt{M} E\left(h_{1,1} \mathbf{I}_{\left(d\sqrt{M}, \infty \right)} \left(\left| h_{1,1} \right| \right) \right) \right|^2 \le \frac{N}{M} \frac{1}{d^2} \left| E\left(\left| h_{1,1} \right|^2 \mathbf{I}_{\left(d\sqrt{M}, \infty \right)} \left(\left| h_{1,1} \right| \right) \right) \right|^2 \\
& \le \frac{N}{M} \frac{1}{d^2} E\left(\left| h_{1,1} \right|^4 \mathbf{I}_{\left(d\sqrt{M}, \infty \right)} \left(\left| h_{1,1} \right| \right) \right).
\end{aligned} \tag{C.32}$$

From (C.29), we know the above expression approaches 0 as M goes to infinity.

For the third term, we have

$$\begin{aligned} E\left(l_{\max}\left(\mathbf{B}^{''}\right)\right) &\le E\left(\frac{1}{M}\sum_{i,j}\left|z_{i,j}\right|^2\right) = NE\left(\left|h_{1,1}\right|^2 \mathbf{I}_{\left(d\sqrt{M},\infty\right)}\left(\left|h_{1,1}\right|\right)\right) \\ &\le \frac{N}{M}\frac{1}{d^2}E\left(\left|h_{1,1}\right|^4 \mathbf{I}_{\left(d\sqrt{M},\infty\right)}\left(\left|h_{1,1}\right|\right)\right). \end{aligned} \tag{C.33}$$

Therefore the above expression also approaches 0 as M goes to infinity.

Denote $a = \sqrt{b} - \sqrt{b'}$. We have

$$\mathbf{I}_{(b,\infty)}\left(l_{\max}\left(\mathbf{B}\right)\right) \le \mathbf{I}_{(b',\infty)}\left(l_{\max}\left(\mathbf{B}^{'}\right)\right) + \mathbf{I}_{(0,b']}\left(l_{\max}\left(\mathbf{B}^{'}\right)\right)\mathbf{I}_{(a,\infty)}\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right). \tag{C.34}$$

According to Markov's inequality, we have

$$\begin{aligned} E\left(\mathbf{I}_{(a,\infty)}\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right)\right) &= P\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)} > a\right) \\ &\le \frac{E\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right)}{a}. \end{aligned} \tag{C.35}$$

Furthermore

$$\begin{aligned} &\left|\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right| = \left\|\frac{1}{\sqrt{M}}\mathbf{H}\right\| - \left\|\frac{1}{\sqrt{M}}\left(\mathbf{Y} - E\left(y_{1,1}\right)\mathbf{1}_N\mathbf{1}_M^T\right)\right\| \\ &\le \frac{1}{\sqrt{M}}\left\|\left(\mathbf{H}-\mathbf{Y}\right)\right\| + \frac{1}{\sqrt{M}}\left\|E\left(y_{1,1}\right)\mathbf{1}_N\mathbf{1}_M^T\right\| = \frac{1}{\sqrt{M}}\left\|\mathbf{Z}\right\| + \left\|\frac{1}{\sqrt{M}}E\left(y_{1,1}\right)\mathbf{1}_N\mathbf{1}_M^T\right\|. \end{aligned} \tag{C.36}$$

Therefore

$$\begin{aligned} &E\left(l_{\max}\left(\mathbf{B}'\right)\mathbf{I}_{(0,b']}\left(l_{\max}\left(\mathbf{B}^{'}\right)\right)\mathbf{I}_{(a,\infty)}\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right)\right) \\ &\le b'E\left(\mathbf{I}_{(a,\infty)}\left(\sqrt{l_{\max}\left(\mathbf{B}\right)} - \sqrt{l_{\max}\left(\mathbf{B}'\right)}\right)\right) \le \frac{b'}{a}\left(\frac{1}{\sqrt{M}}\left\|\mathbf{Z}\right\| + \left\|\frac{1}{\sqrt{M}}E\left(y_{1,1}\right)\mathbf{1}_N\mathbf{1}_M^T\right\|\right) \\ &\le \frac{2b'}{a}\left(\frac{N}{M}\right)^{1/2}\frac{1}{d}E^{1/2}\left(\left|h_{1,1}\right|^4 I_{\left(d\sqrt{M},\infty\right)}\right). \end{aligned} \tag{C.37}$$

Again by (C.29), the above expression also approaches 0 as M goes to infinity. Finally we have

$$E\left(l_{\max}(\mathbf{B}')\mathbf{I}_{(b',\infty)}\left(l_{\max}(\mathbf{B}')\right)\right)=b'E\left(\frac{l_{\max}(\mathbf{B}')}{b'}\mathbf{I}_{(b',\infty)}\left(l_{\max}(\mathbf{B}')\right)\right)$$
$$\le b'E\left(\left(\frac{l_{\max}(\mathbf{B}')}{b'}\right)^{k}\mathbf{I}_{(b',\infty)}\left(l_{\max}(\mathbf{B}')\right)\right)\le b'E\left(\frac{l_{\max}(\mathbf{B}')}{b'}\right)^{k}, \quad \text{(C.38)}$$

By (C.30), we have

$$\lim_{M\to\infty}E\left(\frac{l_{\max}(\mathbf{B}')}{b'}\right)^{k}=0. \quad \text{(C.39)}$$

Therefore from (C.34)-(C.39), we can obtain

$$\lim_{M\to\infty}E\left(l_{\max}(\mathbf{B}')\mathbf{I}_{(b,\infty)}\left(l_{\max}(\mathbf{B})\right)\right)=0. \quad \text{(C.40)}$$

From (C.31), (C.32), (C.33) and (C.40), we can get $\lim_{M\to\infty}E\left(l_{\max}(\mathbf{B})\mathbf{I}_{(b,\infty)}\left(l_{\max}(\mathbf{B})\right)\right)=0$.

Thus

$$\lim_{M\to\infty}E\left(l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^{H}\right)\right)=\left(1+\sqrt{r}\right)^{2}. \quad \text{(C.41)}$$

In a similar fashion discussed above, we can show $\lim_{M\to\infty}E\left(l_{\max}^{2}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^{H}\right)\right)=\left(1+\sqrt{r}\right)^{4}$,

therefore

$$\lim_{M\to\infty}s^{2}\left(l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^{H}\right)\right)=\lim_{M\to\infty}\left(E\left(l_{\max}^{2}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^{H}\right)\right)-E^{2}\left(l_{\max}\left(\frac{1}{M}\mathbf{H}\mathbf{H}^{H}\right)\right)\right)=0. \quad \text{(C.42)}$$

■