Partial Channel Knowledge Based Precoding for MIMO and Cooperative Communications

By

Hamid Reza Bahrami

Department of Electrical and Computer Engineering
McGill University, Montreal

December 2007

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Your file Votre référence
ISBN: 978-0-494-50772-8
Our file Notre référence
ISBN: 978-0-494-50772-8

Abstract

Linear precoding can be viewed as a convenient way to enhance the throughput and performance of a Multi-Input-Multi-Output (MIMO) communications system for its linear nature that facilitates its low-complexity implementation in a transceiver. Nevertheless, linear precoding designs usually assume instantaneous channel responses perfectly known at the transmitter, which is unrealistically difficult (if not impossible) in fast time-varying channels.

An alternative to the assumption of full channel state knowledge is to consider the availability of partial channel knowledge at the transmitter, including information that changes much slower than the instantaneous channel responses such as the channel statistical parameters. The work presented in this thesis focuses on the design of precoding schemes that are mainly based on partial channel information.

Starting with point-to-point MIMO systems, we show that it is possible to design suitable precoding schemes that achieve a considerable gain in terms of performance or throughput by considering the *spatial* and *path* correlation matrices of the frequency-flat and frequency-selective fading channels, respectively.

Next, we investigate the problem of transmission in point-to-multipoint MIMO systems based on partial channel information. We show that in a MIMO broadcast system, using a partial channel knowledge-based user selection scheme in conjunction with precoding can provide an asymptotic optimum sum-rate performance for a growing number of users.

Using the results obtained in the previous steps, we consider the problem of transmission in cooperative relay networks. We first examine and identify the similarities and differences between MIMO and cooperative relay systems especially in term of diversity and multiplexing gain and their trade-off. We then develop the possible solutions for transmission and reception in the relay networks for both cases: single-antenna and multiple-antenna relay nodes. When a large number of single-antenna relay nodes available in the system, we show that a partial channel-knowledge based relay selection scheme in conjunction with a distributed BLAST transmission scheme can achieve the optimum multiplexing-diversity trade-off. In the case of multi-antenna relay,

we first derive the optimum combining scheme at the destination for both amplify-and-forward (AF) and decode-and-forward (DF) relay systems, and then address the optimal precoding designs at the source and relay nodes. We show that a generalized maximum ratio combining (GMRC) in conjunction with linear precoding can offer the optimum received SNR for both AF and DF relay systems.

Sommaire

Le précodage linéaire peut être vu comme une façon commode d'améliorer le débit et la performance d'un système de communication à *entrée multiple sortie multiple* (MIMO) par sa nature linéaire qui facilite son implémentation peu complexe dans un émetteur-récepteur. Néanmoins, des conceptions de précodage linéaire habituellement des réponses de canaux instantanées parfaitement connues à l'émetteur, qui sont d'un niveau de difficulté pratiquement irréalisable (voire impossible) dans des canaux rapides à temps variables.

Une alternative à l'hypothèse de connaissance d'un état plein canal est de considérer la disponibilité de connaissance de canal partiel à l'émetteur, y compris les informations qui changent beaucoup plus lentement que les réponses de canaux instantanés comme les paramètres statistiques du canal. Le travail présenté dans cette thèse met l'accent sur la conception de schémas de précodage qui sont basés principalement sur l'information de canal partielle.

En commençant par un système MIMO point-à-point, nous démontrons qu'il est possible de concevoir des schémas de précodage appropriés qui réalisent un gain considérable en termes de performance ou de débit en considérant les *matrices de corrélation spatiale et de chemin d'accès*[1] des canaux à évanouissement de fréquence en palier et de fréquence sélective, respectivement.

Ensuite, nous recherchons le problème de transmission dans les systèmes MIMO point-à-multipoint basé sur l'information de canal partielle. Nous démontrons qu'un système de diffusion MIMO, utilisant un plan de sélection d'utilisateurs à base de connaissances de canal partiel en conjonction avec le précodage peut fournir une performance de taux de calcul asymptotique optimum pour un nombre grandissant d'utilisateurs.

Utilisant les résultats obtenus dans les étapes précédentes, nous considérons le problème de transmission dans des réseaux de relais coopératifs. Nous examinons et identifions, tout d'abord, les similarités et les différences entre les systèmes MIMO et les

[1] the *spatial* and *path* correlation matrices

réseaux de relais coopératifs plus particulièrement en termes de diversité et de gain de multiplexage et de leur compensation. Nous développons alors les solutions possibles pour la transmission et la réception dans les réseaux de relais pour les deux cas : les nœuds de relais à antenne unique, et à antenne multiple. Quand un grand nombre de noeuds de relais d'antenne unique est disponible dans le système, nous démontrons qu'un plan de sélection de relais à base de connaissances de canal partiel en conjonction avec un plan de transmission réparti BLAST peut réaliser la compensation de multiplexage-diversité optimum. Dans le cas d'un relais d'antenne multiple, nous dérivons tout d'abord le schéma combinant optimum à la destination, pour à la fois, les systèmes de relais AF[2] (amplifier et faire parvenir) et DF[3] (décoder et faire parvenir), et ensuite la conception de précodage optimum à la source et aux nœuds de relais. Nous démontrons qu'une *combinaison à rapport maximum généralisée* (GMRC)[4] en conjonction avec du précodage linéaire peut offrir le *rapport signal sur bruit* (SNR)[5] optimum reçu pour, à la fois, les systèmes de relais AF et DF.

[2] (AF) amplify-and-forward
[3] (DF) decode-and-forward
[4] (GMRC) generalized maximum ratio combining
[5] (SNR) signal to noise ratio

Acknowledgements

I am in debt to my advisor, Professor Tho Le-Ngoc for his precious guidance that provided me with invaluable insight into the research issues. What I have learned from him is not indeed limited to research issues; I have also been inspired by his motivation, enthusiasm and perseverance necessary in academic work.

I am thankful to NSERC CRD grant with InterDigital Canada for partially supporting this research. I am also sincerely grateful to the members of my Ph.D. advisory committee, Professor Harry Leib and Professor Xiao-Wen Chang whose productive and in-depth inputs helped me a lot to complete this thesis. I would also like to thank Professor Amir K. Khandani, Professor Fabrice Labeau, Professor Ioannis Psaromiligkos and Professor Sofiene Affes for their precise comments that significantly improved the quality of final manuscript.

I would like to thank my friends in the Wireless Communications Laboratory at McGill University. A part of this thesis is the result of long and fruitful discussions with them especially during weekly meetings we had regularly during these four years.

Besides academic life, I have had wonderful moments with my Iranian friends in Montreal. Being with them, I have learned that one's contentment is the direct sum of his friends' happiness. I would not even have the chance to see a few of them at least some time soon; but they will be in my heart and mind for the rest of my life. I wish them a happy and blessed life which they truly deserve.

Nothing can express my gratitude to my parents and sister for their sincere support during some ten years of my academic career. Without their empathy, I could not be at this point to write these lines. I am also in debt to my father-in-law, mother-in-law and brothers-in-law for all their sympathy and assistances during the last months of my study. They have been the peace of mind for me to smoothly finish this thesis.

Last but not least, I am deeply grateful to my wife, Hanieh, who has been the source of inspiration and motivation for me during the last months of my Ph.D. work. Her understanding, patience and encouragement that helped me finish the remaining tasks in a timely manner, mean a lot to me.

Contents

List of Figures

List of Symbols

^{H}	Conjugate transpose (Hermitian)
^{T}	Transposition
$\otimes$	Kronecker product
$\lVert \mathbf{A} \rVert$	Euclidean norm of **A**
$\langle \mathbf{a} , \mathbf{b} \rangle$	Scalar (inner) product of vectors **a** and **b**
$\mathbf{x}$	Transmitted vector
$\mathbf{y}$	Received vector
$\mathbf{n}$	Noise vector
$\mathbf{H}$	Channel matrix
M	Number of transmit antenna
N	Number of receive antenna
L	Number of channel paths
$\mathbf{R}_T$	Transmit correlation matrix
$\mathbf{R}_R$	Receive correlation matrix
$\mathbf{R}_P$	Path correlation matrix
$\mathbf{W}$	Precoding matrix
E_s	Signal power
N_0	Noise power density
$\overline{\mathbf{H}}$	Composite channel matrix for frequency-selective channel
$\mathbf{H}_l$	Channel matrix for the l^{th} path
$\mathbf{R}_{T,l}$	Transmit correlation matrix of path l
$\mathbf{R}_{R,l}$	Receive correlation matrix of path l

$\rho_{m,k}$	Spatial correlation coefficient between antenna *m* and *k*
$\rho^{P}_{i,j}$	Path correlation coefficient between paths *i* and *j*
$\mathbf{W}_T$	Precoding matrix for frequency-selective channel
$\mathbf{W}_R$	Decoding matrix
$\mathbf{I}(\mathbf{x},\mathbf{y})$	Mutual information between **x** and **y**
P_l	Channel power of path *l*
$\mathbf{P}$	path power matrix
μ	KKT power coefficient
n	Number of available user in a multi-user system
N_i	Number of antenna of user *i*
$\mathbf{s}_i$	Transmit vector for user *i*
$\mathbf{y}_i$	Receive vector at user *i* terminal
$\mathbf{H}_i$	Channel matrix of user *i*
$\mathbf{W}_i$	Precoding matrix for user *i*
s_i	Transmit symbol for user *i* with single antenna
y_i	Receive symbol at user *i* terminal with single antenna
$\mathbf{h}_i$	Channel vector for user *i* with single antenna
$\mathbf{w}_i$	Precoding vector for user *i* with single antenna
P	Total available power at the transmitter
P_i	Power allocated to user *i*
C_{sum}	Sum rate of users in a multi-user system
R	Achievable rate
$\mathbf{U}$	Orthonormal basis
$\varepsilon(\mathbf{a},\mathbf{b})$	Orthogonality measure between vectors **a** and **b**
$F_X(x)$	CDF of X
$f_X(x)$	PDF of X
$\chi^2(k)$	χ^2 distribution with *k* degree of freedom
K	Number of active relay nodes in a relay system
l_1	Length of the first transmit interval in relay transmission
l_2	Length of the second transmit interval in relay transmission

$\mathbf{Y}_1$	Receive symbol matrix in the first transmit interval
$\mathbf{N}_1$	Noise matrix in the first transmit interval
$\mathbf{Y}_2$	Receive symbol matrix in the second transmit interval
$\mathbf{N}_2$	Noise matrix in the second transmit interval
$\mathbf{h}_0$	Channel vector between source and destination in a relay system
$\mathbf{h}_i$	Channel vector between relay i and destination in a relay system
g_i	Channel gain between source and relay i in a relay system
β_i	Repetition gain of relay i
r	Multiplexing gain
$d(r)$	Diversity gain at given multiplexing gain of r
x	Transmitted symbol from source
$\widetilde{x}$	Transmitted symbol from relay
y_1	Receive vector in the first symbol interval
y_2	Receive vector in the second symbol interval
$\mathbf{n}_1$	Noise vector in the first symbol interval
$\mathbf{n}_2$	Noise vector in the second symbol interval
$\mathbf{w}_1$	Source precoding vector
$\mathbf{w}_2$	Relay precoding vector
$\tilde{\mathbf{w}}_1$	Combining vector in the first symbol interval
$\tilde{\mathbf{w}}_2$	Combining vector in the second symbol interval
$\mathbf{W}_2$	Relay precoding matrix in AF mode
$\lambda_{\max}(\mathbf{A})$	Largest eigenvalue of $\mathbf{A}$
$\mathbf{u}_{\max}(\mathbf{A})$	Eigenvector of $\mathbf{A}$ corresponding to its largest eigenvalue ($\lambda_{\max}(\mathbf{A})$)

Chapter 1

Introduction

While high data rates and reliability are the key requirements of broadband wireless communication systems, limited spectral resources and the complex nature of the radio propagation channels presently hinder their development and call for advanced multi-dimensional signal processing and transmission techniques such as space-time and multi-input multi-output (MIMO) schemes to provide interference cancellation, diversity or multiplexing gain. A natural question in a MIMO system is how it is possible to achieve the capacity or performance (bit error rate) per transmit and receive antenna pairs equal or close to that of a single-input single-output (SISO) channel. One way to achieve this goal is to decouple the channel into some independent subchannels. It is usually referred as diagonalization of MIMO channel matrix. To explain it more, let us consider a simple illustrative example. Assume that both the transmitter and receiver are equipped with M antennas. Therefore, there are M^2 channel links between transmit and receive antennas. Let us further assume each transmit antenna transmits an independent signal. The signal received at each element of receive antenna is a linear combination of all signals transmitted by M transmit antennas. We are interested in separating M independent transmitted signals at the receiver. A convenient way to separate these M independent signals is to construct M independent parallel sub-channels from available M^2 links. In this case, the channel is decoupled to M parallel SISO channels and therefore, it is easily possible to achieve the capacity or performance per transmit and receive antenna close or equal to that if SISO channel. As the above MIMO channel can be modeled as an $M \times M$ matrix whose entries are the channel responses between each pair of transmit and receive antenna, this approach in fact leads to orthogonalization of the channel matrix.

On the other hand, due to the dynamic nature of transmission media, an adaptive scheme that adapts to channel variations is of great significance in order to increase the link performance. Channel-state information at the transmitter (CSIT) can be used in the design of the transmission schemes for possible further system performance improvement or/and complexity reduction. The term "channel state information" can refer to either the instantaneous channel response, or channel parameter(s), or channel statistics. Channel state information can be estimated at the receiver and fed back to the transmitter (assuming the feedback delay is much less than the channel coherence time).

Linear precoding is an approach that makes use of the available channel information at the transmitter to remove the inter-relation between MIMO links and hence construct independent parallel channels. A linear precoder can be viewed as a matrix with complex elements that is multiplied by a vector of input data and can result in the diagonalization of channel matrix. In this sense, there is an analogy between precoding in MIMO channels and pre-equalization in SISO inter symbol interference (ISI) channels. While a pre-equalizer tries to remove/reduce the interference between subsequent symbols, a MIMO precoder aims to eliminate the inter-relation between sub-channels in a MIMO channel. Pre-equalization can convert SISO ISI channel into an additive white Gaussian noise (AWGN) channel and hence the performance of the system after pre-equalization is equal to that of AWGN channel while precoder transforms the channel into some independent parallel channels, each of which can provide a performance equal to that of a SISO link.

Linear precoding and decoding schemes are scalable to any number of antennas and are simpler to implement when compared to nonlinear schemes [1]. Detection (decoding) of precoding schemes can be done using linear processing at the receiver and hence is simpler, which makes it suitable for the application where receiver volume and complexity is important (mobile units). Precoding can be applied to both uncoded and coded systems.

Design of linear precoders has usually assumed perfect knowledge of instantaneous channel response at the transmitter to improve the performance (i.e., higher information rates or lower bit error rates) [1], [2], [19]. In fast time-varying channels, perfect channel knowledge (of instantaneous channel response) at the transmitter is very difficult and hence is not a realistic assumption. At first, it requires accurate channel estimation at the receiver. Secondly, even if channel estimation at the receiver is extremely accurate, the delay can make feedback information out-dated when it is applied to the precoding.

It is more reasonable to assume that transmitter only has partial channel knowledge, for example, in terms of transmit and receive correlation matrices. For this reason, linear precoding designs based on partial channel information that is slowly time-varying, are desired and more suitable to practical applications. In this case, the transmitter can easily achieve long-term information of the channel via a low-speed feedback channel from receiver. The performance gain of such schemes would not be as high as full channel knowledge schemes, yet their simplicity and feasibility makes them appealing to use in MIMO wireless links.

As correlation is a long-term statistical parameter of the channel, it can be easily estimated and tracked (by receiver and/or transmitter). On the other hand, correlation (in space or path …) has been proved to be a main source of performance degradation in MIMO communication systems [20], [21], [42]. Having the correlation information at the transmitter, it is possible to design the schemes that aim to improve the system performance and/or throughput. Our stress in this thesis is therefore on the precoder schemes that uses spatial and path correlation information at the transmitter to improve the system performance.

As briefly pointed out above, partial channel knowledge covers a wide range of information that can be obtained at the transmitter and it is not necessarily limited to correlation. For example there are different precoder designs based on channel mean knowledge (precoding on mean CSIT) [100]-[104]. Also, it is possible to consider the availability of both channel mean and correlation information at the transmitter (precoding on both mean and correlation CSIT) [105], [106]. Other types of partial CSIT include channel K (Ricean) factor [109], channel condition number [110] and phase distribution [109].

In this thesis, the key idea is to use partial channel information in designing suitable transmission schemes in MIMO systems. More specifically, we propose to mainly investigate the channel statistical properties rather than the actual channel time-varying space-frequency responses and use them as partial channel knowledge at the transmitter for developing MIMO precoding schemes in single-user and multi-user environments. Partial channel information can refer to a wide range of channel parameters such as matrix channel condition number, number of channel paths or statistics such as channel mean, correlation in transmit and receive space, path, and sub-carrier. In this research work, we will first concentrate on the use of spatial and path correlations for precoder design as they have been proved to be very important in the performance and capacity of a MIMO link. We consider the precoder design based on partial channel knowledge for different frequency-flat and frequency-selective fading scenarios.

We first develop precoder designs for *point-to-point* transmission in both frequency-flat and frequency-selective channels. Next, we consider the problem of *point-to-multipoint* transmission and investigate the design of suitable transmission schemes for broadcast MIMO systems. In the last step, we focus on the scenario of *cooperative relay networks* in which one or some relay nodes help transmitter to send its information to the receiver. We propose precoder structures that can be applied to amplify-and-forward (AF) and decode-and-forward (DF) protocols.

1.1. Precoder Designs for Point-To-Point Transmission

In the case of point-to-point transmission, we consider MIMO precoder designs for frequency-flat and frequency-selective channels using *spatial* and *path* correlation matrices to be presented in Chapters 2 and 3.

As mentioned earlier in this chapter, design of linear precoders (precoding matrices or beamformers) usually assumes full channel knowledge at the transmitter and aims to improve performance (i.e., higher information rates or lower bit error rates) by optimal allocation of resources such as power and bits over multiple antennas, based on the channel properties [1]- [3], [6], [19]. Different performance criteria can be considered in the design of precoders. The most common and important criteria, however, include system capacity, error exponent, pairwise error probability (PEP) and detection mean square error (MSE). In [19], linear precoders were designed under average and peak power constraints, by using minimum mean squared-error (MMSE) and bit error rate (BER) criteria. All those cases can be viewed as a weighted MMSE criterion with a proper set of weight coefficients as proved in [1]. It was shown that the optimum linear precoder and decoder usually decompose the MIMO channel into its eigen sub-channels (or eigen-modes) and allocate power on these sub-channels according to a water-pouring (water-filling) strategy. In other words, the linear precoder at the transmitter beamforms into the eigen-modes of the channel and hence, is also called the eigen-beamformer in the literature. The number of eigen-modes is determined by the rank of the channel matrix.

All of the above schemes assume full channel knowledge or, equivalently, perfect channel state information (CSI) at both the transmitter and the receiver. In practice, however, perfect CSI is rarely available at the transmitter [2], [3]. Hence, it is more reasonable to assume that transmitter only has partial channel knowledge, for example, in terms of transmit and receive correlation matrices. In this case, the transmitter can achieve long-term information of the channel via a low-speed feedback channel from receiver. Knowing the transmit correlation matrix and assuming an identity receive correlation matrix, an optimal linear precoder was designed based on the pair-wise error probability (PEP) criterion in [3] and [6]. Using capacity criterion and knowledge of the channel correlation matrix at the transmitter, [7] proposed a precoder for multiple-input-single-output (MISO) channels. In [8], the eigen-decomposition of the average MIMO channel has been used and thereupon implemented a water-filling approach across the eigen-modes of the correlation channel matrix with i.i.d. rows and correlated columns (i.e., identity receive correlation matrix).

The assumption of an identity receive correlation matrix is based on the rich scattering environment near mobile units in a wireless link. However, such assumed environmental conditions are not common in wireless uplink channels, and therefore the receive correlation matrix cannot be always neglected.

In [9], based on the knowledge of transmit and receive correlation at the transmitter, a precoder structure was designed for a predetermined orthogonal space-time block code (STBC). In [10], the so-called Kronecker product model in vector form was used to derive the closed-form expression of a precoder for a coded multi-input, single-output (MISO) system. Due to the channel model used in these papers, the effects of transmit and receive correlation matrices could not be seen separately. Based on PEP criterion, [11] and [12] presented precoder designs for orthogonal STBC using a MIMO channel model that allows the consideration of different receive correlation matrices for each of the transmit antennas.

Based on capacity criterion, [13] derived a precoder for a 2×2 correlated MIMO system. In [14], another precoder design based on capacity criterion aims to find an optimal water-pouring policy based on capacity regions. The final result is not an explicit formula based on transmit and receive correlation matrices.

Chapter 2 of the thesis considers precoder designs based on both transmit and receive correlation matrices and derives the structures for optimal linear precoders under three criteria: minimum pairwise error probability (PEP), minimum mean squared-error (MMSE), and maximum ergodic (mean) channel capacity.

In Section 2.1, we study a MIMO channel model that includes transmit and receive correlation matrices in its structure. This model is well-known as Kronecker model in the literature. The Kronecker correlation model has been experimentally verified in indoor environments for up to 3×3 antenna configurations [96], [97] and in outdoor environments for up to 8×8 configurations [98], [99]. The key idea in Kronecker model is that all receive antenna elements "see" the same transmit correlation matrix. Other more general channel models are also available that assume different transmit correlation matrices for each receive antenna elements [107], [108]. Here, mainly due to practical considerations, we however confine ourselves to the Kronecker channel model.

In Section 2.2 and under the maximum ergodic (mean) channel capacity, this chapter also starts the analysis by using an upper-bound on capacity, similar to the approach in [13]. However, unlike [13] that only considers two special cases for high SNR and low SNR separately, this paper presents a direct, generalized solution for all cases by applying different factorizations and transformations, and shows that the optimum precoder based on this criterion for a general MIMO system is an eigen-beamformer and the power allocation policy on its eigen-values follows a water-pouring strategy, which is different from the results in [13] for a 2×2 channel.

Under MMSE criterion, we consider a general MIMO system and derive the optimum precoder that beamforms into the eigen-modes of the transmit and receive correlation matrices.

Under the minimum pairwise error probability (PEP) criterion, while a unity receive correlation matrix was assumed in [3] and [6], this chapter takes into account the effect of receive correlation matrix to derive objective function for computing the optimum precoding matrix applicable to both general uncoded and coded MIMO systems. It is further shown that, for orthogonal ST coding, the PEP-based optimum precoder also has the structure of an eigen-beamformer based on the eigen-modes of the transmit and receive correlation matrices.

In other words, the optimal linear precoder for any uncoded and coded MIMO system based on the MMSE or ergodic capacity criterion, or for an orthogonal ST coded MIMO system based on the minimum PEP criterion, is shown to be an eigen-beamformer that transmits the signal along eigenvectors of the transmit correlation matrix. Based on the eigen-values of both the transmit and receive correlation matrices, power loading across the eigen-beams is determined and can be viewed as a kind of water-pouring policy. Moreover, by simulations, we show that the receive antenna correlation matrix has a minor effect on the design of the optimal transmitter.

Section 2.3 considers the same scenario while optimizing a tight upper-bound on ergodic capacity of a MIMO system. The results are interesting because it gives us the same precoder structure in Section 2.2 except for a power-loading matrix. In fact, the precoder turns out to be an eigen-beamformer with different power loadings over eigen-modes compared to the case of PEP and MMSE.

Finally, in Sections 2.4 and 2.5 illustrative results and chapter summary are presented. The results in this chapter have been partly presented in [21], [22], [86], [87] and [91].

While Chapter 2 deals with designs for frequency-flat channel, Chapter 3 is allocated to the scenarios in which the underlying channel is frequency-selective. Frequency-flat fading channels have been considered for various precoding designs based on partial channel knowledge. There are, however, very few works addressing the frequency-selective fading environment. In [5], a precoder has been designed for an OFDM based MIMO system exploiting the transmit- and path-correlation properties. However, in the final design, the effect of path-correlation matrix was neglected. In practice, multi-path signals can be correlated in various cases. For example, they are generated by scattering clusters seen in a narrow angular range and at a long distance from the transmit (or receive) antenna, or they are faded signals before a pinhole and after a pinhole with delay spread [5], [20]. Due to the transmit and receive filtering structures, the fading coefficients from various paths with different delays can be statistically correlated and their correlation can be significant sometimes in some part of the channel (e.g., near the keyhole). Other types of precoder design for frequency-selective channels are also possible, which are not necessarily linear. For example, a non-linear precoding scheme for MIMO frequency-selective channel using the Tomlinson-Harashima precoder was proposed in [111].

Chapter 3 considers MIMO precoding designs for general frequency-selective fading channels based on both spatial and path correlation matrices and derives the structures for capacity-approaching linear precoders. Three separate cases are considered: (i) *uncorrelated* channel paths with *similar* spatial correlation, (ii) *uncorrelated* channel paths with *different* spatial correlation, and (iii) *correlated* channel paths. For *uncorrelated* channel paths, it is shown that the precoder is composed of a number of parallel precoders for frequency-flat fading channels. The power allocation to each precoder is based on the power of channel paths and the eigenvalues of transmit correlation matrix and can be calculated based on a statistical water-pouring policy. Furthermore, in the case of similar spatial correlation, these parallel precoders have the same structure. For *correlated* paths, we show that the selected precoder is an eigen-beamformer and the power allocation on each eigenmode follows a statistical water-pouring strategy based on the product of the eigenvalues of transmit and path correlation matrices. Capacity improvement of the proposed precoders based on partial channel knowledge is investigated in different propagation scenarios such as correlated and uncorrelated channel paths and transmits antennas.

The rest of Chapter 3 is organized as follows. Based on the representation of a frequency-selective fading channel by a model with L effective paths, Section 3.1 develops a comprehensive model suitable to analyze a general MIMO system in a frequency-selective fading environment with emphasis on the spatial and path correlations. We consider different propagation scenarios and develop the channel model that fits to the three cases mentioned above. Using this model, Section 3.2 formulates the optimization problem to develop precoder/decoder designs that maximize an upper bound on the ergodic capacity under the transmitted power constraints, based on the knowledge of spatial and path correlation matrices at the transmitter. The optimum precoding structures are derived for both uncorrelated and correlated channel paths. In Section 3.3, performance in terms of achievable ergodic capacity versus SNR of the proposed precoders is evaluated and compared with that of systems using either no precoding or spatial precoding in various scenarios by simulation. Finally, conclusion is presented in Section 3.4.

Parts of the analysis and results studied in this chapter have been already presented in [86], [88], [89], [90] and [91].

1.2. Precoder Designs for Point-to-Multipoint Transmission

In Chapter 4, we focus on the problem of transmission in point-to-multipoint MIMO systems based on partial channel information. More specifically, Chapter 4 proposes efficient user-selection and zero-forcing precoding schemes based on partial channel knowledge at the transmitter, suitable for broadcast MIMO wireless systems. The design criterion is based on the average user sum rate.

We stated that linear precoding is an effective tool to achieve better performance and higher throughput in multiple-input-multiple-output (MIMO) communication systems. It is shown [19], [22], [58] that MIMO precoding techniques for single-user systems can be applied to obtain the achievable capacity of a MIMO channel[1]. The MIMO channel capacity can be computed by optimizing the transmit covariance matrix that can be interpreted as a linear precoding matrix. In other words, derivation of the capacity of a point-to-point MIMO channel and optimum linear precoder design to achieve the capacity of a single-user system are in fact dual problems. The duality between precoder design and system capacity calculation, however, does not exist in MIMO broadcast (BC) systems, i.e., there is no *linear* scheme that achieves the capacity of MIMO BC channels. In [59], the achievable region of MIMO BC channels for two single-antenna users was evaluated using an approach known as dirty paper coding (DPC) [60]. Later, [61] and [62] suggested a duality between multi-access (MA) and MIMO BC channels, and used it to calculate the rate region and capacity of a general MIMO BC system with an arbitrary number of multi-antenna users, which can be approached by DPC [61]. DPC based on the nonlinear structure of generalized decision feedback equalizer (GDFE) can be implemented using a successive interference pre-cancellation technique at the transmitter [65]. However, due to the complex nature of optimization processes in use and required assumptions, application of DPC is limited in real communication systems with relatively large number of users and time-varying MIMO wireless channels.

While it is shown that channel knowledge at transmitter is very important to achieve high performance and capacity [62], similar to the point-to-point transmission, the assumption of *full* knowledge of actual channel responses in multi-user wireless systems is impractical since such knowledge is very costly to acquire due to a large number of user channels and quickly outdated due to the channel time-varying nature.

[1] limited by the minimum number of the transmit and receive antennas [53], [57].

Therefore, *partial* channel knowledge at transmitter based on channel *statistical* parameters with much slower variation than the actual channel responses, has been recently considered as a realistic assumption in multi-user communication systems. In [63] and [67], a random beamforming scheme based on partial channel state information (CSI) has been suggested, in which, as the number of users increases to infinity, its sum rate can achieve the same scaling factor as the capacity obtained with perfect CSI using DPC, i.e.,

$$\lim_{n\to\infty}\frac{R}{C_{sum}}=\lim_{n\to\infty}\frac{R}{R_{DPC}}=1 \tag{1.1}$$

where n is the number of the users, R is the achievable rate of the proposed scheme and C_{sum} is the sum capacity of the system assuming full channel knowledge at the transmitter which is the same as the rate achieved by DPC (R_{DPC}).

In Chapter 4, we investigate downlink precoding schemes that can achieve the capacity of a MIMO broadcast channel in which a multiple-antenna transmitter communicates with a number of mobile units, based on the assumption that the transmitter just has a partial knowledge of users' channels.

In [66], using full channel knowledge at the transmit side, a zero-forcing precoding scheme has been introduced and its performance has been evaluated. It has also been proved that when the number of users tends to infinity, the zero-forcing scheme is optimal in term of capacity, i.e., it can achieve the ergodic sum capacity of BC channels when the number of users is large. The problem with this scheme lies, however, in the need for full channel knowledge of the channel at the transmitter, which is difficult to obtain in fast-varying wireless channels. Another difficulty with the scheme proposed in [66] is the complexity of the transmit operations. The transmitter has to find a set of best channels in each transmission interval that have the most-orthogonal channel vectors. Note that, the optimality of the above approaches in [63] and [66] is asymptotical, i.e., the optimality decreases with reduced number of users. An interesting zero-forcing approach which is also similar to the one proposed in [66] but with the assumption of partial channel information has been proposed in [112] and [113]. The authors prove that this scheme can also be asymptotically optimum at the limit of large number of users.

Our study shows that a careful selection of channel side information is very important in the sense that it can reduce the feedback cost and transmit complexity while still provides a considerable performance. We propose a zero-forcing transmission scheme that uses only partial channel information with low feedback load and also facilitates the algorithm of selecting the best users at the transmitter. The proposed scheme achieves the same ergodic sum capacity as the scheme in [66] (and hence DPC) with much reduced feedback load and algorithm complexity.

By distributing the processing loads among users communicating in the network, the proposed scheme greatly relaxes the feedback load needed to calculate the optimal precoding vectors, and the complexity of selecting the best user algorithm at the transmitter. This reduction in feedback overhead and complexity as compared to other schemes is increased with the number of users and/or transmits antennas and hence, makes the proposed scheme a desired candidate for distributed network management systems.

Four different strategies for user selection, power allocation and precoding are proposed. Each of these strategies is suitable for a specific propagation scenario and channel condition. However, as the structure and algorithms of all of these schemes are identical with minor difference, it is possible to implement all of them at the base station and switch amongst them when necessary. This gives a degree of robustness to the system that can cope with channel impairments and changes.

Furthermore, in [66], users with high channel gains (e.g., closer to the base station) always have a better chance to be selected. Such unfairness is avoided in the proposed schemes since the chance of the selection of a user does not depend on the channel gain (power) and the distance to the base station, i.e., all the users have equal chance to be selected as the best (winner) users.

This chapter is organized as follows. In Section 4.1 we propose a simple model suitable for multi-user MIMO broadcast systems. In Section 4.2, the achievable sum rate and capacity of a MIMO broadcast system is investigated and the idea of zero-forcing precoding is proposed. It is shown that under certain circumstances zero-forcing precoding is asymptotically optimum in term of sum rate of the system achieved by applying dirty paper coding (DPC). In Section 4.3 and 4.4, the main zero-forcing precoding scheme that uses very limited channel information is proposed and it is shown that it can also achieve to the optimum sum rate at the limit of high number of active users. Some practical and design considerations of the proposed scheme such as feedback load and complexity are studied in Section 4.5. Section 4.6 considers some numerical results while 4.7 provide chapter summary and some concluding remarks.

Note that the scheme proposed in Chapter 4 has been partly presented in [92] and [93].

1.3. Precoder Designs for Cooperative Relay Transmission

Cooperation has been subject of intensive study in different branches of science. Nevertheless, the extensive use of this concept in communications is relatively new and has recently gained a considerable attention in the related literature [80].

Chapters 5 and 6 of this thesis concentrate on precoder designs for cooperative relay networks. Relay networking [2]-[6] is one of the frameworks in which the concept of cooperation becomes meaningful. In these systems, a source node tries to send its corresponding information to the destination node with the help of one or a number of relay node(s). Cooperative relaying targets additional diversity and coding gain and provides additional level of reliability particularly when the forward link (channel between source and destination) is not reliable enough, i.e. forward channel is poorly conditioned.

In Chapters 5 and 6, we develop collaborative precoding strategies for cooperative MIMO relay systems. In the cooperative systems, several terminals collaborate to assist each other in transmission of their signals. More specifically, in cooperative relay systems, a source node transmits a signal to a number of relays. These relays resend the processed version of the received signal to the intended destination node. The destination node then can combine and process all the received signals (may also include the signal directly transmitted from source node) to enhance the detection performance. Because of the nature of the wireless channels and the fact that the channels between transmit, receive and relay nodes are independent, the cooperative relaying system can achieve a diversity order that is a function of the number of independent channels. In a multi-user environment, cooperative relaying systems can also mitigate the interference by optimizing the transmission scheduling and processing done in the cooperative relays to minimize mutual interference and to facilitate information transmission by cooperation. Diversity and interference mitigation can result in a higher performance and throughput. Cooperative diversity concept, therefore, promises a power-efficient solution for future wireless communications systems to achieve broader coverage and to mitigate channel impairments without the need to use large power at the transmitter.

It has been shown that there exists a fundamental trade-off between spatial multiplexing and diversity gain in multiple-input-multiple-output (MIMO) communication systems. It has been proved that for a MIMO system composed of m transmit and n receive antennas, using a sufficiently large code length ($l \geq m+n-1$) for transmission, for a given rate r, the optimal diversity order $d^*(r)$ that can be considered as the optimum trade-off is [68]:

$$d^*(r) = (m-r)(n-r) \;\; ; \;\; 0 \leq r \leq \min\{m,n\} \tag{1.2}$$

In particular, the maximum diversity order of the system and the maximum multiplexing gain are $d^*_{\max} = mn$ and $r_{max} = \min\{m, n\}$, respectively. The maximum diversity order of mn is achieved at multiplexing gain of $r = 0$ and vice-versa maximum multiplexing gain is obtained when there is no diversity gain. Furthermore, increasing diversity order will decrease multiplexing gain and any increase in multiplexing gain comes at the cost of reduced diversity order of the system.

A relay system [69], [39], [70] can be also considered as a multiple antenna system in which a source node transmits its information to the destination node with the help of the relay node(s). This system is in fact a MIMO system with m equal to the sum of the antennas at the source and relay node(s) [71]. However, the main difference from a MIMO system is that the transmit antennas cannot *fully cooperate*. This fact shows that the optimal trade-off between diversity and multiplexing gain of a MIMO system is in fact an upper-bound on that of a relay system. It means that regardless of the network topology and transmission scheme, it is not possible to achieve to the diversity-multiplexing trade-off of the *same-scale* MIMO system.

In the literature, two different kinds of relay systems have been recognized. In the first strategy, the relay node(s) does not decode the received signal but reflects it to the receiver with a specific weight. This is called *amplify-and-forward* (AF) relay network (e.g. [35]-[37]). It is also sometimes called *transparent* relay. On contrary, it is also possible that the relay node(s) decode the received signal from source node and transmit a decoded version of the signal to the destination node. This strategy is usually called *decode-and-forward* (DF) relaying (also *regenerative* relay) (e.g., [73], [74]). The application of relay networks covers a wide range from multiple-access (MA) and broadcast (BC) systems to sensory and ad-hoc networks [69], [78], [79].

In Chapter 5, our focus is in the case of amplify-and-forward relay networks, in which all the nodes are assumed to transmit in a *half-duplex* mode, i.e., the units cannot transmit and receive at the same time. This is a common assumption in the literature related to the practical complication and difficulty in preventing transmit-to-receive interference in a transceiver due to the large difference in operating power in the same frequency. While this half-duplex constraint is restrictive to protocol development, we nevertheless consider it throughout this paper.

On the other hand, to avoid mutual interference between channels and to make the system design easier, some works assume transmission in orthogonal channels. This assumption, however, results in a suboptimal use of resources such as bandwidth. We, on the other hand, assume non-orthogonal channels in which all the corresponding nodes transmit in the same frequency slot. Therefore, our study is on the *non-orthogonal* amplify-and-forward (NAF) relay systems.

Consider a two-hop, single-relay AF network, in which the source node transmits in the first interval while both relay and destination nodes listen and in the second interval relay node sends a weighted version of the signal received from the source while the source node remains silent. It has been shown in [75] [76] (and also [77] for a more general case) that this two-hop, single-relay AF network can achieve the following diversity-multiplexing trade-off

$$d^*(r) = (1-r) + (1-2r)^+ \ ; \ 0 \le r \le 1/2 \quad (1.3)$$

where $^+$ represents the positive part, i.e., $[x]^+=\max(0,x)$. In other words, (1.3) shows that the maximum achievable diversity order of this two-hop, single-relay AF network is 2 while its maximum multiplexing gain is just ½ as compared to an equivalent 2x1 MISO system with the maximum diversity and multiplexing gain of 2 and 1, respectively. As mentioned, this is partly due to non-collaborative nature of transmit antennas in relay systems and partly due to the assumption of half-duplex terminals. In [76], it has also been proved that (1.3) is in fact the optimum trade-off that can be achieved in an amplify-and-forward half-duplex single-relay system.

While there is an extensive study on the relay systems in which each individual terminal is equipped with single antenna, there has not been a great attention to the cases that some or all terminals are equipped with multiple antennas. It still remains an open problem to exploit the ability of multiple antennas to increase the multiplexing gain and diversity order of the relay systems. In particular, as the main target in relay systems is to achieve higher diversity orders, how we can optimize the diversity order while still keeping the multiplexing gain at an acceptable level.

This question is the main focus of our study in Chapter 5. More specifically, while we assume the transmit and relay nodes to be single antenna (to avoid complexity of the terminals), the receive node is assumed to be equipped with multiple antenna. Generalization to the case of multi-antenna transmit and relay nodes is straight forward. While this set-up is more suitable for the uplink of wireless communication systems and multiple access (MA) schemes, it can also be applied to ad-hoc and sensory networks, in which the access point (AP) is equipped with multiple antennas.

In Chapter 5, after providing a simple system model in Section 5.1, we briefly study a relay system from a theoretical point of view in Section 5.2. We investigate the question of diversity-multiplexing trade-off in a relay system with multi-antenna source and destination nodes, and single-antenna relay nodes. In Section 5.3, we propose a relay-assisted transmission scheme and using diversity-multiplexing trade-off concept to derive the minimum number of relay nodes that can approach the optimum diversity-multiplexing trade-off. In the rest of Chapter 5, we study to realize this promising trade-off using a BLAST-like algorithm [114], [57], [42] but in a *distributed* sense. We show that it is possible to achieve the optimality by a simple reception technique given that a suitable relay selection is applied to the system.

Our study shows that at any specific source rate smaller than one, careful selection of the number of relays and the relay nodes themselves can result in the optimum trade-off. In addition at relay nodes a *distributed* BLAST-like transmission scheme for sending the information already broadcast from source node can preserve the optimality. While detection of this distributed BLAST-like scheme is complex, it can be simplified by suitable selection of relays. We show that for a suitable selection of relays, a simple *successive nulling and cancellation* method can achieve optimality. Finally, we propose the idea of *distributive precoding* to combat non-optimalities in relay selection. For the problem of relay selection, a partial knowledge-based scheme (as discussed in Chapter 4) can be applied. This selection significantly reduces the amount of the feedback needed for relay selection.

In other words, the transmission scheme studied in Chapter 5 consists of two main aspects. First, based on the number of receive antennas, the optimal *number* of relay nodes and the optimum *relay nodes* amongst all possible combinations in the network are selected. Second, it applies a BLAST-like approach in conjunction with a distributive precoding scheme to the transmission of source information. The *precoder* is a simple linear transformation matrix that is calculated at the destination node (e.g. base station or access point) and is fed back to the source node. Numerical results and chapter summary are presented in Sections 5.3 and 5.4, respectively. Some of the results in Chapter 5 have been presented in [94].

We continue our study on cooperative relay systems in Chapter 6 from a more practical point of view. Two natural questions in the context of relay networks are the problems of transmission and reception strategies. Mainly, how the source and relay node(s) should send the information to the destination node and how the destination should optimally combine the source information with replicate version(s) of information from relay node(s). Obviously, one can not answer to these two questions separately. In other words, transmission and reception strategies should be jointly optimized.

A major practical issue that should be addressed in the design of transmission and reception schemes in relay networks is complexity. Unlike point-to-point transmission schemes in which transmitter and receiver are responsible for the recovery of their own information, in a relay system, other parts of the network are also engaged in the communications. Hence, it is very desirable that the communication strategy imposes minimum level of computational load on the relay nodes. This point becomes one of the constraints that should be taken into account in the design of strategies for relay systems.

While there is a vigorous body of work on the relay systems in which each individual terminal is equipped with single antenna, the case of multi-antenna nodes has not been studied extensively. [83] shows that the relay systems with MIMO capability offer a promising capacity and this capacity scales linearly with the number of antennas at source/destination and logarithmically with the number of relay nodes or antennas. [84] also considers a similar case and proposes a cooperative beamforming approach that can achieve the network capacity in the limit of a large number of relay nodes. The effect of relay-assisted transmission on the capacity of rank-deficient MIMO systems was considered in [85]. While all the above results and most of those reviewed in Chapter 5 are attractive from a theoretical point of view, there are also the needs for practical transmission and reception schemes to practically realize multi-antenna relay networks in order to provide higher capacity and better performance than systems with single-antenna terminals.

The lack or scarcity of such a study on design of communication schemes for multi-antenna relay systems in the literature is the main motivation of our study in

Chapter 6 for development of practical transmission and reception schemes for both AF and DF protocols and in a *half-duplex* operation mode. More specifically, throughout Chapter 6, source, destination and relay nodes are assumed to be all equipped with multiple antennas. Our goal is to find optimal transmission and reception schemes for this set-up while avoiding a huge complexity, especially at relay node.

We show that a maximum ratio combining scheme at the receiver in conjunction with suitable linear precoding techniques at transmit and relay node can lead to this end. Our study shows that the proposed scheme is optimal in term of received SNR (and capacity) while maintaining an acceptable computational load at all nodes. In addition, the technique can be applied to both AF and DF protocols with small modifications. This feature can facilitate switching between two protocols when necessary. On the other hand, structures of source and relay nodes are identical, which enables a node to play the role of the transmitter or relay in different time instants without the need of additional software or hardware overhead.

We further show that a generalized maximum ratio combiner (GMRC) at the destination is optimum for both AF and DF protocols in term of SNR. Furthermore, for DF protocol, the precoders at the source and relay nodes should send the information in the direction of the eigenvector associated strongest eigenvalues of the channel matrices. While it is straight forward to derive the structure of precoders in the case of DF protocol, the case of AF can not be elaborated easily. We instead propose a relay selection scheme that can result in the best possible received SNR.

Chapter 6 is organized as follows. In Section 6.1, we present the system model of the multi-antenna relay network. In Section 6.2, maximum ratio combining schemes for different setups such as point-to-point MIMO, multipoint-to-point system, AF and DF relaying are studied. Section 6.3 is dedicated to precoder designs for source and relay nodes. Numerical results and chapter summary are presented in Sections 6.4 and 6.5, respectively. Materials of Chapter 6 have been partly presented in [95].

Finally, Chapter 7 concludes this thesis with an overall summary of the proposed schemes and results and a brief discussion of potential further work.

Our work in this thesis starts by the simple case of point-to-point transmission and moves to the more general case of point-to-multipoint transmission (broadcast system) and finally lands in the scenario of cooperative relay network. Figure 1.1 shows that one can interpret a cooperative relay network as a combination of two broadcast and

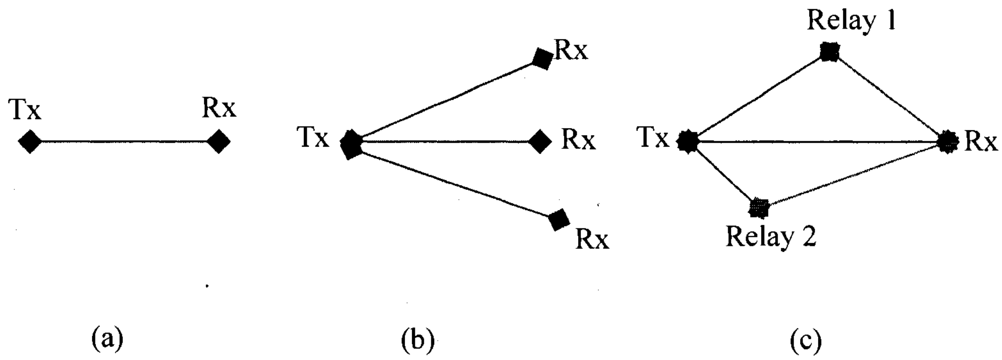


Figure 1.1: Different system scenarios considered in this thesis: (a) point-to-point, (b) point-to-multipoint and (c) cooperative relay transmissions

multi-access system. Therefore, cooperative relay networks can be viewed as a generalization to the case of broadcast system. This argument sheds more lights on the sequence of our research work presented in this thesis.

1.4. Summary of Contributions

- Precoder design for point-to-point MIMO systems based on partial channel information knowledge:

  - ✓ Closed-form derivation and comparison of the precoder structures for frequency-flat fading channels with transmit and receive correlation matrices known at the transmitter

  - ✓ Proposing the use of path correlation matrices as a partial channel information in frequency-selective fading channels

  - ✓ Derivation of precoder structure for frequency-selective fading channels in different propagation scenarios

- Precoder design for point-to-multipoint multi-user MIMO systems:

  - ✓ Development of a partial channel knowledge-based user selection scheme using pre-defined orthonormal set in MIMO broadcast systems

  - ✓ Development of a combined precoding and user-selection scheme that assymptotically achieve the optimum transmission in MIMO broadcast systems

- ✓ Comparison of different user-selection techniques in term of complexity, feedback load and design considerations

- Precoder design for cooperative relay systems:
  - ✓ Derivations of diversity-multiplexing trade-off in cooperative relay networks
  - ✓ Development of a distributed BLAST scheme with distributive precoding for transmission in relay systems
  - ✓ Derivations of the optimum combining schemes for AF and DF systems
  - ✓ Derivations of the optimum precoding structures for AF and DF systems

Chapter 2

Partial Channel Knowledge-Based Precoding in Frequency-Flat Channels[1]

2.1. MIMO Frequency-Flat Channel Model

In the design of communication systems, appropriate channel modeling is of great importance. As much as the channel model is close to reality and it considers different propagation phenomena, the transmission/detection schemes designed based on that specific channel model can work better in reality. Here, our main emphasis is on the models that consider the long-term statistics in addition to the short-term statistics of the channel. Long-term channel statistics are those parameters that do not change as frequently as instantaneous channel. For example, in a MIMO communication link, the correlation between different subchannels is a long-term statistics because it changes slowly (slower than channel instantaneous response (gain)). There are several channel parameters that are fixed or change slowly in time and hence can be considered as long-term information of the channel. Some of them are channel mean, variance, number of paths in multi-path channels, transmit and receive antenna correlation in MIMO channels.

In this section, we develop the channel models suitable for precoder design based on partial channel knowledge at the transmitter for a frequency-flat channel. Later in Chapter 3, we present a suitable model for frequency-selective channels as well. Partial channel knowledge at the transmitter is in fact those parameters that are related to the

[1] This chapter has been partially presented in [21], [22], [86] and [87].

long-term statistics (behavior) of the channel. Compare to the full channel knowledge case (transmitter knows the channel instantaneously) partial channel knowledge is much easier to obtain at the transmitter because it does not change very fast and can be tracked or fedback from receiver to the transmitter easier.

We mainly propose the models that consider the spatial correlation for frequency flat channel model and path correlation for frequency selective case as partial channel knowledge. Correlation (either spatial or path) has been proved to be a main source of performance degradation in MIMO communication systems [20], [21], [42]. Proper modeling of correlation is therefore, very important in the design of MIMO transmission schemes. On the other hand, as correlation is a long-term statistical parameter of the channel, it can be easily estimated and tracked (by receiver and/or transmitter). Hence, obtaining correlation information at the transmitter is easy.

Using the correlation information, it is possible to design the schemes that aim to improve the system performance and/or throughput to a high extent. Our stress in this thesis proposal is therefore on the precoder schemes that uses spatial and path correlation information at the transmitter to improve the system performance. Hence, in this chapter we propose the channel models that are suitable for this purpose.

Figure 2.1 shows a typical MIMO channel in which transmitter and receiver are equipped with a number of antennas. A traditional frequency flat MIMO channel model for a system composed of M transmit and N receive antenna is simply an $N \times M$ matrix, $\mathbf{H}$ whose components are random variables. Each entry represents a single-input-single-output (SISO) channel between each pair of the transmit and receive antenna arrays sometimes known as subchannels and they are usually assumed to be zero-mean complex Gaussian random variable. For example, h_{nm} is the channel from the m^{th} transmit to the n^{th} receive antenna. The system model can then be written as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.1}$$

where $\mathbf{x}$ is the $M \times 1$ input vector and $\mathbf{y}$ and $\mathbf{n}$ are the $N \times 1$ output and noise vectors, respectively.

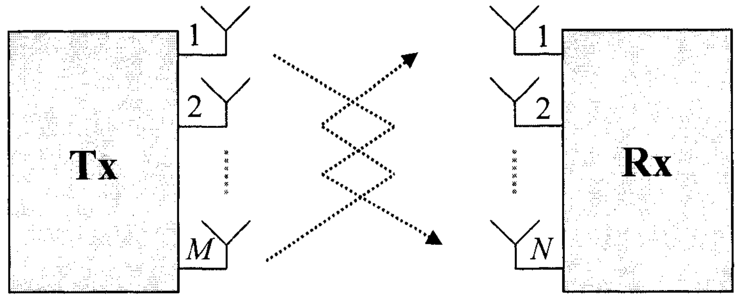


Figure 2.1: A Typical MIMO Channel

Traditional channel models for MIMO systems, however, do not consider spatial correlation separately. Spatial correlation is in fact the correlation between the elements of the MIMO channel matrix. They simply model channel as a matrix with independent random variables entries. Each entry specifies a SISO channel between a pair of transmit and receive antenna. It clearly does not take into account the effect of the interaction between subchannels. Nevertheless, because of the proximity of the antenna elements in channel, the signals on different subchannels interfere with each other and hence, it will result in correlation between subchannels in a MIMO system.

Here, to consider the effect of the spatial correlation and facilitate the design of the precoder based on spatial correlation, the following widely used channel model is used. In many applications, the transmit and/or the receive antennas can be correlated. For example, in cellular systems, the base-station antennas are typically unobstructed and have no local scatterers. This induces correlation across the base-station antennas, as a result of which the MIMO channel matrix entries do not fade independently. Antenna correlations tell us about the available spatial diversity in a MIMO channel. If the antennas are highly correlated, very little spatial diversity gain can be extracted from the channel and vice-versa if the antennas are uncorrelated [2].

We consider the transmitter and receiver equipped with M and N antennas, respectively. The MIMO channel is represented by the $N\times M$ matrix $\mathbf{H}$ and its entry, h_{nm} is the complex-valued gain from the m^{th} transmit to n^{th} receive antenna. Assuming that the transmit and receive scattering radii are large compared to the distance between the transmitter and the receiver, the channel matrix model $\mathbf{H}$ can be decomposed as [4], [20]

$$\mathbf{H} = \mathbf{R}_R^{1/2}\mathbf{G}\mathbf{R}_T^{1/2} \tag{2.2}$$

where $\mathbf{G}$ is an $N\times M$ matrix with i.i.d. zero-mean complex Gaussian entries and ½ variance per dimension. Furthermore, $\mathbf{R}_R = \mathbf{R}_R^{1/2}(\mathbf{R}_R^{1/2})^H = E\{\mathbf{HH}^H\}$ and $\mathbf{R}_T = \mathbf{R}_T^{1/2}(\mathbf{R}_T^{1/2})^H = E\{\mathbf{H}^H\mathbf{H}\}$ are the receive and transmit correlation matrices, respectively. Entries of $\mathbf{R}_R$ and $\mathbf{R}_T$ are determined from the receive and transmit antenna spacing and angular spread [4]. The superscript ^{H} denotes the Hermitian transposition (i.e., the operation of transposition combined with complex conjugation).

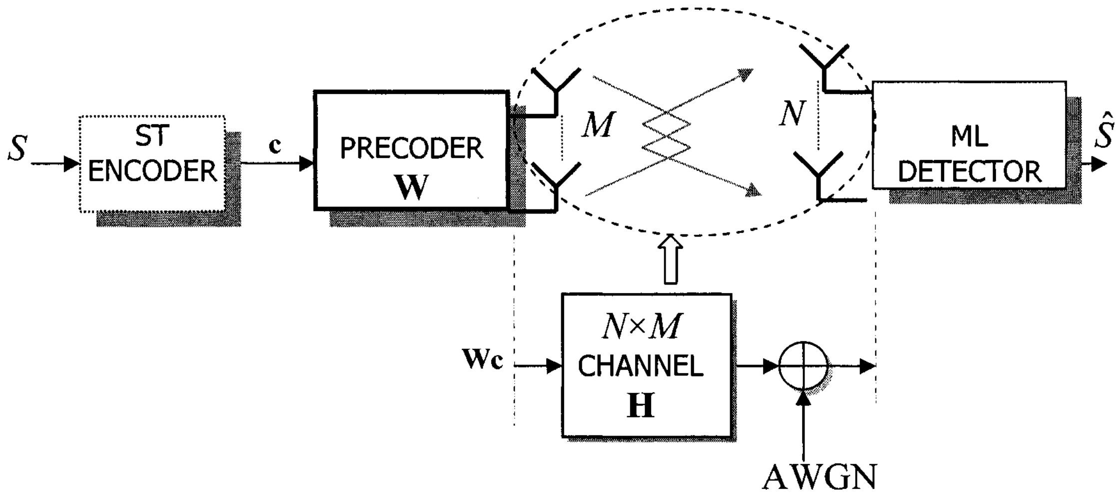


Figure 2.2: MIMO System Block Diagram

The channel model structure of (2.2) facilitates the following analysis in deriving the optimum precoding structures based on the transmit and receive correlation matrices, and the performance evaluation to examine their individual effects. This Kronecker model [43], [44] has been widely used in the literature for correlated MIMO systems [3], [6], [13], [44] although it may have limitations in representing some propagation phenomena such as keyhole phenomenon related to the media correlations corresponding to scatterers that are not local [4], [20] and propagation characteristic changes in different transmit antennas [11], [12], [21].

Figure 2.2 shows the block diagram of the MIMO system under consideration. The transmitter includes a ST encoder followed by a linear precoder governed by an $M\times M$ matrix $\mathbf{W}$. Input data, S, is first mapped to a codeword $\mathbf{c}$ by the ST encoder, and subsequently undergoes the linear precoding to form the codeword $\mathbf{Wc}$ to be transmitted over the MIMO channel **H**. At the receive side, the noisy codeword is recovered by maximum likelihood (ML) decoding.

The transmitter is assumed to know $\mathbf{R}_T$ and $\mathbf{R}_R$ matrices. Based on this information, the transmitter design chooses the optimum **W** according to the design criteria.

2.2. Optimal Precoding Using PEP and MMSE Criteria

2.2.1. Derivations of Optimal Precoding using PEP Criterion

In this subsection, we will first derive the general objective function based on PEP criterion, which can be used to compute the optimum precoding matrix for a general coded or uncoded MIMO system. We will then demonstrate that, for orthogonal space-time codes, the optimum precoding matrix structure turns out to be an eigen-beamformer with orthogonal beams pointing to the eigenvectors of the transmit correlation matrix.

Based on the model described in the previous section, the pair-wise error probability (PEP), i.e., the probability that a transmitted codeword **c** is erroneously received as a different codeword **e**, can be upper-bounded as

$$PEP(\mathbf{c},\mathbf{e}) \leq \exp\left(-\frac{E_s}{4N_0}\mathrm{tr}(\mathbf{HW}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^{\mathrm{H}}\mathbf{W}^H\mathbf{H}^H)\right) \tag{2.3}$$

where $N_0/2$ is the noise variance per dimension, E_s is the average constellation energy, **c** and **e** are $M{\times}K$ codeword matrices (K is a constant that depends on the code structure). Furthermore, the power constraint $\mathrm{tr}\left(\mathbf{WW}^H\right)=M$ should be imposed on **W** to limit the total transmitted power.

By using (2.2), the equality $\mathrm{tr}(\mathbf{AB})=\mathrm{tr}(\mathbf{BA})$ and the eigen-decomposition of the nonnegative definite matrices $\mathbf{R}_T^{1/2}\mathbf{W}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H\mathbf{W}^H\left(\mathbf{R}_T^{1/2}\right)^H=\mathbf{U\Lambda U}^H$, $\mathbf{R}_R^{1/2}\left(\mathbf{R}_R^{1/2}\right)^H=\mathbf{VDV}^H$, (2.3) can be rewritten as

$$PEP \leq \exp\left(-\frac{E_s}{4N_0}\mathrm{tr}\left(\mathbf{VDV}^H\mathbf{GU\Lambda U}^H\mathbf{G}^H\right)\right) \tag{2.4}$$

Furthermore $\mathrm{tr}\left(\mathbf{VDV}^H\mathbf{GU\Lambda U}^H\mathbf{G}^H\right)=\mathrm{tr}\left(\mathbf{DV}^H\mathbf{GU\Lambda U}^H\mathbf{G}^H\mathbf{V}\right)=\mathrm{tr}\left(\mathbf{D\tilde{G}\Lambda\tilde{G}}^H\right)$

where $\tilde{\mathbf{G}}=\mathbf{V}^H\mathbf{GU}$.

Since $\mathbf{V}$ and $\mathbf{U}$ are unitary matrices, the entries of $\tilde{\mathbf{G}}$ and $\mathbf{G}$ have the same distribution. The diagonal entries, d_i, $i=1,..,N$ and λ_j, $j = 1,..,M$ of the diagonal matrices $\mathbf{D}$ and $\mathbf{\Lambda}$, respectively, are nonnegative since they are the eigen-values of the nonnegative definite matrices $\left\{\mathbf{R}_T^{1/2}\mathbf{W}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H\mathbf{W}^H\left(\mathbf{R}_T^{1/2}\right)^H\right\}$ and $\left\{\mathbf{R}_R^{1/2}\left(\mathbf{R}_R^{1/2}\right)^H\right\}$, respectively. Therefore,

$$\mathrm{tr}\left(\mathbf{D}\tilde{\mathbf{G}}\mathbf{\Lambda}\tilde{\mathbf{G}}^H\right)=\mathrm{tr}\left(\left(\mathbf{D}^{1/2}\tilde{\mathbf{G}}\mathbf{\Lambda}^{1/2}\right)\left(\mathbf{D}^{1/2}\tilde{\mathbf{G}}\mathbf{\Lambda}^{1/2}\right)^H\right)=\sum_{i=1}^{N}\sum_{j=1}^{M}d_i\lambda_j\left|\tilde{g}_{ij}\right|^2 \tag{2.5}$$

where $\tilde{g}_{ij}$ is the (i,j) entry of $\tilde{\mathbf{G}}$. Using (2.5), (2.4) can be expressed as

$$PEP\leq\prod_{i=1}^{N}\prod_{j=1}^{M}\exp\left(-\frac{E_s}{4N_0}d_i\lambda_j\left|\tilde{g}_{ij}\right|^2\right) \tag{2.6}$$

Note that $\left|\tilde{g}_{ij}\right|$'s are independent Rayleigh-distributed components. By averaging (2.6) with respect to $\left|\tilde{g}_{ij}\right|$, we obtain the upper bound on average PEP

$$\overline{PEP}\leq\prod_{i=1}^{N}\prod_{j=1}^{M}\frac{1}{\left(1+\left(\frac{E_s}{4N_0}\right)d_i\lambda_j\right)}=\left(\det\left(\mathbf{I}+\left(\frac{E_s}{4N_0}\right)(\mathbf{D}\otimes\mathbf{\Lambda})\right)\right)^{-1} \tag{2.7}$$

where $\otimes$ stands for the Kronecker product.

Our goal is to find the optimum $\mathbf{W}$ that minimizes the right-hand side of (2.7). Toward this objective, we perform the following manipulations

$$\begin{aligned}\det\left(\mathbf{I}+\left(\frac{E_s}{4N_0}\right)(\mathbf{D}\otimes\mathbf{\Lambda})\right)&=\det\left(\mathbf{I}+\left(\frac{E_s}{4N_0}\right)\left(\mathbf{D}\otimes\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\right)\right)\\&=\det\left(\mathbf{I}+\left(\frac{E_s}{4N_0}\right)\left(\mathbf{D}\otimes\left(\mathbf{R}_T^{1/2}\mathbf{W}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H\mathbf{W}^H\left(\mathbf{R}_T^{1/2}\right)^H\right)\right)\right)\\&=\det\left(\mathbf{I}+\left(\frac{E_s}{4N_0}\right)(\mathbf{D}\otimes\mathbf{\Lambda})\left(\mathbf{I}\otimes\left(\mathbf{\Phi}^H\mathbf{\Psi}\right)\right)(\mathbf{I}\otimes\mathbf{P})\left(\mathbf{I}\otimes\left(\mathbf{\Phi}^H\mathbf{\Psi}\right)^H\right)\right)\end{aligned} \tag{2.8}$$

In deriving the above expression, we used the eigen-decomposition of the nonnegative-definite Hermitian matrices: $\mathbf{W}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H\mathbf{W}^H=\mathbf{\Psi}\mathbf{P}\mathbf{\Psi}^H$ and $\mathbf{R}_T=\mathbf{\Phi}\mathbf{\Delta}\mathbf{\Phi}^H$.

So far, the derivations are for a general case of $\mathbf{c}$ and hence applicable to both general uncoded and coded MIMO systems to compute the optimum precoding matrix by minimizing the right-hand side of (2.8).

Further simplification is possible for the orthogonal case, i.e., $(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H = \alpha\,\mathbf{I}$ where α is a scalar. It follows that $\mathbf{\Psi P \Psi}^H = \mathbf{W}(\mathbf{c}-\mathbf{e})(\mathbf{c}-\mathbf{e})^H \mathbf{W}^H$ and the power constraint translates to $\mathrm{tr}(\mathbf{P}) = \alpha M$.

Now, turning our attention back to (2.8), our optimization problem takes the following form

$$\max_{\mathbf{P}} \det\left(\mathbf{I} + \left(\frac{E_s}{4N_0}\right)(\mathbf{D}\otimes\mathbf{\Lambda})\tilde{\mathbf{U}}(\mathbf{I}\otimes\mathbf{P})\tilde{\mathbf{U}}^H\right) \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = \alpha M \tag{2.9}$$

Since $\tilde{\mathbf{U}} = \mathbf{I}\otimes(\mathbf{\Phi}^H\mathbf{\Psi})$ is unitary, the Hadamard's inequality [23] suggests that $\tilde{\mathbf{U}} = \mathbf{I}$ or $\mathbf{\Psi} = \mathbf{\Phi}$. Therefore, the optimal precoder $\mathbf{W}_{opt}$ is

$$\mathbf{W}_{opt} = (\alpha^{-1/2})\mathbf{\Phi}\mathbf{P}^{1/2}\mathbf{\Gamma} \tag{2.10}$$

where $\mathbf{\Gamma}$ is an arbitrary unitary matrix which has no effect on the system performance and therefore can be set to identity. In other words, the optimal linear precoding matrix $\mathbf{W}_{opt}$ turns out to be an eigen-beamformer with orthogonal beams pointing to the eigenvectors of the transmit correlation matrix, $\mathbf{R}_T$.

Furthermore, the power loading policy across the eigen-beams can be obtained by substituting $\tilde{\mathbf{U}} = \mathbf{I}$ in (8) as

$$\max_{P} \det\left(\mathbf{I} + \left(\frac{E_s}{4N_0}\right)(\mathbf{D}\otimes\mathbf{\Lambda})(\mathbf{I}\otimes\mathbf{P})\right) \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = \alpha M \tag{2.11}$$

2.2.2. Derivations of Optimal Precoding using MMSE Criterion

Corresponding to the transmitted vector $\mathbf{s}_i$, $\hat{\mathbf{s}}_i$ is detected at the receiver. The mean squared error (MSE) is defined as $MSE = E\{(\hat{\mathbf{s}}_i - \mathbf{s}_i)(\hat{\mathbf{s}}_i - \mathbf{s}_i)^H\}$ and can be expressed in terms of $\mathbf{H}$, $\mathbf{W}$, $\mathbf{U}$, $\mathbf{V}$, $\mathbf{G}, \tilde{\mathbf{G}}$, $\mathbf{D}$ and $\mathbf{\Lambda}$ (as previously defined and developed in Section 2.2.1) as

$$MSE = \mathrm{tr}\left(E_s\left(\mathbf{I}+\frac{E_s}{N_0}\mathbf{W}^H\mathbf{H}^H\mathbf{H}\mathbf{W}\right)^{-1}\right) = \mathrm{tr}\left(E_s\left(\mathbf{I}+\frac{E_s}{N_0}\mathbf{V}\mathbf{D}\mathbf{V}^H\mathbf{G}\mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\right)^{-1}\right)$$
$$= \mathrm{tr}\left(E_s\left(\mathbf{I}+\frac{E_s}{N_0}\mathbf{D}\tilde{\mathbf{G}}\mathbf{\Lambda}\tilde{\mathbf{G}}^H\right)^{-1}\right) \quad (2.12)$$

At low SNRs, Using $(\mathbf{I}+\mathbf{A})^{-1} = \mathbf{I}-\mathbf{A}+\mathbf{A}^2 - + \ldots \approx \mathbf{I}-\mathbf{A}$ with $\mathbf{A} = \frac{E_s}{N_0}\mathbf{D}\tilde{\mathbf{G}}\mathbf{\Lambda}\tilde{\mathbf{G}}^H$ in (2.12), we obtain $MSE \approx E_s \mathrm{tr}\left(\mathbf{I}-\frac{E_s}{N_0}\mathbf{D}\tilde{\mathbf{G}}\mathbf{\Lambda}\tilde{\mathbf{G}}^H\right)$. In this case, for independent Rayleigh-distributed components $|\tilde{g}_{ij}|$'s, we obtain the average MSE by averaging the right-hand side of the above equation, i.e.,

$$\overline{MSE} \approx E_s \sum_{i=1}^{M}\sum_{j=1}^{N}\left[1-\frac{E_s}{2N_0}d_i\lambda_j\right] = E_s \ \mathrm{tr}\left[\left(\mathbf{I}-\frac{E_s}{2N_0}(\mathbf{D}\otimes\mathbf{\Lambda})\right)\right] \quad (2.13)$$

Our goal is to obtain the minimum MSE (MMSE) in (2.13) or equivalently, to minimize $\mathrm{tr}\left(\mathbf{I}-\frac{E_s}{2N_0}(\mathbf{D}\otimes\mathbf{\Lambda})\right)$ or maximize $\mathrm{tr}\left(\frac{E_s}{2N_0}(\mathbf{D}\otimes\mathbf{\Lambda})\right)$. Following the same steps as for PEP, the precoder structure will be the same as (2.10) and the power across eigen-modes can be computed from the following optimization problem:

$$\max_{\mathbf{P}} \ \mathrm{tr}\left(\frac{E_s}{2N_0}(\mathbf{D}\otimes\mathbf{\Lambda})(\mathbf{I}\otimes\mathbf{P})\right) \qquad \textbf{s.t.} \ \ \mathrm{tr}(\mathbf{P}) = \alpha M \quad (2.14)$$

At high SNRs, the second term in (2.12) becomes dominant and $MSE \approx E_s \mathrm{tr}\left(\frac{E_s}{N_0}\mathbf{D}\tilde{\mathbf{G}}\mathbf{\Lambda}\tilde{\mathbf{G}}^H\right)^{-1}$. After taking its expectation, we obtain the average MSE, i.e., $\overline{MSE} \approx E_s \mathrm{tr}\left(\kappa\left(\mathbf{D}^{-1}\otimes\mathbf{\Lambda}^{-1}\right)\right)$ where κ is a constant that can be calculated by taking the expectation of the components of $\tilde{\mathbf{G}}^{-1}$. Following the same steps as PEP, we achieve to the same optimal precoder as in (2.10) and the optimization problem to the power loading policy across the eigen-beams can be obtained from:

$$\min_{\mathbf{P}} \ \mathrm{tr}\left(\kappa(\mathbf{D}\otimes\mathbf{\Lambda})(\mathbf{I}\otimes\mathbf{P})\right)^{-1} \quad \textbf{s.t.} \ \ \mathrm{tr}(\mathbf{P}) = \alpha M \quad (2.15)$$

In other words, in both high and low SNRs (2.10) is also applicable to the problem using MMSE criterion, i.e., the MMSE and PEP criteria lead to the same optimum precoder design except for the power loadings. However, the derivations for MMSE criterion do not assume orthogonality and hence the optimum linear precoder is applicable to both general coded and uncoded MIMO systems.

2.2.3. Power Loading Policy under PEP and MMSE criteria

In the following we consider the special cases that lead to a simpler form for the power loading policy for PEP case.

No Receive Correlation: For $\mathbf{R}_R = \mathbf{I}$, we have $\mathbf{D} = \mathbf{I}$ and (2.11) becomes

$$\max_{\mathbf{P}} \left(\det\left(\mathbf{I} + \left(\frac{E_s}{4N_0} \right) \mathbf{\Delta P} \right) \right)^N \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = \alpha M \tag{2.16a}$$

It can be shown that (2.16a) has the following solution for $\mathbf{P}$

$$p_i = \left(v - \left(\left(\frac{E_s}{4N_0} \right) \delta_i \right)^{-1} \right)^+ \tag{2.16b}$$

where δ_i' s are the diagonal entries of $\mathbf{\Delta}$ and v is a constant determined by the power constraint. This clearly has the form of the well-known water-pouring policy and coincides with the solution in [3]. From an intuitive point of view, when $\mathbf{R}_T$ is full rank (i.e., all δ_i' s are positive), as SNR increases, the second term on right hand side of (2.16b) decreases and allows the power to be distributed evenly among all eigen-values. Nevertheless, when some of the δ_i's are zero, the total power is divided among all other eigen-values of $\mathbf{R}_T$, according to the water-pouring policy.

No Transmit Correlation: For $\mathbf{R}_T = \mathbf{I}$, $\mathbf{\Delta} = \mathbf{I}$ and (2.11) becomes

$$\max_{\mathbf{P}} \det\left(\mathbf{I} + \frac{E_s}{4N_0} (\mathbf{D} \otimes \mathbf{P}) \right) = \prod_{i=1}^{N} \prod_{j=1}^{M} \left(1 + \left(\frac{E_s}{4N_0} \right) d_j p_i \right) \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = \alpha M \tag{2.17}$$

Due to perfect symmetry with respect to p_i, for all variations of $\mathbf{D}$, the expression adopts its maximum when $\mathbf{P} = \alpha\, \mathbf{I}$, i.e., the final solution is independent of $\mathbf{D}$.

General Case: (2.11) can be equivalently expressed in log-domain as

$$\max_{\mathbf{P}} \sum_{i=1}^{N} \sum_{j=1}^{M} \log\left(1 + \left(\frac{E_s}{4N_0} \right) d_j (\delta_i p_i) \right) \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = \alpha M$$

Apply Karush-Kuhn-Tucker (KKT) optimization conditions [56] on each of the elements, p_i's, of **P**, i.e., $\forall i: \sum_{j=1}^{N} \frac{\left(\frac{E_s}{4N_0}\right) d_j \delta_i}{1+\left(\frac{E_s}{4N_0}\right) d_j (\delta_i p_i)} = \mu$ where μ is a constant for KKT equations.

When receive correlation matrix is well-conditioned, the KKT conditions can be approximated as $\forall i: \frac{1}{1+\left(\frac{E_s}{4N_0}\right)\frac{1}{N}\mathrm{tr}(\mathbf{D})\delta_i p_i} \approx \mu$. This leads to the following approximation of the solution

$$p_i = \left(v - \left(\left(\frac{E_s}{4N_0}\right)\frac{1}{N}\mathrm{tr}(\mathbf{D})\delta_i\right)^{-1}\right)^{+} \tag{2.18}$$

where δ_i and v are the same parameters defined in the special cases. This solution is in complete agreement with the solutions of two above special cases.

For MMSE, based on the same discussions for PEP, in low SNR case, the optimum water-pouring policy which is the solution to (2.14) is to poor water on the strongest eigen-mode of transmit correlation matrix. In high SNRs, using KKT conditions, one can achieve to the following solution for the optimization problem in (2.15):

$$p_i = \left(v\delta_i^{-1/2} - \left(\left(\frac{E_s}{2N_0}\right)\frac{1}{N}\mathrm{tr}(\mathbf{D})\delta_i\right)^{-1}\right)^{+} \tag{2.19}$$

2.3. Optimal Precoder Using Ergodic Capacity Criterion

2.3.1. Derivations of Optimal Precoder using Maximum Ergodic Capacity Criterion

Ergodic capacity is the mean capacity over all channel realizations for a specific average SNR:

$$C = E\{C(\overline{\gamma})\}$$

where $E\{.\}$ denotes expectation with respect to the channel matrix distribution and $\bar{\gamma}$ is the average SNR. Ergodic capacity is used to see how the average capacity changes as a function of average SNR. The ergodic capacity of the MIMO system in Figure 2.2 is defined as

$$C = E\left[\log_2\left(\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{W}^H\mathbf{H}^H\mathbf{H}\mathbf{W}\right)\right)\right] \tag{2.20}$$

Again, the power constraint $\mathrm{tr}\left(\mathbf{W}\mathbf{W}^H\right) = M$ should be imposed on $\mathbf{W}$ to limit the total transmit power.

Getting expectation from the log function in (2.20) is very hard (if not impossible). By applying the Jensen's inequality [16] to the concave $\log_2[\det(.)]$ function, i.e., $E\{\log_2[\det(\mathbf{A})]\} \le \log_2[\det(E\{\mathbf{A}\})]$, one can obtain the following upper-bound on capacity

$$C \le C_{UB} = \log_2\left(\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{W}^H E\{\mathbf{H}^H\mathbf{H}\}\mathbf{W}\right)\right)$$

The above upper bound is simple but not very tight. Although it can provide a fast computation of the precoding matrix, it may result in a suboptimum precoder design. In the following, we will derive a tighter upper-bound.

Since $\log_2(.)$ function is also concave, we can use the Jensen's inequality $E\{\log_2[\det(\mathbf{A})]\} \le \log_2[E\{\det(\mathbf{A})\}]$, to obtain a tighter upper bound on the ergodic capacity as follows:

$$\begin{aligned} C \le C_{UB} &= \log_2\left(E\left\{\det\left(I + \frac{1}{N_0}\mathbf{W}^H\mathbf{H}^H\mathbf{H}\mathbf{W}\right)\right\}\right) \\ &= \log_2\left[E\left(\det\left(I + \frac{1}{N_0}\mathbf{W}^H\left(\mathbf{R}_T^{1/2}\right)^H\mathbf{G}^H\left(\mathbf{R}_R^{1/2}\right)^H\mathbf{R}_T^{1/2}\mathbf{G}\mathbf{R}_T^{1/2}\mathbf{W}\right)\right)\right] \end{aligned} \tag{2.21}$$

We will show that it is possible to get the expectation from the above equation with a reasonable complexity. Using the same relations developed in Section 2.2 for the matrices $\mathbf{T}$, $\mathbf{W}$, $\mathbf{U}$, $\mathbf{V}$, $\mathbf{G}, \tilde{\mathbf{G}}$, $\mathbf{D}$, and $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$, we have

$$C_{UB} = \log_2\left[E\left\{\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{\Lambda U}^H\mathbf{G}^H\mathbf{V}^H\mathbf{DVGU}\right)\right\}\right]$$
$$= \log_2\left[E\left\{\det\left(\mathbf{I} + \frac{1}{N_0}\mathbf{\Lambda}\tilde{\mathbf{G}}^H\mathbf{D}\tilde{\mathbf{G}}\right)\right\}\right] \tag{2.22}$$

The operation $E\{\det(.)\}$ can be computed by using the following determinant expansions [17]

$$\det(\mathbf{I}+\mathbf{A}) = \sum_{k=0}^{n}\sum_{\hat{\alpha}_k}\det(\mathbf{A})_{\hat{\alpha}_k}^{\hat{\alpha}_k}$$

and

$$\mathbf{A}_{(k\times k)} = \prod_{k=1}^{n}\mathbf{A}_i \rightarrow \det(\mathbf{A}) = \sum_{\hat{\alpha}_k^1}\sum_{\hat{\alpha}_k^2}\cdots\sum_{\hat{\alpha}_k^{n-1}}\det(\mathbf{A}_1)_{\hat{\alpha}_k^1}^{\{1,2,\dots,k\}}\det(\mathbf{A}_2)_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^1}\dots\det(\mathbf{A}_n)_{\{1,2,\dots,k\}}^{\hat{\alpha}_k^{n-1}}$$

where $\det(\mathbf{A})_{\hat{\alpha}_k^j}^{\hat{\alpha}_k^i}$ denotes the determinant of a sub-matrix of **A** obtained by selecting the row and column subset from the matrix **A** indexed by $\hat{\alpha}_k^i = \{\alpha_1^i, \alpha_2^i, \dots, \alpha_k^i\}$ and $\hat{\alpha}_k^j = \{\alpha_1^j, \alpha_2^j, \dots, \alpha_k^j\}$, respectively. The cardinalities of the subsets $\hat{\alpha}_k^i$ and $\hat{\alpha}_k^j$ are k. Furthermore, for a diagonal matrix, $\det(\mathbf{A})_{\hat{\alpha}_i}^{\hat{\alpha}_j} = 0 \; ; \; \hat{\alpha}_i \neq \hat{\alpha}_j$.

Since the matrices $\mathbf{\Lambda}$ and **D** are diagonal, by using the above matrix expansions, (2.22) can be expressed as

$$C_{UB} = \log_2\left[E\left(\sum_{k=0}^{K}\sum_{\hat{\alpha}_k^1}\sum_{\hat{\alpha}_k^2}\left(\frac{1}{N_0}\right)^k \det(\mathbf{\Lambda})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1}\det(\mathbf{D})_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^2}\det(\widetilde{\mathbf{G}})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^2}\det(\widetilde{\mathbf{G}}^H)_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^1}\right)\right] \tag{2.23}$$

where $K = \min\{M, N\}$.

Consider a matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ where $\mathbf{x}_j$ is an $m\times 1$ vector with zero mean i.i.d. complex Gaussian entries with ½ variance per dimension. It follows that the product $\mathbf{Y}=\mathbf{XX}^H$ has a Wishart distribution and $E(\det(\mathbf{Y})) = \frac{n!}{(n-m)!}$ [46]. The entries of submatrices $(\widetilde{\mathbf{G}})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^2}$ and $(\widetilde{\mathbf{G}}^H)_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^1}$ in (2.23) are zero mean i.i.d. complex Gaussian random variables with ½ variance per dimension. Therefore, $E\{\det(\tilde{\mathbf{G}})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^2} \det(\tilde{\mathbf{G}}^H)_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^1}\} =$ $E\{\det((\tilde{\mathbf{G}})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^2}(\tilde{\mathbf{G}}^H)_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^1})\} = \frac{k!}{(k-k)!} = k!$ and

$$C_{UB} = \log_2\left[\sum_{k=0}^{K}\sum_{\hat{\alpha}_k^1}\sum_{\hat{\alpha}_k^2} k!\left(\frac{1}{N_0}\right)^k \det(\mathbf{\Lambda})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1} \det(\mathbf{D})_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^2}\right]$$

$$= \log_2\left[\sum_{k=0}^{K}\sum_{\hat{\alpha}_k^1}\sum_{\hat{\alpha}_k^2} k!\left(\frac{1}{N_0}\right)^k \det\left(\mathbf{R}_T^{1/2}\mathbf{W}\mathbf{W}^H\left(\mathbf{R}_T^{1/2}\right)^H\right)_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1} \det(\mathbf{D})_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^2}\right] \quad (2.24)$$

$$= \log_2\left[\sum_{k=0}^{K}\sum_{\hat{\alpha}_k^1}\sum_{\hat{\alpha}_k^2} \frac{k!}{[N_0]^k} \det\left(\mathbf{\Lambda}\mathbf{\Phi}^H\mathbf{\Psi}\mathbf{P}\mathbf{\Psi}^H\mathbf{\Phi}\right)_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1} \det(\mathbf{D})_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^2}\right]$$

The above derivation uses the relations between $\mathbf{W}, \mathbf{\Psi}, \mathbf{P}, \mathbf{T}, \mathbf{\Phi}$ and $\mathbf{\Delta}$ as discussed in Section III.

Since $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are unitary matrices, the maximum value of C_{UB} in (2.24) is achieved when the product $\mathbf{\Delta\Phi}^H\mathbf{\Psi P\Psi}^H\mathbf{\Phi}$ is diagonal. For this purpose, we can set $\mathbf{\Psi} = \mathbf{\Phi}^H$ and consider the following structure of precoder matrix $\mathbf{W} = \mathbf{\Phi P}^{1/2}\mathbf{\Gamma}$.

Similar to the previous case, $\mathbf{\Gamma}$ is an arbitrary unitary matrix that has no effect on the optimization problem and therefore can be set to **I**. The linear precoding matrix **W** turns out to be an eigen-beamformer with orthogonal beams pointing to the eigenvectors of the transmit correlation matrix, $\mathbf{R}_T$.

In this case, the power loading policy across the eigen-beams can be obtained by setting $\mathbf{\Psi\Phi}^H = \mathbf{I}$ in (2.24), i.e.,

$$\max_{\mathbf{P}} \log_2\left[\sum_{k=0}^{K} k!\left(\frac{1}{N_0}\right)^k \sum_{\hat{\alpha}_k^1}(\det(\mathbf{\Delta})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1}\det(\mathbf{P})_{\hat{\alpha}_k^1}^{\hat{\alpha}_k^1})\sum_{\hat{\alpha}_k^2}\det(\mathbf{D})_{\hat{\alpha}_k^2}^{\hat{\alpha}_k^2}\right] \quad \textbf{s.t.} \quad \mathrm{tr}(\mathbf{P}) = M \quad (2.25)$$

2.3.2. Power Loading Policy under Maximum Ergodic Capacity Criterion

In the following, we will consider some special cases as well as the general case.

Capacity Bound: For a MIMO system without any precoding, if there is no correlation in the channel (i.e., the correlation matrices are equal to identity), we can find an upper-bound for capacity as

$$C_{UB} = \log_2\left[\sum_{k=0}^{K} \frac{k!}{[N_0]^k}\binom{M}{k}\binom{N}{k}\right] \tag{2.26}$$

The above result has also been derived in [13].

No Receive Correlation: For simplicity, consider the case of $M = N = 2$. For $\mathbf{R}_R = \mathbf{I}$, we have **D**=**I**. Furthermore, for the sake of brevity, the factor $[N_0]^{-k}$ in (2.24) can be omitted by simply multiplying each diagonal elements of matrix **Δ** by $\frac{1}{N_0}$. The upper bound in (2.24) is reduced to $C_{UB} = \log_2[1 + \text{tr}(\mathbf{\Delta P}) + 2\det(\mathbf{\Delta P})]$. For 2×2 diagonal matrices **Δ** and **P**, it can be verified that $\det(\mathbf{I}+\mathbf{\Delta P})=1+\text{tr}(\mathbf{\Delta P})+\det(\mathbf{\Delta P})$. Therefore, $C_{UB} = \log_2[\det(\mathbf{I} + \mathbf{\Delta P}) + \det(\mathbf{\Delta P})]$. By using the Minkowski's inequality, $\det(A + B) \geq \det(A) + \det(B)$ [23], we get

$$C_{UB} \leq \log_2[\det(\mathbf{I} + 2\mathbf{\Delta}\,\mathbf{P})] \tag{2.27a}$$

Now the structure of the maximization problem changes to a simple water-pouring problem. Its solution is

$$p_i = \left(v - \frac{N_0}{2\delta_i}\right)^+ \tag{2.27b}$$

where the constant v has to be found such that the power constraint on the summation of p_i's is satisfied and δ_i is the i^{th} diagonal entry of **Δ**. Note that the above result is similar to that obtained for the PEP criterion with no receive correlation shown by (2.27b).

No Transmit Correlation: In this case, we have $\mathbf{R}_T = \mathbf{I}$ and consequently, $\mathbf{\Delta} = \mathbf{I}$. Therefore, (2.24) becomes

$$C_{UB} = \log_2[1 + \text{tr}(\mathbf{P})\text{tr}(\mathbf{D}) + 2\det(\mathbf{P})\det(\mathbf{D})] \tag{2.28}$$

The right-hand side of (2.28) has perfect symmetry with respect to diagonal entries of **D**. Furthermore, $\text{tr}(\mathbf{P})$ is constant due to the power constraint. This implies that **P**=α**I** is the solution to the optimization problem, similar to the same case for PEP criterion discussed in the previous section. This indicates that the receive correlation matrix has weaker effect in the power allocation problem in comparison with the transmit correlation matrix.

General Case: Using the Karush-Kuhn-Tucker conditions [56], for a MIMO system, the solution can be found to have a water-pouring form rather than that proposed in [13] and can be expressed as

$$p_i = \left(v - N_0 \frac{\det(\mathbf{D})}{\mathrm{tr}(\mathbf{D})\delta_i} \right)^+ \tag{2.29}$$

where δ_i is the i^{th} diagonal entry of $\mathbf{\Delta}$ and the constant v can be found via a water-pouring process such that the power constraint on the summation of p_i's is satisfied. This answer is in agreement with the above special cases.

Note that, the above derivations for capacity criterion do not assume orthogonality and hence the optimum linear precoder is applicable to both general coded and uncoded MIMO systems.

2.4. Numerical Results

2.4.1. BER Performance of PEP/MMSE Precoder:

We evaluate the "BER versus SNR" performance of MIMO/BPSK systems using Alamouti full-rate [18] STBC with M transmit antennas and N receive antennas. We consider 4 cases: (i) full correlation at both Tx and Rx sides, (ii) full correlation at the Tx side *only*, (iii) *partial* correlation at both Tx and Rx sides, and (iv) *partial* correlation at the Tx side *only*. By '*partial* correlation', we mean that the correlation between each different pair of antenna is less than one. In simulation, we consider ρ=0.5 for adjacent antennas and we assume ρ decreases linearly with distance. For each case, results in Section 2.2 are used to derive the corresponding optimum PEP/MMSE precoder. As discussed before, these two criteria lead to the same design for precoders.

Figures 2.3-2.5 show the simulation results for systems using Alamouti full-rate STBC scheme, in 4 cases (i), (ii), (iii), and (iv), respectively. We assume 2 receive antennas, i.e., N=2, in all cases. The PEP/MMSE precoder provides a gain of 3dB (for M=2) to 5dB (for M=4) in the *full* correlation cases (Figs. 2.3-2.4). This precoding gain is reduced to 2.5dB (for M=2) to 4.5dB (for M=4) in the *partial* correlation cases (Figs. 2.5-

2.6). Comparing Figures 2.3 and 2.4 (or equivalently, Figures 2.5 and 2.6) indicates that precoding gain is mostly due to the correlation in transmit side[2].

The precoder aims to remove the correlation effect by sending signals over the independent eigen-modes of the channel. As shown by the channel model in (2.2), the rightmost matrix is for transmit correlation. Therefore, the precoder tends to send signals on the eigen-modes of the transmit correlation matrix and to load power on its strongest eigen-modes. As the channel model used here has the same receive correlation for all transmit antennas (and hence all eigen-modes), the effect of receive correlation is the same for all transmit eigen-modes.

[2] We also evaluated the "BER versus SNR" performance of MIMO/BPSK systems using Tarokh rate-¾ code [41]. Although the simulation results show a larger precoding gain as compared to systems using Alamouti full-rate STBC scheme, they indicate the same trends, i.e., Precoding gain in the *full* correlation cases is larger than that in the *partial* correlation cases and it is mainly due to the Tx correlation.

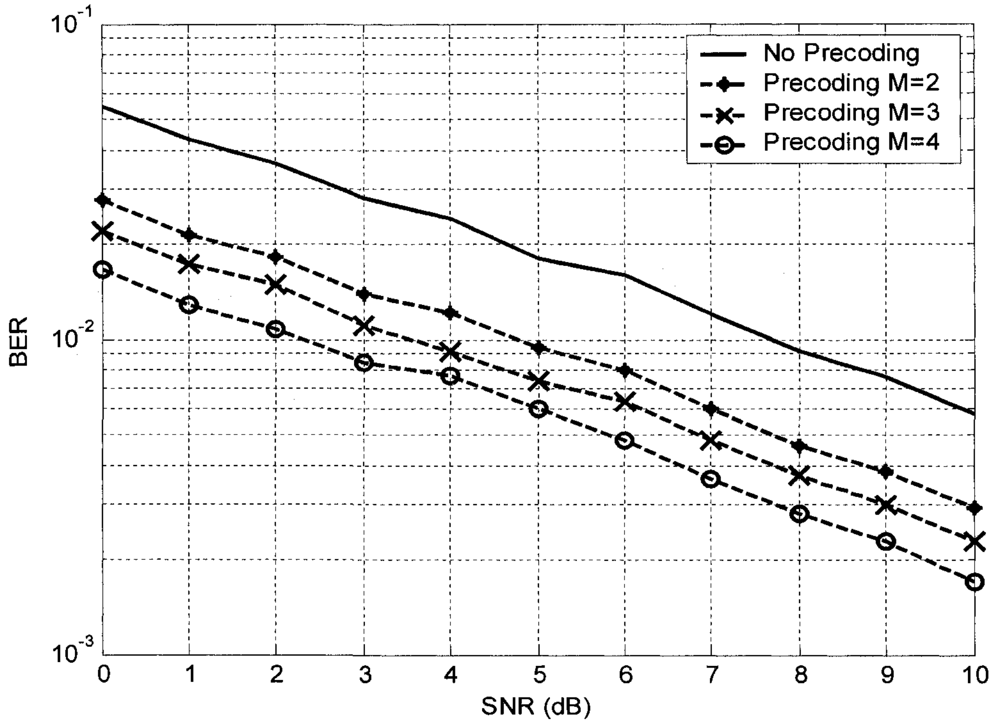


Figure 2.3: BER Performance for *full* correlation at both Tx and Rx sides (with Alamouti Code [18])

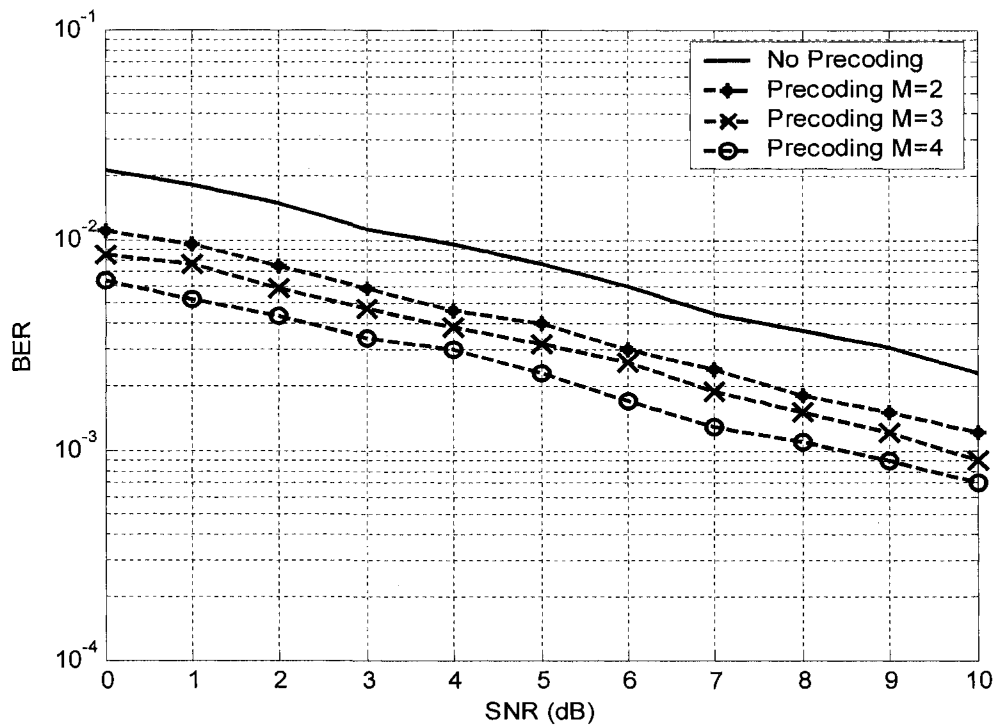


Figure 2.4: BER Performance for *full* correlation at the Tx side *only*

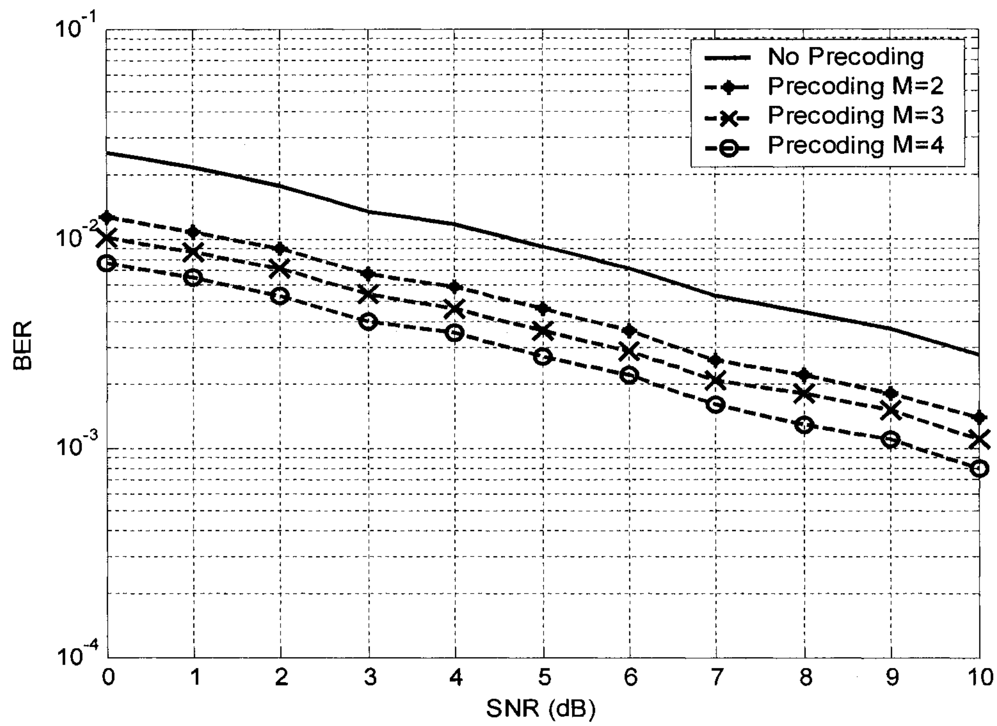


Figure 2.5: BER Performance for *partial* correlation at both Tx and Rx sides

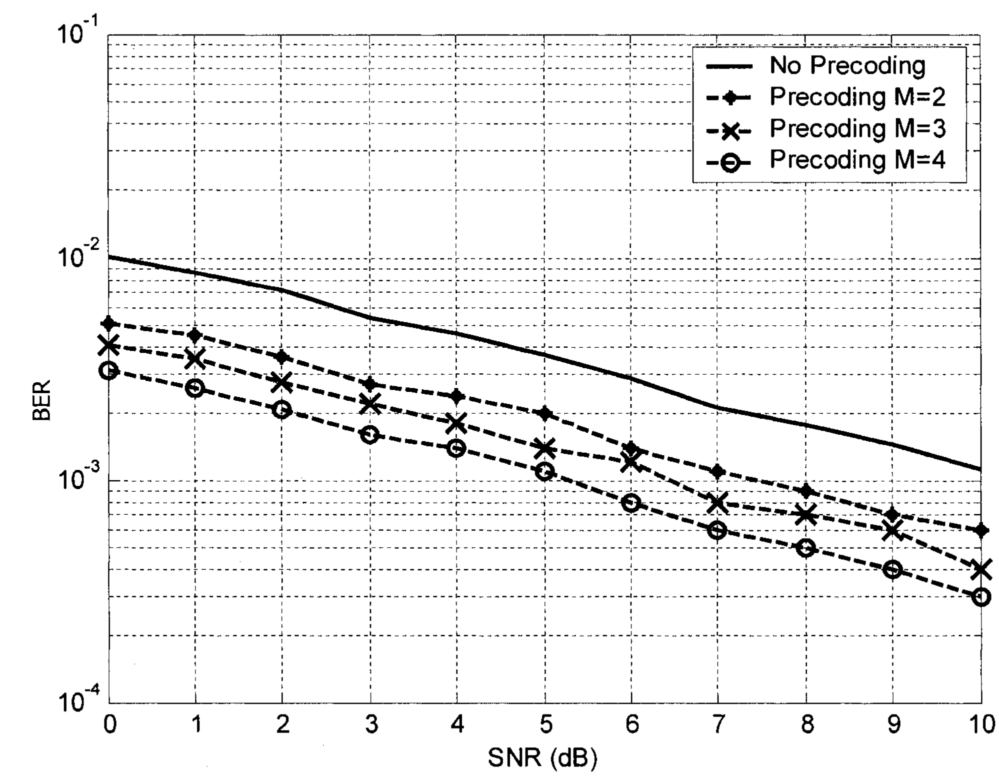


Figure 2.6: BER Performance for *partial* correlation at the Tx side *only*

2.4.2. Performance of Ergodic Capacity Precoder

In this part, we evaluate the ergodic capacity in terms of bits/channel/use versus SNR. First, we compare the ergodic capacity of 2×2 and 3×3 MIMO systems with uncorrelated channel transmission matrix obtained by simulations with its upperbound given by (2.26). We examine the performance of 2×2 systems with and without precoding in 4 cases: (i) full correlation at both Tx and Rx sides, (ii) full correlation at the Tx side *only*, (iii) *partial* correlation at both Tx and Rx sides, and (iv) *partial* correlation at the Tx side *only*, as shown in Figures 2.7-2.10, respectively. For each case, results in Section 2.3 are used to derive the corresponding optimum ergodic-capacity precoder.

In general, as SNR increases, the precoder provides a larger improvement in ergodic capacity. A similar trend is observed: Capacity improvement offered by precoding in full correlation cases (Figures 2.7-2.8) is larger than that in *partial* correlation cases (Figures 2.9-2.10) and correlation at the Tx side has a major effect on the system performance. For example, for an SNR of 5dB, the precoder provides a capacity improvement of 12.9% (from 3.1 to 3.5 bits/channel/use) for the case of full correlation at both Tx and Rx sides (Figure 2.7). The precoding capacity improvement at SNR of 5dB becomes 8.3% (from 3.6 to 3.9 bits/channel/use) for the case of full correlation at the Tx side *only* (Figure 2.8), 7.7% (from 3.9 to 4.2 bits/channel/use) for the case of *partial* correlation at both Tx and Rx sides (Figure 2.9), and 4.9% (from 4.1 to 4.3 bits/channel/use) for the case of *partial* correlation at the Tx side *only* (Figure 2.10).

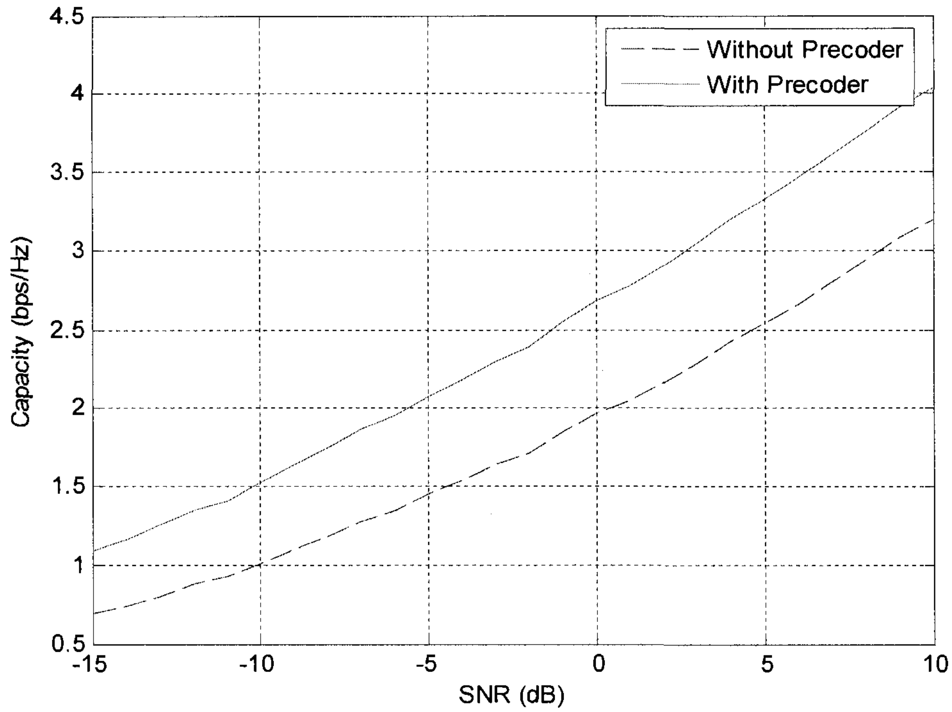


Figure 2.7: Ergodic capacity for *full* correlation at both Tx and Rx sides

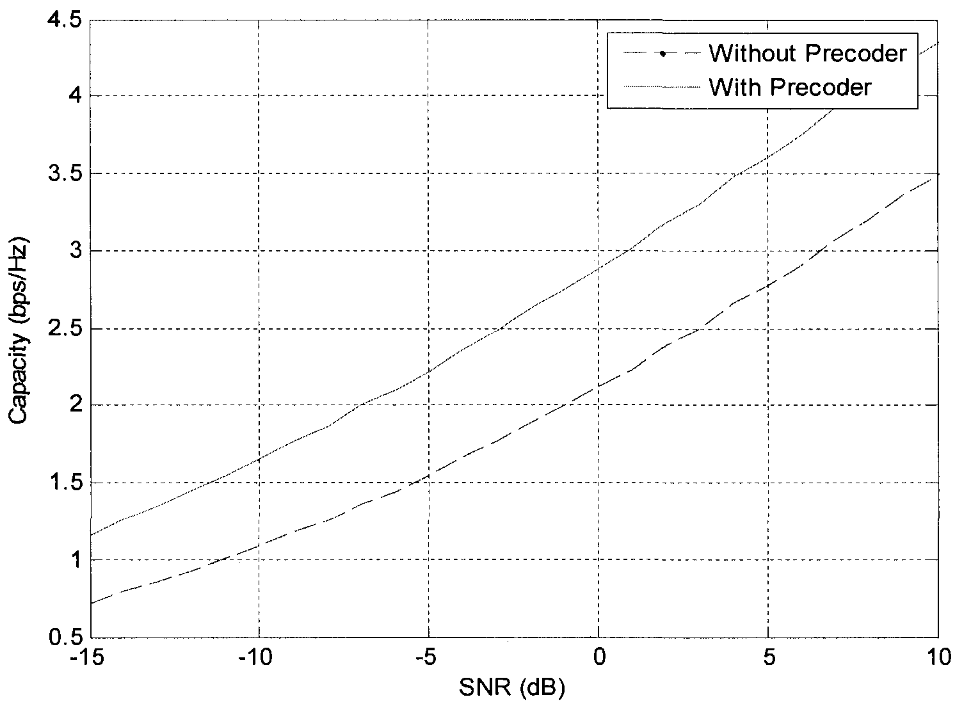


Figure 2.8: Ergodic capacity for *full* correlation in at the Tx side *only*

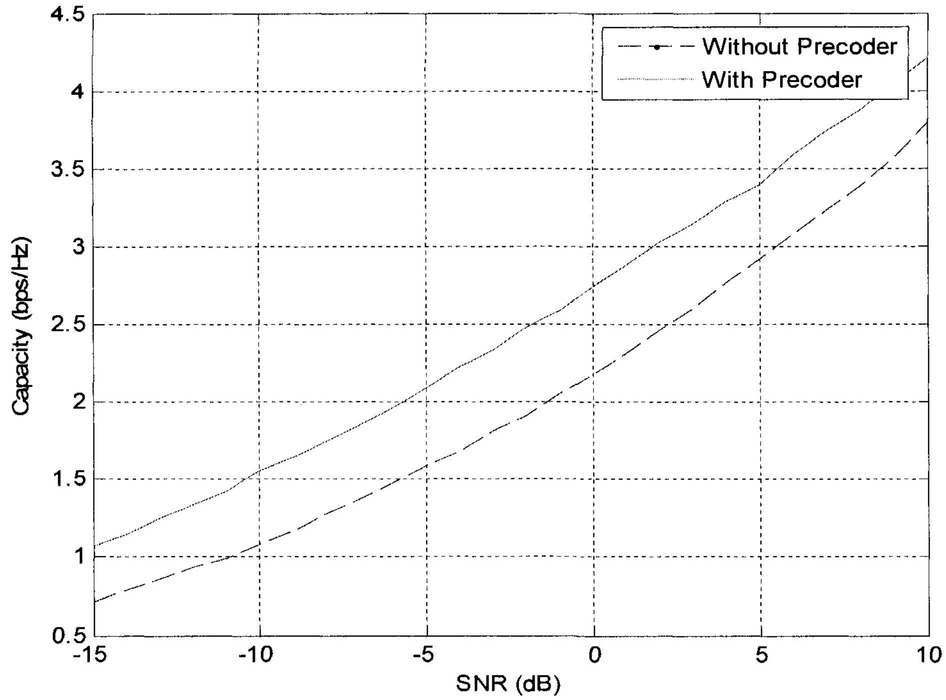


Figure 2.9: Ergodic capacity for *partial* correlation at both Tx and Rx sides

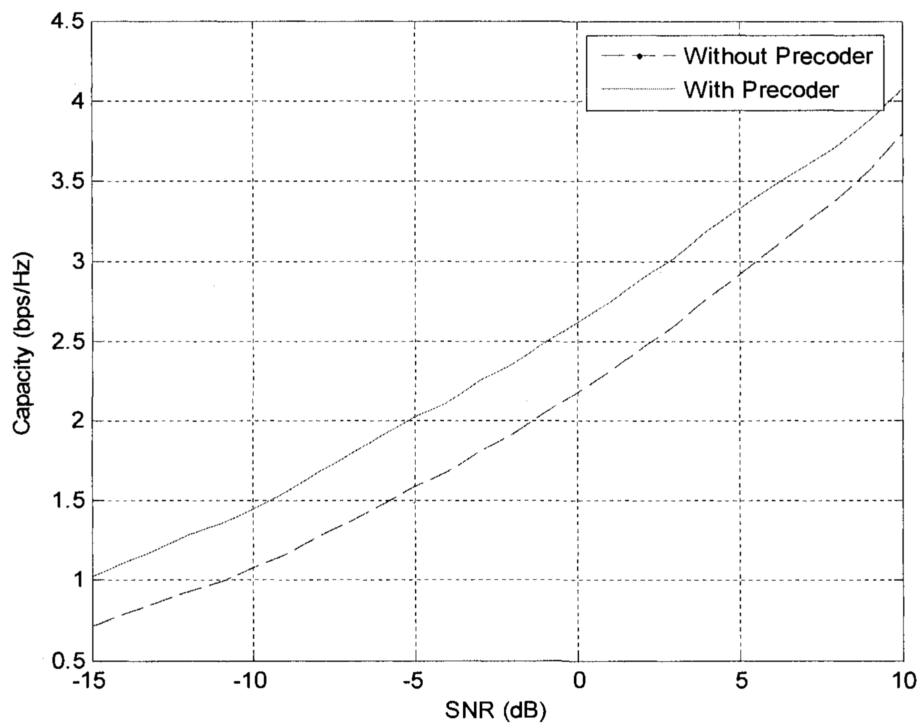


Figure 2.10: Ergodic capacity for *partial* correlation at the Tx side *only*

2.5. Chapter Summary

In this chapter, we investigated three optimal multi-antenna transmitters based on the knowledge of only transmit and receive correlation matrices of the underlying MIMO channel using PEP, MMSE and ergodic capacity as the performance criteria. Also, an upperbound on capacity of MIMO systems was introduced. The optimal precoders are eigen-beamformers, which transmit the signal along the eigenvectors of the transmit correlation matrix and power loading across the eigenbeams is determined based on the eigenvalues of the transmit (and receive) correlation matrix.

Simulation results on performance of MIMO systems with and without precoding in various channel conditions show a significant performance improvement provided by the precoder, especially in the fully correlated cases, i.e., low-rank transmit and receive correlation matrices.

Chapter 3

Partial Channel Knowledge Based Precoding in Frequency-Selective Channels[1]

3.1. MIMO Frequency-Selective Channel Model

MIMO precoder design based on full channel knowledge at the transmitter for performance enhancement in frequency selective channels has been investigated [19]. As mentioned before, it is usually assumed that the transmitter does have the instantaneous channel information and based on that the precoder tries to optimize a performance metric related to the instantaneous channel realizations. The performance metric includes a wide range of criteria such as pair-wise error probability (PEP), mean square error (MSE) and ergodic or outage capacity.

In fast fading channels, however, because of the fast varying nature of the channel, full channel knowledge may be difficult to obtain at the transmitter and hence, it is more reasonable to assume partial channel knowledge at transmitter. Although researches have mainly focused on the problem of precoder design based on partial channel knowledge for frequency flat channels, there are few works in the literature that talk about precoder design for general frequency selective environments based on partial channel knowledge.

There are several reasons for the existence of path correlation in a multi-path communication channel. For example, the scatterers that are far from the transmit or receive antenna can result in a temporal correlation between the multi-path signals that

[1] This chapter has been partially presented in [86], [88], [89] and [90].

are scattered from these scatterers and hence will result in a correlation on each channel link. Also the paths in a keyhole channel [20] are correlated as they all experience the same propagation scenario in some part of the channel (near the pinhole or keyhole). In these channels, a transmitted signal experiences the channel fading before the pinhole and it experiences the delay spread after the pinhole [21].

In [21], a precoder has been designed for an OFDM based MIMO system exploiting transmit and path correlation. However, in the final design the effect of path correlation matrix has been neglected. In this chapter, we consider the precoder designs for general frequency selective channels based on both spatial and path correlation matrices and derive the structures for optimal linear precoders under maximum ergodic (mean) channel capacity. We show that the optimum precoder based on this criterion is an eigen-beamformer and the power allocation policy on each eigen-mode follows a water-pouring strategy that depends on the product of the eigenvalues of transmit and path correlation matrices.

Capacity improvement of the proposed precoders based on partial channel knowledge is investigated in different propagation scenarios such as correlated and uncorrelated channel paths and transmit antennas. Note that for the sake of brevity, we just consider the capacity criterion in this subchapter. Based on the system model described in this section, it is straight-forward to derive the precoders based on PEP and MMSE criteria in frequency selective channels. The derivations are similar to the capacity criterion which is discussed here.

We consider a general MIMO communication system with M-element transmit and N-element receive antenna arrays in a wireless frequency-selective fading environment as shown in Figure 3.1. The equivalent complex baseband channel, including the RF transceivers and broadband wireless frequency-selective fading MIMO environment, is represented as a multi-path model with L effective paths where each path introduced by a scattering cluster can be modeled as an $N \times M$ matrix denoted by $\mathbf{H}_l$ with $l = 0,1,\ldots, L$-1.

Assuming proper pulse shaping at the transmitter and filtering and sampling at proper instances in the receiver, we develop the following discrete time MIMO channel model

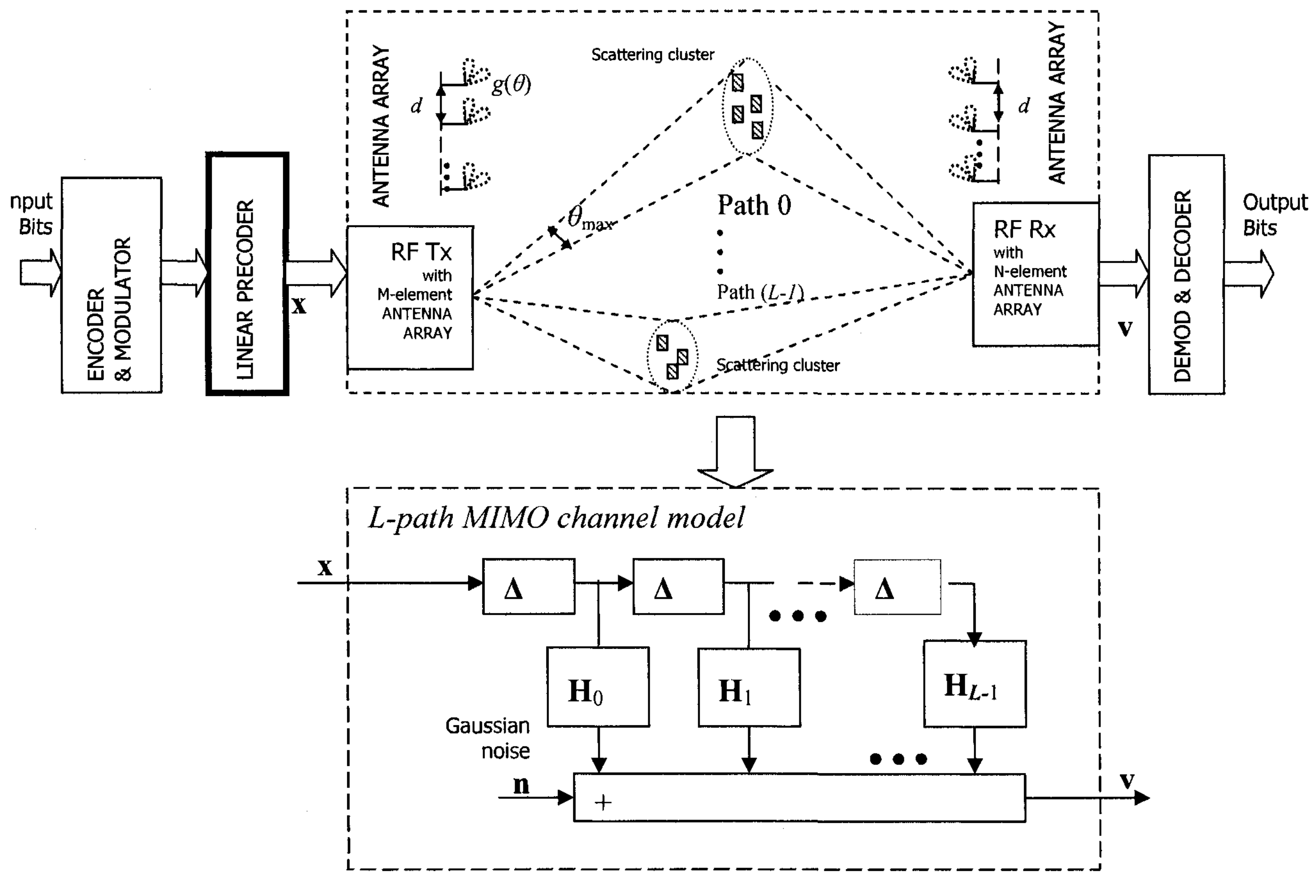


Figure 3.1: A MIMO communication system in a frequency-selective fading environment

for frequency selective channels. At each time instant, the $M\times1$ transmitted complex baseband signal vector, the $N\times1$ received complex baseband signal vector and the $N\times1$ complex baseband Gaussian noise vector at time instant k can be written as: $\mathbf{x}(k)=[x_1(k),x_2(k),...,x_M(k)]^T$, $\mathbf{y}(k)=[y_1(k),y_2(k),...,y_N(k)]^T$ and $\mathbf{n}=[n_1(k),n_2(k),...,n_N(k)]^T$, respectively. The superscript $[.]^T$ denotes the transpose operator.

Consider a transmitted block of $P+L$ vectors of size $M\times1$ where P is an arbitrary integer. We stack them in an $M(P+L)\times1$ vector, $\bar{\mathbf{x}}(k)=[\mathbf{x}^T(k(P+L),...,\mathbf{x}^T(k(P+L)+P+L-1)]^T$. We also stack P received snapshots in a $PN\times1$ vector, $\bar{\mathbf{y}}(k)=[\mathbf{y}^T(k(P+L)+L),...,\mathbf{y}^T(k(P+L)+P+L)]^T$, in which we eliminated the first L vectors to cancel the inter-block interference (IBI)[2]. We obtain the following relation:

[2] A similar assumption was used in [19].

$$\bar{\mathbf{y}}(k) = \bar{\mathbf{H}}.\bar{\mathbf{x}}(k) + \bar{\mathbf{n}}(k) \tag{3.1}$$

where $\bar{\mathbf{n}}(k) = [\mathbf{n}^T(k(P+L)+L),...,\mathbf{n}^T(k(P+L)+P+L)]^T$ is the noise vector, and $\bar{\mathbf{H}}$ is the $N(P+1)\times M(P+L)$ block-Toeplitz channel matrix:

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_{L-1} & \mathbf{H}_{L-2} & \cdots & \mathbf{H}_0 & 0 & \cdots & 0 \\ 0 & \mathbf{H}_{L-1} & \mathbf{H}_{L-2} & \cdots & \mathbf{H}_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \mathbf{H}_{L-1} & \mathbf{H}_{L-2} & \cdots & \mathbf{H}_0 & 0 \\ 0 & \cdots & 0 & \mathbf{H}_{L-1} & \mathbf{H}_{L-2} & \cdots & \mathbf{H}_0 \end{bmatrix} \tag{3.2}$$

The $N\times M$ matrix $\mathbf{H}_l(k)$ represents the spatial response corresponding to path l, $l = 0,1,\ldots, L-1$, at the time instant k, i.e., its entry, $h_{nm}(k)$ is the complex-valued random gain from the m^{th} transmit to n^{th} receive antenna over the effective path l at the instant k, assumed to be unchanged during a frame transmission. Due to the different delays between L effective paths, (3.2) can provide a frequency-selective fading MIMO channel model, while each individual $\mathbf{H}_l(k)$ just represents a frequency-flat fading MIMO channel.

For each resolvable *path*, its transmit (or receive) spatial correlation matrix, $\mathbf{R}_T$ (or $\mathbf{R}_R$) can be formulated with its $(m,k)^{\text{th}}$ component representing the spatial coefficient, ρ_{mk}, between the m^{th} and k^{th} transmit (or receive) antenna elements, e.g., for a simple *uniform* antenna array with equal small antenna spacing d (see, for example, equation (3), page 588 of [47], equation (6) page 176 of [48]),

$$\rho_{m,k} = \int_{-\theta_{\max}/2}^{+\theta_{\max}/2} e^{-j2\pi\frac{(m-k)d}{\lambda}\sin\theta} P_{ang}(\theta)\,d\theta \tag{3.3}$$

where λ is the signal wavelength, $P_{ang}(\theta)$ is the angular power spectrum (APS), and $\theta_{\max}$ represents the angular spread of the path, defined as the maximum angle at which the antenna array can "*see*" the path as shown in Figure 3.1.

The existence of correlation between channel paths has been investigated by some works [21], [49]. Measurement and experimental results show that in many practical cases the assumption of independence of channel paths is not valid. One simple and easy-to-imagine case is the so-called *keyhole* or *pinhole* channels [20] in which due to sheer number of scatterers there is just a small aperture in the middle of the channel that all the signals should go through before being received at the destination. In this case, all of the paths experience the same fading scenarios and therefore are correlated at the receiver.

Modeling path correlation in term of channel physical parameters is a non-trivial task and has not been done extensively in the literature. As an illustrative example to obtain a rough idea of the relation between the physical parameters and path correlation, consider each resolvable path formed by a scattering cluster with the same transmit (or receive) small angular spread ϕ_T (or ϕ_R). By assuming the same simple model of (3.3) in which two clusters play the same role as two antenna elements, the $(i,j)^{\text{th}}$ component of the $L \times L$ path correlation matrix can be approximated as

$$\rho_{ij}^{P} = \rho_{ij}^{T}\rho_{ij}^{R}, \quad \rho_{ij}^{S} = \int_{-\phi_S}^{+\phi_S} e^{-j2\pi\frac{d_{ij}}{\lambda}\sin\theta} P_{ang}(\theta)\,d\theta, \quad S = T \text{ or } R \tag{3.4}$$

where d_{ij} is the average distance between two clusters i and j. It indicates that as d_{ij} and/or ϕ_T (or ϕ_R) decreases, the correlation is increased. In other words, it implies that the replicas of transmitted signals that are caused by scattering from two nearby clusters (i.e., with small d_{ij}) that are far away from transmit and receive antenna arrays (i.e., with small ϕ_T (or ϕ_R)), are highly correlated.

Based on the general channel model in (3.2), the following paragraphs present further derivations of the detailed channel models[3] to characterize path and spatial correlation in the frequency-selective MIMO channels in three different propagation scenarios.

3.1.1. Uncorrelated Paths with the Same Spatial Correlation:

Spatial correlation has been modeled by using a Kronecker product model, in which the channel correlation is the product of the transmit and receive correlation matrices. Kronecker model is mathematically tractable and suitable in many propagation scenarios and hence has been exploited by many works. It is also worth noting that its ability to model some other propagation environments has been questioned in the literature. However, it has been shown that Kronecker model is still valid in many scenarios and its performance and accuracy is acceptable compared to more complex correlation models, e.g. [50]. Therefore, we adopt the Kronecker model for spatial correlation in this chapter. As mentioned in the previous subsection, the fading correlation is governed by the angle spread, antenna spacing, and wavelength [4]. Note that, this structure results from an assumption that transmit and receive scattering radii are large enough and only immediate surroundings of the antennas on one side (transmitter or receiver) have an impact on its

[3] to be used in precoder designs in Section 3.2.

antenna correlation without affecting that of the other side. In this part, we further assume that the angles of departure (AoD) and arrival (AoA) of all channel paths are almost the same. In this case, for the l^{th} path, the path matrix $\mathbf{H}_l(k)$ can be decomposed as:

$$\mathbf{H}_l(k) = \mathbf{R}_R^{1/2}\mathbf{G}_l(k)\mathbf{R}_T^{1/2} \tag{3.5}$$

where $\mathbf{G}_l(k)$ is an $N\times M$ matrix with i.i.d. zero-mean complex Gaussian entries. Note that their variances are not the same since the paths can have different gains. More specifically, $E\{\text{tr}(\mathbf{G}_l\mathbf{G}_l^H)\} = P_l$, where P_l is the power of the l^{th} path. Furthermore, $\mathbf{R}_R$ and $\mathbf{R}_T$ are $N\times N$ and $M\times M$ receive and transmit spatial correlation matrices, respectively. Entries of $\mathbf{R}_R$ and $\mathbf{R}_T$ can be determined from the receive and transmit antenna spacings, angular spread and angular power spectrum of the channel [4] and are the same for the L paths. When the departure and arrival angles of all channel paths are almost equal, the spatial correlation matrices can be the same for all channel paths.

Consider the first row of $\bar{\mathbf{H}}$ in (3.2), $\mathbf{H} = [\mathbf{H}_{L-1}, \mathbf{H}_{L-2}, \ldots, \mathbf{H}_0, 0, \ldots 0]$ as an $N\times M(P+L)$ matrix that consists of $(P+L)$ matrices of size $(N\times M)$. The second row of $\bar{\mathbf{H}}$ in (3.2) is the right-shifted version of the first row by one $(N\times M)$-matrix and can be written as $\mathbf{HE}$, where $\mathbf{E}$ is an $M(P+L)\times M(P+L)$ matrix, $\mathbf{E} = \begin{bmatrix} \mathbf{0}_{M(P+L-1)\times M} & \mathbf{I}_{M(P+L-1)} \\ \mathbf{I}_M & \mathbf{0}_{M\times M(P+L-1)} \end{bmatrix}$, $\mathbf{I}_{M(P+L-1)}$ and $\mathbf{I}_M$ are the $M(P+L\text{-}1)\times M(P+L\text{-}1)$ and $M\times M$ identity matrices, and $\mathbf{0}_{M(P+L-1)\times M}$ and $\mathbf{0}_{M\times M(P+L-1)}$ are the $M(P+L\text{-}1)\times M$ and $M\times M(P+L\text{-}1)$ zero-matrices, respectively. Similarly, the $(j+1)^{th}$ row of $\bar{\mathbf{H}}$ in (3.2) is the right-shifted version of the j^{th} row by one $(N\times M)$-matrix and can be written as the product of the j^{th} row by $\mathbf{E}$ for j=2,3,…, P. In other words, the channel model can be written as:

$$\bar{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \mathbf{HE} \\ \mathbf{HE}^2 \\ \vdots \\ \mathbf{HE}^P \end{bmatrix} = (\mathbf{I}_{P+1} \otimes \mathbf{H}).\begin{bmatrix} \mathbf{I}_{M(L+P)} \\ \mathbf{E} \\ \mathbf{E}^2 \\ \vdots \\ \mathbf{E}^P \end{bmatrix} = (\mathbf{I}_{P+1} \otimes \mathbf{H}).\bar{\mathbf{E}} \tag{3.6}$$

where $\mathbf{I}_{P+1}$ is the $(P+1)\times(P+1)$ identity matrix and $\otimes$ stands for Kronecker product. $\mathbf{E}$ has the following properties:

- $\forall i,\ 0 < i \leq P : \det(\mathbf{E}^i) = \pm 1$, $\mathbf{E}^i(\mathbf{E}^i)^T = \mathbf{I}$ and hence $\mathbf{E}^i(\mathbf{E}^i)^T$ has the same eigenvalues and eigenvectors as the identity matrix. The eigenvalues of $\mathbf{E}^i$ have the same absolute values as those of the identity matrix.
- Given an $M(P+L)\times M(P+L)$-matrix $\mathbf{A}$ and the eigen-decompositions of $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ and $\mathbf{B} = \mathbf{A}\mathbf{E}^i = \mathbf{U}_1\mathbf{\Lambda}_1\mathbf{U}_1^H$, $\mathbf{\Lambda}_1 = \mathbf{\Lambda}$ and $\mathbf{U}_1 = \mathbf{U}\mathbf{E}^i$.

  Proof: Because of the first property of $\mathbf{E}^i$, $\mathbf{A}$ and $\mathbf{B}$ are in fact *similar*. Two similar matrices have the same set of eigenvalues [23]. For any $\mathbf{x}$ to be an eigenvector of $\mathbf{B}$ corresponding to eigenvalue λ, one can then write $\mathbf{B}\mathbf{x} = \lambda\mathbf{x}$, then $(\mathbf{E}^i)^T\mathbf{A}\mathbf{E}^i\mathbf{x} = \lambda\mathbf{x}$, or $\mathbf{A}\mathbf{E}^i\mathbf{x} = \lambda\,\mathbf{E}^i\mathbf{x}$ and hence $\mathbf{E}^i\mathbf{x}$ is an eigenvector of $\mathbf{A}$. Vice versa, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then $(\mathbf{E}^i)^T\mathbf{x}$ will be an eigenvector of $\mathbf{B}$.

■

From the above properties, (3.5) and (3.6), the channel model can be written as:

$$\bar{\mathbf{H}} = (\mathbf{I}_{P+1} \otimes \mathbf{R}_R^{1/2}\mathbf{G}(\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{1/2})).\bar{\mathbf{E}} \tag{3.7}$$

in which $\mathbf{R}_R$ and $\mathbf{R}_T$ are the $N\times N$ and $M\times M$ receive and transmit correlation matrices defined in (3.5), and $\mathbf{G} = [\mathbf{G}_{L-1}, \mathbf{G}_{L-2}, \ldots, \mathbf{G}_0, 0, \ldots, 0]$ is an $N\times M(L+P)$-matrix composed of $(P+L)$ matrices of size $N\times M$.

3.1.2. Correlated Paths with the Same Spatial Correlation

We consider now the case of L correlated channel paths with the same spatial transmit and receive correlation matrices for all L paths. From (3.6), the channel model can be written as:

$$\bar{\mathbf{H}} = (\mathbf{I}_{P+1} \otimes \mathbf{R}_R^{1/2}\mathbf{G}(\mathbf{R}_P^{1/2} \otimes \mathbf{R}_T^{1/2})).\bar{\mathbf{E}} \tag{3.8}$$

where $\mathbf{G}$ is the same $N\times M(L+P)$ matrix as in the previous case except that, for simplicity, the entries of $\mathbf{G}_l$'s are assumed to be i.i.d. zero-mean complex Gaussian with ½ variance per dimension (in fact, we consider the power of paths in the diagonal entries of path correlation matrix). $\mathbf{R}_R$ and $\mathbf{R}_T$ are the $N\times N$ and $M\times M$ receive and transmit correlation matrices and $\mathbf{R}_P$ is the $(P+L)\times(P+L)$ path correlation matrix. Elements of $\mathbf{R}_P$ show the correlation between different paths. Note that $\mathbf{R}_P$ has only L^2 non-zero elements. Therefore, to feedback the path correlation matrix $\mathbf{R}_P$ to the transmitter, it is sufficient to send L^2 values (not $(P+L)^2$) and other entries of this matrix are zero.

An example of this scenario is when the paths generated from the two scattering clusters travel in the same directions and therefore would experience possibly correlated propagation situations. It will result in correlated fading between these two paths. For tractability in analysis, path correlation is assumed to be separated from the transmit and receive antenna correlation, which corresponds to the situation of multipaths caused by scattering clusters seen in a narrow angular range and at a long distance from the transmit and receive antenna surroundings [21].

3.1.3. Uncorrelated Paths with Different Spatial Correlation

Unlike the previous case, the spatial correlation matrices in this case are not the same for all channel paths, i.e., (3.5) is not valid. We can write each channel path matrix as:

$$\mathbf{H}_l(k) = \mathbf{R}_{R,l}{}^{1/2}\mathbf{G}_l(k)\mathbf{R}_{T,l}{}^{1/2} \tag{3.9}$$

where $\mathbf{R}_{R,l}$ and $\mathbf{R}_{T,l}$ are the $N\times N$ and $M\times M$ receive and transmit correlation matrices associated with l^{th} channel path.

A possible propagation scenario leading to this channel structure is when the channel is composed of a number of local scattering clusters located in the vicinity of transmitter and receiver, but with different angles of arrival (AoA) and departure (AoD). In this scenario, the spatial correlation matrices are different for different paths. As a result, the channel structure of the previous case in (3.6) cannot be used. In other words, the effects of transmit and receive correlation need to be considered as follows. Assuming that every channel path has different transmit and receive correlation matrices, one can develop the following model:

$$\overline{\mathbf{H}} = (\mathbf{I}_{P+1}\otimes\mathbf{H}).\overline{\mathbf{E}} = (\mathbf{I}_{P+1}\otimes\mathbf{R}_R^{1/2}\mathbf{G}\mathbf{R}_T^{1/2})\overline{\mathbf{E}} = (\mathbf{I}_{P+1}\otimes\mathbf{R}_R^{1/2})(\mathbf{I}_{P+1}\otimes\mathbf{G})(\mathbf{I}_{P+1}\otimes\mathbf{R}_T^{1/2})\overline{\mathbf{E}} \tag{3.10}$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{L-1} & \mathbf{0}_{N\times M} & \mathbf{0}_{N\times M} & \cdots & \cdots & \mathbf{0}_{N\times M} & \mathbf{0}_{N\times PM} \\ \mathbf{0}_{N\times M} & \mathbf{G}_{L-2} & \mathbf{0}_{N\times M} & \cdots & \vdots & \vdots & \vdots \\ \vdots & \mathbf{0}_{N\times M} & \ddots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \cdots & \ddots & \mathbf{0}_{N\times M} & \mathbf{0}_{N\times PM} \\ \mathbf{0}_{N\times M} & \mathbf{0}_{N\times M} & \cdots & \cdots & \mathbf{0}_{N\times M} & \mathbf{G}_0 & \mathbf{0}_{N\times PM} \\ \mathbf{0}_{PN\times M} & \mathbf{0}_{PN\times M} & \cdots & \cdots & \mathbf{0}_{PN\times M} & \mathbf{0}_{PN\times M} & \mathbf{0}_{PN\times PM} \end{bmatrix}$$

is an $N(P+L)\times M(P+L)$ block diagonal matrix whose elements, $\mathbf{G}_l$'s are $N\times M$ matrices with i.i.d. zero-mean, complex Gaussian entries and ½ variance per dimension. The remaining entries are zero, i.e., $\mathbf{0}_{N\times M}$ denotes an $N\times M$ zero matrix. Furthermore, $\mathbf{R}_R$ and $\mathbf{R}_T$ are $N\times N(P+L)$ and $M(P+L)\times M(P+L)$ receive and transmit correlation matrices with the following structures:

$$\mathbf{R}_R = [\mathbf{R}_{R_{L-1}}, \mathbf{R}_{R_{L-2}}, \ldots, \mathbf{R}_{R_0}, 0, \ldots, 0]$$

$$\mathbf{R}_T = \begin{bmatrix} \mathbf{R}_{T,L-1} & & & & & & \\ & \mathbf{R}_{T,L-2} & & & 0 & & \\ & & \ddots & & & & \\ & & & \mathbf{R}_{T,0} & & & \\ & & & & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \tag{3.11}$$

where $\mathbf{R}_{R,l}$ and $\mathbf{R}_{T,l}$ are, respectively, the $N\times N$ receive and $M\times M$ transmit correlation matrices associated with the l^{th} channel path as defined in (3.9).

Now the system model consists of a channel with M and N transmit and receive antennas and L channel paths and precoding/decoding matrices at the transmitter and receiver can be written as:

$$\bar{\mathbf{y}}(k) = \mathbf{W}_R \bar{\mathbf{H}}(k) \mathbf{W}_T \bar{\mathbf{x}}(k) + \mathbf{W}_R \bar{\mathbf{n}}(k) \tag{3.12}$$

where $\bar{\mathbf{y}}(k)$ and $\bar{\mathbf{n}}(k)$ is the $N(P+1)\times 1$ output vector by stacking $P+1$ subsequent $N\times 1$ receive and noise vectors $\mathbf{y}$ and $\mathbf{n}$ in long vectors (for suitable elimination of the IBI effects as previously discussed), $\bar{\mathbf{x}}(k)$ is the $M(P+L)\times 1$ transmitted vector that can be constructed by stacking $P+L$ subsequent $M\times 1$ input vectors $\mathbf{x}$ in (3.1) in a vector, $\bar{\mathbf{H}}$ is $N(P+1)\times M(P+L)$ composite channel matrix defined in (3.1) and $\mathbf{W}_T$ and $\mathbf{W}_R$ are the $M(P+L)\times M(P+L)$ and $M(P+L)\times N(P+1)$ precoder and decoder matrices, respectively. After applying $\mathbf{W}_R$ over the received vectors an ML decoder detects the signal vectors. Our objective is to find $\mathbf{W}_T$ and $\mathbf{W}_R$ based on a performance metric while we assume that we have full channel knowledge at the receiver but just spatial and path correlation matrices at the transmitter.

Transmit and receive correlation matrices can be estimated by taking time average over measured samples of channel path matrices, i.e., for the l^{th} path, $\mathbf{R}_{T,l} = E\{\mathbf{H}_l^H \mathbf{H}_l\}$ and $\mathbf{R}_{R,l} = E\{\mathbf{H}_l \mathbf{H}_l^H\}$. In practice, they can be continuously updated by using the simple exponential-smoothing moving average technique[4] applied to the past K channel path matrix samples as

$$\mathbf{R}_{T,l}(n) = (1-\beta)\mathbf{R}_{T,l}(n-1) + \beta \frac{1}{K} \sum_{k=(n-1)K}^{nK-1} \mathbf{H}_l^H(k)\mathbf{H}_l(k) \tag{3.13}$$

and

$$\mathbf{R}_{R,l}(n) = (1-\beta)\mathbf{R}_{R,l}(n-1) + \beta \frac{1}{K} \sum_{k=(n-1)K}^{nK-1} \mathbf{H}_l(k)\mathbf{H}_l^H(k) \tag{3.14}$$

where $\mathbf{R}_{T,l}(i)$ and $\mathbf{R}_{R,l}(i)$ are the l^{th} channel path transmit and receive correlation matrices at the i^{th} time frame, $\mathbf{H}_l(k)$ is the l^{th} channel path matrix at time instant k. The forgetting factor β and number of frames K are adjusted based on channel correlation changes.

It can be verified from (3.8) that $\mathbf{R}_P \otimes \mathbf{R}_T = \frac{1}{L-1} E\{\bar{\mathbf{H}}_1^H \bar{\mathbf{H}}_1\}$ where $\bar{\mathbf{H}}_1$ is the first block row of $\bar{\mathbf{H}}$. The same moving average technique can be used to calculate $E\{\bar{\mathbf{H}}_1^H \bar{\mathbf{H}}_1\}$ and $\mathbf{R}_T = E\{\mathbf{H}_l^H \mathbf{H}_l\}\,(\forall l)$ to derive the path correlation matrix $\mathbf{R}_P$.

Note that in real systems, usually receiver calculates the correlation matrices and feeds them back to the transmitter via a low-speed feedback channel.

3.2. Capacity-Approaching Precoding Structure

We assume that the receiver has the perfect channel information but transmitter knows only spatial and path correlation matrices. Our objective is to design the precoder ($\mathbf{W}_T$) matrix to maximize the ergodic capacity for a given total transmit power. Using the channel matrix descriptions in the previous subsection, the capacity proof is a simple extension of the well-known capacity result for single-input single-output (SISO)

[4] This simple technique is presented as an illustrative example. Estimation techniques for correlation matrices and properties are a separate issue beyond the scope of this chapter and can be found, for example, in [52].

channels with memory, similar to the scenario discussed in [51]. Subsequently, for sufficiently large values of P, the ergodic capacity per dimension can be derived as[5]

$$\begin{aligned} C &\cong \frac{1}{P} E\{\log_2 \frac{\det(\mathbf{W}_R(\mathbf{I}+\gamma\bar{\mathbf{H}}\mathbf{W}_T\mathbf{W}_T^H\bar{\mathbf{H}}^H)\mathbf{W}_R^H)}{\det(\mathbf{W}_R\mathbf{W}_R^H)}\} \\ &= \frac{1}{P} E\{\log_2 \det(\mathbf{I}_{N(P+1)}+\gamma\mathbf{W}_R\bar{\mathbf{H}}\mathbf{W}_T\mathbf{W}_T^H\bar{\mathbf{H}}^H\mathbf{W}_R^H(\mathbf{W}_R\mathbf{W}_R^H)^{-1})\} \end{aligned} \tag{3.15}$$

where $\mathbf{W}_T$ is the $M(P+L)\times M(P+L)$ precoding matrix, $\mathbf{W}_R$ is the $M(P+L)\times N(P+1)$ decoding matrix and γ is the channel signal to noise ratio (SNR), $E\{.\}$ denotes expectation over different realizations of channel matrix $\bar{\mathbf{H}}$ and the superscript ^{H} denotes the Hermitian transposition. Assuming that the spatial and path correlation matrices are available at the transmitter, we want to design the precoding and decoding matrices $\mathbf{W}_T$ and $\mathbf{W}_R$ to maximize (3.15) under the power constraint criterion on $\mathbf{W}_T$ based on the limited power at the transmitter.

At first, consider the optimization of decoder matrix $\mathbf{W}_R$. With the assumed knowledge of the *instantaneous* channel information at the receiver, a reasonable criterion to design a linear receiver $\mathbf{W}_R$, for given $\bar{\mathbf{H}}$ and $\mathbf{W}_T$, is to maximize *instantaneous* mutual information (rather than its *average*):

$$I(\bar{\mathbf{y}};\bar{\mathbf{x}}) = \log_2(\det(\mathbf{I}_{N(P+1)}+\gamma\mathbf{W}_R\bar{\mathbf{H}}\mathbf{W}_T\mathbf{W}_T^H\bar{\mathbf{H}}^H\mathbf{W}_R^H(\mathbf{W}_R\mathbf{W}_R^H)^{-1})) \tag{3.16}$$

***Lemma* 3.1:** Any decoding matrix of form $\mathbf{W}_R = \mathbf{\Gamma}\mathbf{W}_T^H\bar{\mathbf{H}}^H$ maximizes the *instantaneous* mutual information in (3.16) with an arbitrary $M(P+L)\times M(P+L)$ matrix $\mathbf{\Gamma}$ which $\mathbf{\Gamma}\,\mathbf{\Gamma}^H=\mathbf{I}$.

Proof: For brevity and simplicity, assume *white* noise at the decoder input with identity noise correlation matrix, i.e., $\mathbf{R}_n=\mathbf{I}$. The mutual information in (3.16) can be written as:

$$\begin{aligned} I(\bar{\mathbf{y}};\bar{\mathbf{x}}) &= \log_2(\det(\mathbf{I}_{M(P+L)}+\gamma\mathbf{W}_T^H\bar{\mathbf{H}}^H\mathbf{W}_R^H(\mathbf{W}_R\mathbf{W}_R^H)^{-1}\mathbf{W}_R\bar{\mathbf{H}}\mathbf{W}_T)) \\ &\leq \log_2(\det(\mathbf{I}_{M(P+L)}+\gamma\mathbf{W}_T^H\bar{\mathbf{H}}^H\bar{\mathbf{H}}\mathbf{W}_T)) \end{aligned} \tag{3.17}$$

[5] The ergodic capacity of a frequency-selective channel based on the model (3.1), can be found by taking the limit on the average mutual information when the dimension of the channel matrix approaches infinity [53]-[55]: $C = \lim_{P\to\infty}\frac{1}{P}E\{\log\det(\mathbf{I}+\gamma\mathbf{H}\mathbf{H}^H)\}$

by using the matrix identity $\det(\mathbf{I}+\mathbf{AB})=\det(\mathbf{I}+\mathbf{BA})$. The inequality in (3.17) comes from the fact that $\mathbf{W}_R$ is an $M(P+L)\times N(P+1)$ matrix and hence, in general, $\mathbf{W}_R^H(\mathbf{W}_R\mathbf{W}_R^H)^{-1}\mathbf{W}_R \leq \mathbf{I}$. It is clear that (17) can achieve its maximum when $\bar{\mathbf{H}}\mathbf{W}_T$ is in the range space of $\mathbf{W}_R^H$, i.e., $\mathbf{W}_R^H = \bar{\mathbf{H}}\mathbf{W}_T\mathbf{\Pi}$ where $\mathbf{\Pi}$ is an arbitrary invertible $M(P+L)\times M(P+L)$ matrix where $\mathbf{\Pi}\mathbf{\Pi}^{\mathrm{H}} = \mathbf{I}$. Therefore, $I(\bar{\mathbf{y}};\bar{\mathbf{x}})$ is maximized when $\mathbf{W}_R = \mathbf{\Gamma}\mathbf{W}_T^H\bar{\mathbf{H}}^H$ with $\mathbf{\Gamma} = \mathbf{\Pi}^H$. More specifically, $\mathbf{\Gamma}$ is a unitary matrix. In all the derivations, we assume that $M(P+L)\leq N(P+1)$. It is, however, easy to generalize the derivations to the case that $M(P+L) > N(P+1)$.

■

Substituting the decoding matrix $\mathbf{W}_R = \mathbf{\Gamma}\mathbf{W}_T^H\bar{\mathbf{H}}^H$ into (11) will result in:

$$C = \frac{1}{P}E\{\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H\bar{\mathbf{H}}^H\bar{\mathbf{H}}\mathbf{W}_T)\} \tag{3.18}$$

Our objective is now finding the optimal precoder matrix $\mathbf{W}_T$, to maximize the ergodic capacity based on the partial channel knowledge of only the spatial and path correlation matrices, $\mathbf{R}_T$, $\mathbf{R}_R$ and $\mathbf{R}_P$, available at the transmitter. Note that, while the decoding matrix, $\mathbf{W}_R$, can be updated at each time instant by using the instantaneous channel information at the receiver, the *precoding* matrix at the transmitter is only needed to re-compute over a long interval whenever the spatial and path correlation matrices are changed.

We will first derive the capacity-approaching precoding matrix for *uncorrelated* paths and then extend the solutions to the *correlated* channel paths.

3.2.1. Uncorrelated Channel Paths

In this case, substituting the channel matrix from (3.7) in (3.18), we obtain:

$$C = \frac{1}{P}E\{\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H\bar{\mathbf{E}}^T(\mathbf{I}_{P+1}\otimes(\mathbf{I}_{P+L}\otimes\mathbf{R}_T^{H/2})\mathbf{G}^H\mathbf{R}_R^{H/2}) \\ (\mathbf{I}_{P+1}\otimes\mathbf{R}_R^{1/2}\mathbf{G}(\mathbf{I}_{P+L}\otimes\mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T)\} \tag{3.19}$$

Getting the expectation of the above equation is very tough if not impossible. Instead, we try to find an upper bound on the capacity, which of course will result in a near-optimum solution to the precoder problem. We use the well-known Jensen's inequality [16] to the sum of these concave functions to optimize an upper bound on the ergodic capacity for its analytical tractability. Applying Jensen's inequality to (3.19) leads to:

$$C \le C_{UP} = \frac{1}{P}\log_2 \det E\{\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H \bar{\mathbf{E}}^T (\mathbf{I}_{P+1} \otimes (\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{H/2})\mathbf{G}^H \mathbf{R}_R^{H/2}) (\mathbf{I}_{P+1} \otimes \mathbf{R}_R^{1/2}\mathbf{G}(\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T\} \tag{3.20}$$

***Lemma* 3.2:** The $M(P+L)\times M(P+L)$ precoding matrix that optimizes the right hand side of (20) can be written as: $\mathbf{W}_T = \text{diag}(\mathbf{W}_i)$, $i = 0,1,\ldots,(L+P\text{-}1)$ where $\mathbf{W}_i$ is an $M\times M$ matrix. The problem can be reduced to a symbol-wise precoding problem in which the selected candidate is an $M\times M$ precoder applies to each of the $P+L$ vectors, separately. Furthermore, candidate $\mathbf{W}_i$'s can be found using the eigen-decomposition of the transmit correlation matrix and have the form $\mathbf{W}_i = \mathbf{\Phi}\mathbf{\Sigma}_i^{1/2}\mathbf{\Gamma}_i$, where $\mathbf{\Gamma}_i$'s are $M\times M$ arbitrary unitary matrices, $\mathbf{\Sigma}_i$'s are $P+L$ diagonal matrices, and $\mathbf{\Phi}$ is the $M\times M$ transmit eigenvector matrix resulting from eigen decomposition of $\mathbf{R}_T$.

Proof: Using the properties of Kronecker product, one can rewrite (3.20) as:

$$C \le \frac{1}{P}\log_2 \det E\{\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H \bar{\mathbf{E}}^T (\mathbf{I}_{P+1} \otimes (\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{H/2})\mathbf{G}^H \mathbf{R}_R^{H/2}\mathbf{R}_R^{1/2} \mathbf{G}(\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T\} \tag{3.21}$$

By using matrix identity $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$ and the eigen-decompositions of the nonnegative definite matrix $\mathbf{R}_R^{H/2}\mathbf{R}_R^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^H$ where $\mathbf{V}$ and $\mathbf{V}^H$ are unitary matrices and $\mathbf{D}$ is a diagonal matrix that contains non-negative eigenvalues, (3.21) can be written as:

$$C \le C_{UB} = \frac{1}{P}\log_2 \det E\{\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H \bar{\mathbf{E}}^T (\mathbf{I}_{P+1} \otimes (\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{H/2})\hat{\mathbf{G}}^H \mathbf{D}\hat{\mathbf{G}}(\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T\} \tag{3.22}$$

Since $\mathbf{V}$ and $\mathbf{V}^H$ are unitary matrices, $\hat{\mathbf{G}} = \mathbf{V}^H\mathbf{G}$ and $\hat{\mathbf{G}}^H = \mathbf{G}^H\mathbf{V}$ have the same distribution as $\mathbf{G}$ and $\mathbf{G}^H$, i.e., they are zero-mean circularly symmetric complex Gaussian random matrices. By taking the expectation in (3.22) and some simple manipulations, we obtain:

$$C \le C_{UB} = \frac{1}{P}\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H \bar{\mathbf{E}}^T [\mathbf{I}_{P+1} \otimes (\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{H/2})] [\mathbf{I}_{P+1} \otimes (\text{tr}(\mathbf{D})\mathbf{P} \otimes \mathbf{I}_M)(\mathbf{I}_{P+L} \otimes \mathbf{R}_T^{1/2})]\bar{\mathbf{E}}\mathbf{W}_T) \tag{3.23}$$

where $\mathbf{P}$ is a $(P+L)\times(P+L)$ diagonal matrix whose first L diagonal entries are powers of channel paths, P_i, i=0,1,…, (L-1), and the P remaining diagonal entries are zero.

The right-hand side matrices in (3.23) can be multiplied and hence it can be written as the sum of P products:

$$C \le C_{UB} = \frac{1}{P}\log_2 \det(\mathbf{I}_{M(P+L)} + \sum_{l=0}^{P} \gamma \mathrm{tr}(\mathbf{D})(\mathbf{E}^l)^T(\mathbf{P}\otimes\mathbf{I}_M)(\mathbf{I}_{P+L}\otimes\mathbf{R}_T)\mathbf{E}^l\mathbf{W}_T\mathbf{W}_T^H) \quad (3.24)$$

Using the properties of $\mathbf{E}$ and based on the special structure of diagonal matrices $(\mathbf{I}_{P+L}\otimes\mathbf{R}_T)$ and $(\mathbf{P}\otimes\mathbf{I}_M)$, it is interesting to see that:

$$\sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{P}\otimes\mathbf{I}_M)(\mathbf{I}_{P+L}\otimes\mathbf{R}_T)\mathbf{E}^l = (\mathbf{P}_T\otimes\mathbf{I}_M)(\mathbf{I}_{P+L}\otimes\mathbf{R}_T) \quad (3.25)$$

where $\mathbf{P}_T\otimes\mathbf{I}_M = \sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{P}\otimes\mathbf{I}_M)\mathbf{E}^l$ is a (P+L)×(P+L) diagonal matrix constructed by summation of entries of $\mathbf{P}$. For an assumed P>L, $\mathbf{P}_T$ has the following form:

$$\mathbf{P}_T = \begin{bmatrix} P_{L-1}+\ldots+P_0 & & & & & & & & & & \\ & P_{L-2}+\ldots+P_0 & & & & & \mathbf{0} & & & & \\ & & \ddots & & & & & & & & \\ & & & P_0 & & & & & & & \\ & & & & P_{L-1} & & & & & & \\ & & & & & P_{L-1}+P_{L-2} & & & & & \\ & & & & & & \ddots & & & & \\ & & \mathbf{0} & & & & & P_{L-1}+\ldots+P_1 & & & \\ & & & & & & & & P_{L-1}+\ldots+P_0 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & P_{L-1}+\ldots+P_0 \end{bmatrix}$$

Now, Consider the eigen-decompositions $\mathbf{W}_T\mathbf{W}_T^H = \mathbf{\Psi}\mathbf{\Sigma}\mathbf{\Psi}^H$ and $\mathbf{R}_T = \mathbf{\Phi}\mathbf{\Delta}\mathbf{\Phi}^H$. Therefore, (3.24) can be rewritten as:

$$C \le C_{UB} = \frac{1}{P}\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathrm{tr}(\mathbf{D})(\mathbf{P}_T\otimes\mathbf{I}_M)(\mathbf{I}_{P+L}\otimes\mathbf{\Phi}\mathbf{\Delta}\mathbf{\Phi}^H)\mathbf{\Psi}\mathbf{\Sigma}\mathbf{\Psi}^H) \quad (3.26)$$

Our aim is to maximize C_{UB} in (3.26) subject to the constraint on the total transmit power, i.e., $\mathrm{tr}(\mathbf{W}_T\mathbf{W}_T^H) \le \mathrm{constant}$. Using Hadamard's inequality [23], it can be shown that we need $\mathbf{\Psi}(\mathbf{I}_{P+L}\otimes\mathbf{\Phi}^H) = \mathbf{\Psi}^H(\mathbf{I}_{P+L}\otimes\mathbf{\Phi}) = \mathbf{I}_{M(P+L)}$ or $\mathbf{\Psi} = (\mathbf{I}_{P+L}\otimes\mathbf{\Phi})$. This identity ensures that the argument of the determinant function in (3.26) is a diagonal matrix. Therefore, the precoding matrix can be written as:

$$\mathbf{W}_T = (\mathbf{I}_{P+L}\otimes\mathbf{\Phi})\mathbf{\Sigma}^{1/2}\mathbf{\Gamma} = \mathrm{diag}(\mathbf{\Phi}\mathbf{\Sigma}_i^{1/2}\mathbf{\Gamma}_i) \qquad i = 0,1,\ldots,P+L-1 \quad (3.27)$$

where $\mathbf{\Gamma}_i$ is an arbitrary unitary matrix that has no effect on the system performance and therefore can be set to identity for simplicity. Therefore, $\mathbf{W}_T$ will be a block diagonal matrix with diagonal block $\mathbf{W}_i = \mathbf{\Phi}\mathbf{\Sigma}_i^{1/2}\mathbf{\Gamma}_i$ and it proves the *Lemma* 3.2.

■

In other words, the selected linear precoding matrix $\mathbf{W}_T$ consists of $P+L$ eigen-beamformers whose orthogonal beams point to the eigenvectors of the transmit correlation matrix, $\mathbf{R}_T$. Furthermore, the diagonal matrices $\mathbf{\Sigma}_i$'s are in fact the power loadings on each of the eigen-beamformers.

The power loading policies across the eigen-beams can be obtained by substituting $\mathbf{\Psi} = (\mathbf{I}_{P+L} \otimes \mathbf{\Phi})$ in (3.26) as:

$$\max_{\mathbf{\Sigma}} \log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathrm{tr}(\mathbf{D})(\mathbf{P}_T \otimes \mathbf{I}_M)(\mathbf{I}_{P+L} \otimes \mathbf{\Delta})\mathbf{\Sigma}) \text{ s.t. } \mathrm{tr}(\mathbf{\Sigma})\text{:constant} \tag{3.28}$$

By solving the optimization problems in (3.24) using Karush-Kuhn-Tucker optimization conditions [56], power loading matrix $\mathbf{\Sigma}$ can be found as:

$$\sigma_i = \left[\mu - (\gamma \mathrm{tr}(\mathbf{D}) p_{T(i \bmod L+P)} \delta_{(i \bmod M)})^{-1}\right]^+ \quad i = 1,2,\ldots,M(P+L) \tag{3.29}$$

where $[x]^+$ =max[0,x] for a scalar x, δ_i's are the eigenvalues of transmit correlation matrix $\mathbf{R}_T$, μ is the constant determined by the power constraint, and σ_i and p_{Ti}'s are the diagonal entries of $\mathbf{\Sigma}$ and $\mathbf{P}_T$, respectively.

The capacity-approaching precoding design for *uncorrelated* channel paths can be summarized as follows:

i) Eigen-decomposition of $M \times M$ transmit correlation matrix and calculation of entries of $\mathbf{P}_T$.

ii) Computation of the power coefficients by solving $M(P+L)$ power constraint equations (3.29).

iii) Construction of the precoding matrix based on (3.27).

Figure 3.2 illustrates the resulting precoding structure that consists of $P+L$ *parallel* precoders. In this case, each $M\times 1$ symbol vectors $\mathbf{x}[k(P+L)+i]$, i=1,2,...,($P+L$-1), can be precoded separately using precoder $\mathbf{W}_i$. The resulting vectors are then can be stacked and transmitted through the channel, respectively.

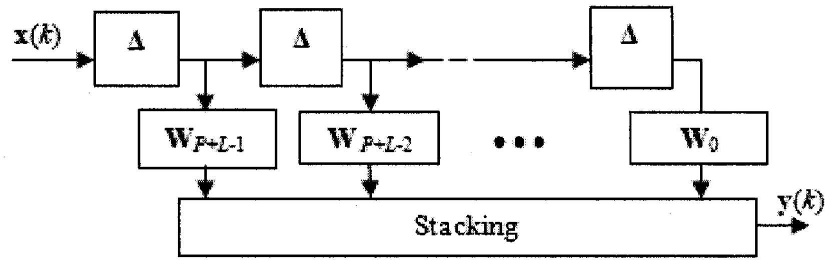


Figure 3.1: Precoder structure for uncorrelated channel paths

We now consider the case of *uncorrelated* channel paths with *different* spatial correlation matrices (described in subsection 3.1.3). The following Lemma specifies the structure of the precoder in this case.

***Lemma* 3.3:** The $M(P+L)\times M(P+L)$ linear precoding matrix, $\mathbf{W}_T$, that maximizes the ergodic capacity in (3.18) for the channel model in (3.10) is a block diagonal matrix $\mathbf{W}_T = \text{diag}(\mathbf{W}_i)$, with $(P+L)$ optimal $M\times M$-matrices $\mathbf{W}_i = \mathbf{\Phi}_i\mathbf{\Sigma}_i^{1/2}\mathbf{\Gamma}_i$, where $\mathbf{\Gamma}_i$'s are $M\times M$ arbitrary unitary matrices, $\mathbf{\Sigma}_i$'s are diagonal matrices, and $\mathbf{\Phi}_i$'s are the $M\times M$ unitary matrices resulting from eigen-decomposition of transmit correlation matrices $\mathbf{R}_{T,l}$'s, l=0,1,...,(L-1).

Proof: Following the same steps as in the previous case, we can derive the ergodic capacity similar to (3.26), i.e.,

$$C \le C_{UB} = \frac{1}{P}\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma\bar{\mathbf{R}}_T\mathbf{\Psi}\mathbf{\Sigma}\mathbf{\Psi}^H)) \tag{3.30}$$

where $\bar{\mathbf{R}}_T = \sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{P}.\text{diag}(\text{tr}(\mathbf{D}_i))\otimes\mathbf{I}_M)\mathbf{R}_T\mathbf{E}^l$. Also, $\mathbf{D}_i$ i=0,1,...,(L-1) come from eigen-decomposition of $\mathbf{R}_{R_i}$ as $\mathbf{R}_{R_i} = \mathbf{V}_i^H\mathbf{D}_i\mathbf{V}$ and $\mathbf{R}_T$ is defined in (3.11).

Considering that $\bar{\mathbf{R}}_T = \text{diag}(\bar{\mathbf{\Phi}}_i\bar{\mathbf{\Lambda}}_i\bar{\mathbf{\Phi}}_i^H)$, $i = 0,1,\ldots,(P+L-1)$, and using the Hadamard's inequality, one can find $\mathbf{\Psi} = \text{diag}(\bar{\mathbf{\Phi}}_i)$, $i = 0,1,\ldots,(P+L-1)$. Therefore, the precoding matrix can be written as:

$$\mathbf{W}_T = \text{diag}(\bar{\mathbf{\Phi}}_i)\mathbf{\Sigma}^{1/2}\mathbf{\Gamma} = \text{diag}(\bar{\mathbf{\Phi}}_i\mathbf{\Sigma}_i^{1/2}\mathbf{\Gamma}_i), \quad i = 0,1,\ldots,(P+L-1) \tag{3.31}$$

where $\mathbf{\Gamma}_i$ is an arbitrary unitary matrix that can be set to identity for simplicity. Therefore, the transmit precoding matrix $\mathbf{W}_T$ is also a block diagonal matrix with $(P+L)$ optimal $M\times M$-matrices $\mathbf{W}_i = \bar{\mathbf{\Phi}}_i \mathbf{\Sigma}_i^{1/2} \mathbf{\Gamma}_i$ where $\bar{\mathbf{\Phi}}_i$ is one diagonal block of the eigenvector matrix of $\bar{\mathbf{R}}_T = \sum_{l=0}^{P} (\mathbf{E}^l)^T (\mathbf{P}.\text{diag}(\text{tr}(\mathbf{D}_i)) \otimes \mathbf{I}_M) \mathbf{R}_T \mathbf{E}^l$.

■

Furthermore, the diagonal entries of $\mathbf{\Sigma}$ can be obtained from an optimization problem by setting $\mathbf{\Psi} = \text{diag}(\bar{\mathbf{\Phi}}_i)$ in (3.30):

$$\max_{\mathbf{\Sigma}} \log_2 \det(\mathbf{I}_{M(P+L)} + \gamma\, \text{diag}(\bar{\mathbf{\Lambda}}_i)\mathbf{\Sigma}) \text{ s.t. } \text{tr}(\mathbf{\Sigma}):\text{constant} \tag{3.32}$$

Solving (3.32) results in the power loading policy of transmit precoding matrix as follows:

$$\sigma_i = \left[\mu - (\gamma\, \bar{\delta}_{(i \bmod M)})^{-1}\right]^+, \; i = 1,2,\ldots,M(P+L) \tag{3.33}$$

where σ_i, $\bar{\delta}_i$'s and p_{Ti}'s are the diagonal entries of $\mathbf{\Sigma}$, $\bar{\mathbf{\Lambda}}_i$ and $\mathbf{P}_T$, respectively, and μ is the constant determined by the power constraint.

In other words, when the spatial correlation matrices are not the same for different channel paths, the structure of the precoder is also block-diagonal and therefore it can be decoupled into $(P+L)$ *frequency-flat* $M\times M$ precoders, similar to Figure 3.2. However, the construction of the $(P+L)$ precoders requires to solve the eigen-decomposition of an $M(P+L)\times M(P+L)$ matrix $\bar{\mathbf{R}}_T$ (or equivalently, L different transmit correlation matrices of size $M\times M$), and hence, it is more complicated than the small $(M\times M)$ eigen-decomposition for transmit correlation matrix $\mathbf{R}_T$ in the previous case of *similar* spatial correlation matrices.

3.2.2. Correlated Channel Paths

Similar to the case of *uncorrelated* paths, getting the expectation of general capacity equation in (3.14) is difficult. Instead, substituting (3.8) into (3.18) and applying Jensen's inequality, we can obtain an upper bound on the capacity,

$$C \le C_{UP} = \frac{1}{P}\log_2 \det E\{\mathbf{I}_{M(P+L)} + \gamma \mathbf{W}_T^H \bar{\mathbf{E}}^T (\mathbf{I}_{P+1} \otimes (\mathbf{R}_P^{1/2} \otimes \mathbf{R}_T^{H/2})\mathbf{G}^H \mathbf{R}_R^{H/2}) (\mathbf{I}_{P+1} \otimes \mathbf{R}_R^{1/2}\mathbf{G}(\mathbf{R}_P^{1/2} \otimes \mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T\} \tag{3.34}$$

Our objective is to find the precoder matrix $\mathbf{W}_T$ in (3.34), which maximizes the upper bound on ergodic capacity under transmitted power constraint, based on the knowledge of only spatial (transmit and receive) and path correlation matrices at the transmitter. *Theorem* 3.1 provides the structure of the precoder in this scenario.

***Theorem* 3.1:** Let $\mathbf{W}_T$ be The $M(P+L)\times M(P+L)$ linear precoding matrix that maximizes the upper bound on ergodic capacity in (34) has the form $\mathbf{W}_T = \mathbf{\Phi}\mathbf{\Sigma}^{1/2}\mathbf{\Gamma}$, where $\mathbf{\Phi}$ is a unitary matrix that can be calculated from Kronecker product of the eigenvector matrices of transmit, $\mathbf{R}_T$ and $\sum_{l=0}^{P}(\mathbf{E}_*^l)^T\mathbf{R}_P\mathbf{E}_*^l$ where $\mathbf{R}_P$ is the path correlation matrix, $\mathbf{\Gamma}$ is an arbitrary unitary matrix, $\mathbf{E}_*^l$ is a $(P+L)\times(P+L)$ column-shifting matrix, $\mathbf{E}_*^l = \begin{bmatrix} \mathbf{0}_{l\times(P+L-l)} & \mathbf{I}_{l\times l} \\ \mathbf{I}_{(P+L-l)\times(P+L-l)} & \mathbf{0}_{(P+L-l)\times l}\end{bmatrix}$, and $\mathbf{\Sigma}$ is a diagonal matrix with the $(Mi+j)^{\text{th}}$ diagonal element found from $\sigma_{Mi+j} = \left[\mu - (\gamma \text{tr}(\mathbf{D}) r_{Ti} r_{Pj})^{-1}\right]^+$ $i = 1,2,...,M, j = 1,2,...,(P+L)$, μ is a constant determined by transmit power constraint and r_{Ti}' s and r_{Pj}' s are the eigenvalues of transmit correlation matrix and $\sum_{l=0}^{P}(\mathbf{E}_*^l)^T\mathbf{R}_P\mathbf{E}_*^l$, respectively.

Proof: Using the same procedure as in the previous case and applying the eigen-decomposition of receive correlation matrix, one can derive an equation similar to (3.23) but involving the path correlation matrix, $\mathbf{R}_P$, as follows

$$C \le C_{UB} = \frac{1}{P}\log_2 \det E\{\mathbf{I}_{M(P+L)} + \gamma \text{tr}(\mathbf{D})\mathbf{W}_T^H \bar{\mathbf{E}}^T (\mathbf{I}_{P+1} \otimes (\mathbf{R}_P^{1/2} \otimes \mathbf{R}_T^{H/2})) \\ (\mathbf{I}_{P+1} \otimes (\mathbf{R}_P^{H/2} \otimes \mathbf{R}_T^{1/2}))\bar{\mathbf{E}}\mathbf{W}_T\} \tag{3.35}$$

As mentioned in Section 3.1, in this case, we assume that the entries of $\mathbf{G}_l$'s are i.i.d. zero-mean complex Gaussian with ½ variance per dimension and the power of paths are taken into account in the diagonal entries of path correlation matrix.

The second term in the expectation in (3.35) can be rewritten as a summation of P terms, i.e.,

$$C \le C_{UB} = \frac{1}{P}\log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \text{tr}(\mathbf{D})\left[\sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{R}_P \otimes \mathbf{R}_T)\mathbf{E}^l\right]\mathbf{W}_T\mathbf{W}_T^H) \tag{3.36}$$

It can be seen that $\sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{R}_P \otimes \mathbf{R}_T)\mathbf{E}^l = \left[\sum_{l=0}^{P}(\mathbf{E}_*^l)^T \mathbf{R}_P \mathbf{E}_*^l\right] \otimes \mathbf{R}_T = \bar{\mathbf{R}}_P \otimes \mathbf{R}_T$. Considering the eigen-decompositions of $\bar{\mathbf{R}}_P = (\mathbf{E}_*^l)^T \mathbf{R}_P \mathbf{E}_*^l = \mathbf{\Phi}_P \mathbf{\Delta}_P \mathbf{\Phi}_P^H$ and $\mathbf{R}_T = \mathbf{\Phi}_T \mathbf{\Delta}_T \mathbf{\Phi}_T^H$ and the fact that the eigenvector matrix of the Kronecker product of two matrices is the Kronecker product of their eigenvector matrices, the eigen-decomposition of $\sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{R}_P \otimes \mathbf{R}_T)\mathbf{E}^l$ can be written as:

$$\sum_{l=0}^{P}(\mathbf{E}^l)^T(\mathbf{R}_P \otimes \mathbf{R}_T)\mathbf{E}^l = (\mathbf{\Phi}_P \otimes \mathbf{\Phi}_T)(\mathbf{\Delta}_P \otimes \mathbf{\Delta}_T)(\mathbf{\Phi}_P^H \otimes \mathbf{\Phi}_T^H)$$

Applying Hadamard's inequality to (3.36) and using the eigen-decomposition $\mathbf{W}_T\mathbf{W}_T^H = \mathbf{\Psi\Sigma\Psi}^H$ will give us an equation for the eigenvector matrix $\mathbf{\Psi}$:

$$\mathbf{\Psi} = \mathbf{\Phi}_P \otimes \mathbf{\Phi}_T \tag{3.37}$$

Therefore, the optimal $\mathbf{W}_T$ structure can be derived as

$$\mathbf{W}_T = (\mathbf{\Phi}_P \otimes \mathbf{\Phi}_T)\mathbf{\Sigma}^{1/2}\mathbf{\Gamma} \tag{3.38}$$

where $\mathbf{\Gamma}$ is an arbitrary $M(P+L)\times M(P+L)$ unitary matrix.

The selected linear precoding matrix $\mathbf{W}_T$ is an eigen-beamformer with orthogonal beams pointing to a matrix that is a function of $\bar{\mathbf{R}}_P \otimes \mathbf{R}_T$. The diagonal matrix $\mathbf{\Sigma}$ can be found via a power loading on each of the eigen-beamformers. Finding the eigen-decomposition of $\bar{\mathbf{R}}_P \otimes \mathbf{R}_T$ is sufficient for calculation of optimal precoding matrix with a complexity close to the case of unequal spatial correlation matrices.

The power loading policy across the eigen-beams can be obtained by substituting (3.37) in (3.36) as:

$$\max_{\mathbf{\Sigma}} \log_2 \det(\mathbf{I}_{M(P+L)} + \gamma \mathrm{tr}(\mathbf{D})\mathbf{\Delta\Sigma}) \text{ s.t. } \mathrm{tr}(\mathbf{\Sigma})\text{:constant} \tag{3.39}$$

The elements of $\mathbf{\Delta}$ are the products of the eigenvalues of $\mathbf{R}_T$ and $\mathbf{R}_P$, and the solution to this optimization problem becomes a water-pouring policy [15] with power loading:

$$\sigma_{Mi+j} = \left[\mu - (\gamma \mathrm{tr}(\mathbf{D})\delta_{Mi+j})^{-1}\right]^+ = \left[\mu - (\gamma \mathrm{tr}(\mathbf{D}) r_{Ti} r_{Pj})^{-1}\right]^+ \quad i = 1..M, j = 1..P+L \tag{3.40}$$

where σ_{Mi+j} and δ_{Mi+j} are the $(Mi+j)^{\text{th}}$ diagonal elements of $\mathbf{\Sigma}$ and $\mathbf{\Delta}$, respectively, r_{Ti}' s and r_{Pj}' s are the eigenvalues of $\mathbf{R}_T$ and $\bar{\mathbf{R}}_P$, and μ is a constant determined by the power constraint.

■

Therefore, based on the results of *Theorem* 3.1, assuming all matrices to be full-rank, the optimal precoding design is reduced to finding $M(P+L)$ eigenvectors of the matrix $\bar{\mathbf{R}}_P \otimes \mathbf{R}_T$ and $M(P+L)$ power coefficients on each of the eigenmodes and can be summarized in four following steps:

i) Eigen-decomposition of $M\times M$ transmit correlation, $\mathbf{R}_T$ and $(P+L)\times(P+L)$ $\bar{\mathbf{R}}_P$ matrices.
ii) Computation of the eigenvector matrix of precoder matrix, $\mathbf{\Psi}$, based on the summation in (3.37).
iii) Calculation of the power coefficients by solving $M(P+L)$ power constraint equations (3.40).
iv) Construction of the precoder matrix based on (3.38).

It can be shown that, in the special case of uncorrelated equal-power channels paths, the first part of *Theorem* 3.1 will be reduced to *Lemma* 3.2 and the precoding structure can be decomposed into $P+L$ similar $M\times M$ precoders, i.e., the case of uncorrelated channel paths with similar spatial correlation matrices is a special case.

3.3. Numerical Results

We will first investigate the effect of correlation on the capacity of MIMO systems with different numbers of transmit antennas and effective channel paths. Next, the achievable ergodic capacity of the *spatial* and *proposed* precoders over frequency-selective fading channels with different numbers of antennas and channel paths is evaluated.

Although actual characteristics of correlation coefficients depend on various factors (e.g., channel power angular spectrum (PAS), angular spread and antenna spacing, propagation environment), as an illustrative example and for simplicity in simulation, we assume that the correlation is linearly decreasing with antenna distance, e.g., for a 4×4 MIMO channel using ULA, we consider the following transmit correlation matrix:

$$\mathbf{R}_T = \begin{bmatrix} 1 & \rho & 0.5\rho & 0.25\rho \\ \rho & 1 & \rho & 0.5\rho \\ 0.5\rho & \rho & 1 & \rho \\ 0.25\rho & 0.5\rho & \rho & 1 \end{bmatrix}$$

where ρ is selected to be 0.5 and 1 for partial and full correlation, respectively.

Figure 3.3 illustrates the effects of spatial and path correlations on the capacity of a MIMO channel with equal-power paths, N=2, M=2, 4 and L=2, 4 and P=6 when precoding is not used. In general, the ergodic channel capacity is increased with L, but is greatly reduced as correlation is increased. Furthermore, increasing the correlation also reduces the capacity increase due to increased L. For example, in the case of full correlation, the three capacity-versus-SNR curves for different values of L are almost the same. This can be explained by the fact that full correlation reduces the channel matrix rank and hence, the number of independent eigen-modes of the channel. This is equivalent to reduction in the number of effective parallel channels and hence in the channel capacity. Figure 3.3 also indicates that the increase in the number of transmit antennas, M, does not have any major effect on the capacity because it is known that the capacity in a MIMO system is proportional to the minimum number of transmit and receive antennas, which is N=2 for all capacity-versus-SNR curves.

Figure 3.4 shows the capacity-versus-SNR curves for three systems: (i) without precoding, (ii) with *spatial* precoder, and (iii) with the *proposed* precoder, when no path correlation exists in the channel but the transmit antennas are fully correlated. Without path correlation, as previously discussed in Section 3.1, the proposed precoder structure can be decomposed into L similar *spatial* precoders. The results confirm that the *proposed* and *spatial* precoders have the same performance. Also, the capacity gain provided by the precoders is increased with the dimension of the uncorrelated channel, M and L.

Figure 3.5 shows the ergodic capacity of a partially correlated channel for two different values of M and L (2 and 4). The simulation results indicate that the *proposed* precoder outperforms the *spatial* precoder and achieves better capacity with increased numbers of channel paths and transmit antennas. Furthermore, the gain in achievable capacity of the proposed precoder is increased with larger values of L and M due to the mitigation of path correlation by the proposed precoder, e.g., as compared to the *spatial* precoder, at SNR=10dB, the *proposed* precoder provides a capacity gain of about 10% for M=L=2, and 25% for M=L=4.

Figure 3.6 shows the capacity for the cases of full spatial and path correlation. In general, compared to the results in Figure 3.5, the achievable channel capacity is reduced due to increased correlation at high SNR. However, the *proposed* precoder still provides large gain in the achievable channel capacity as compared to the spatial precoder and system without precoding. As shown in Figure 3.5 and Figure 3.6, this gain is larger with increased values of L and M and increased correlation.

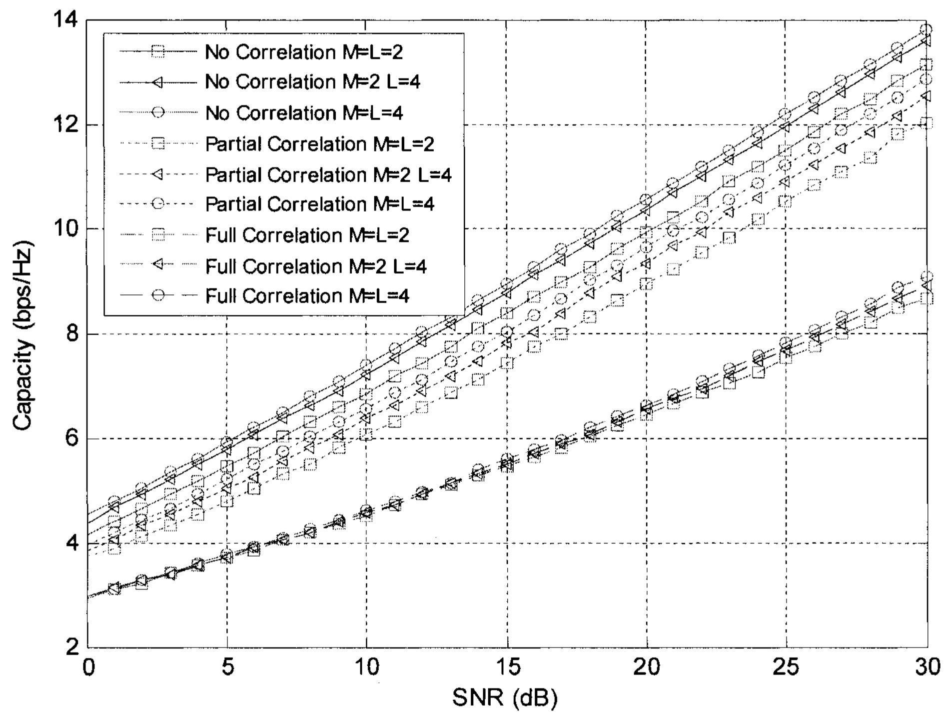


Figure 3.3: Ergodic capacity with different correlation levels and numbers of transmit antennas and channel paths

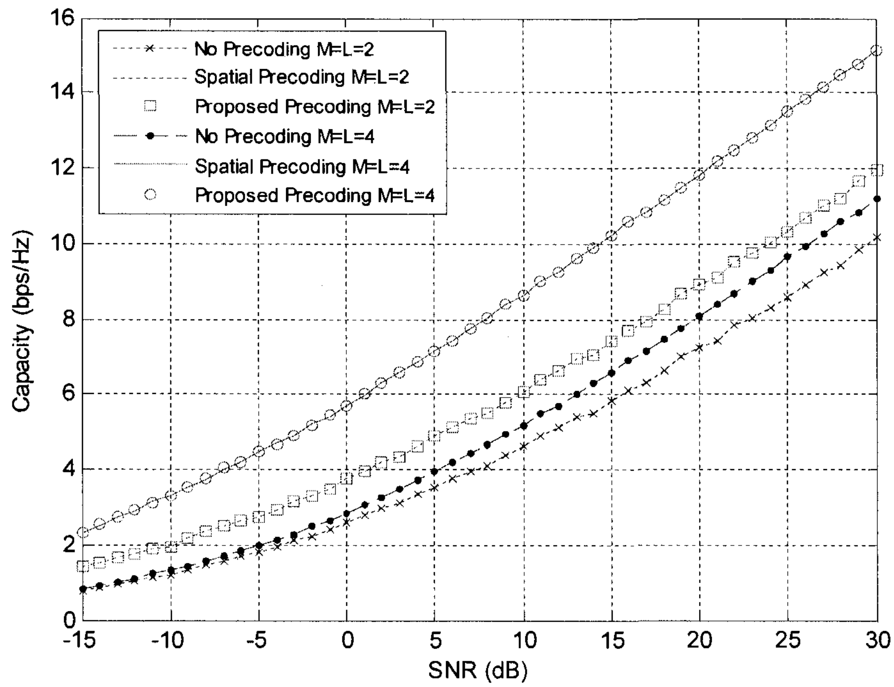


Figure 3.4: Ergodic capacity with *fully correlated* transmit antennas but *no* path correlation

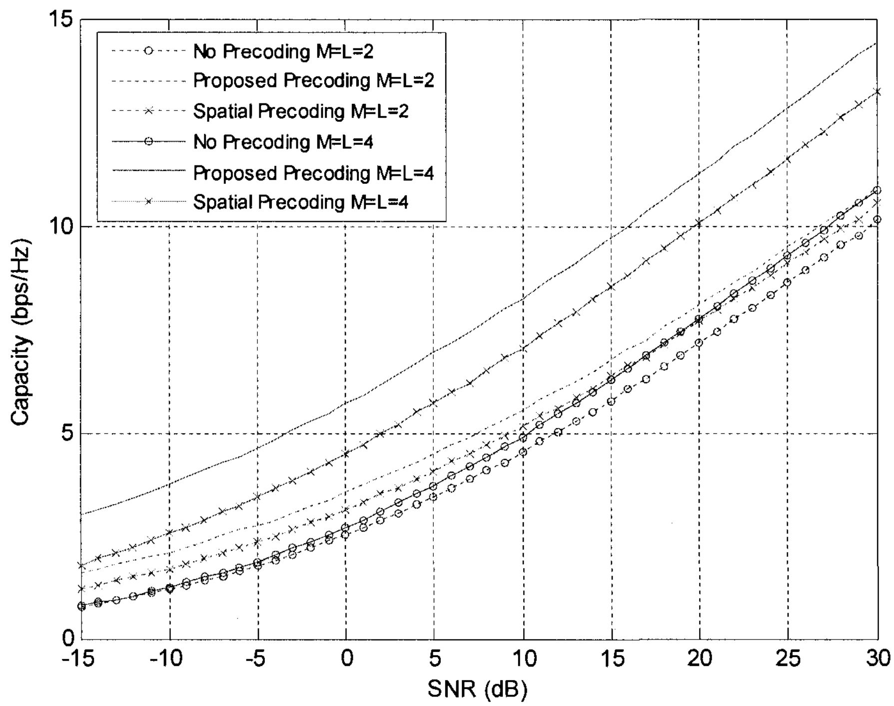


Figure 3.5: Ergodic capacity with *partially correlated* transmit antennas and channel paths

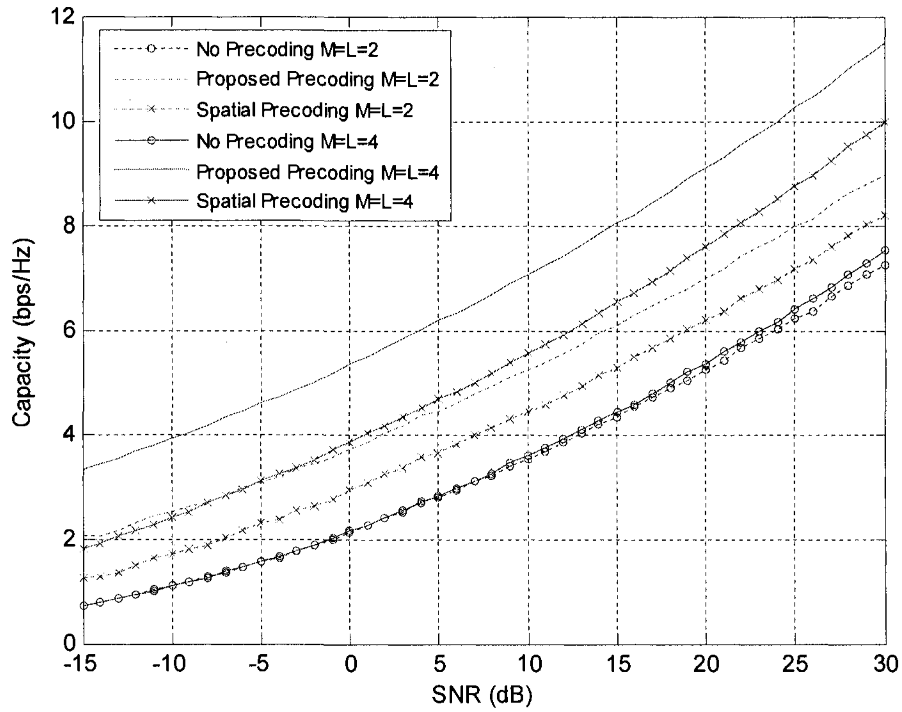


Figure 3.6: Ergodic capacity with *fully correlated* transmit antennas and channel paths

3.4. Chapter Summary

In this chapter, we proposed a precoder structure for MIMO systems in a frequency-selective fading environment, based on the knowledge of transmit and receive antenna and channel path correlations. Optimum precoding designs to maximize the ergodic capacity under constraint on transmitted power were developed for three different cases: uncorrelated channel paths with similar spatial correlation, uncorrelated channel paths with different spatial correlation, and correlated channel paths. The precoder structures in the cases of uncorrelated channel paths are composed of $P+L$ parallel precoders for frequency-flat fading channels. The power assignment to each precoder and the power allocation over the eigen-modes of each precoder was calculated based on the power of channel paths and eigenvalues of transmit correlation matrix. In the case of correlated channel paths, the precoder structure is in fact an eigen-beamformer with the beams refer to the eigenvectors of the Kronecker product of path and transmit correlation matrices. Furthermore, the power allocated to each eigen-mode can be obtained from a water-pouring policy which is specified by the product of eigenvalues of transmit antenna and path correlation matrices.

Simulation results for MIMO systems in a frequency-selective fading environment with different scenarios indicate that the proposed precoder can increase the system ergodic capacity in presence of spatial and path correlations and its offered capacity gain is increased with the level of correlation and numbers of antennas and channel paths. The effectiveness of the proposed precoder is more pronounced in the environment with severe spatial and path correlation.

Chapter 4

User Selection and Precoding in Point-to-Multipoint Systems[1]

4.1. Multi-user System Model

In this chapter, we investigate downlink precoding schemes that can achieve the capacity of a MIMO broadcast channel in which a multiple-antenna transmitter communicates with a number of mobile units, based on the assumption that the transmitter just has a partial knowledge of users' channels.

Our study shows that a careful selection of channel side information is very important in the sense that it can reduce the feedback cost and transmit complexity while still provides a considerable performance. We propose a zero-forcing transmission scheme that uses only partial channel information with low feedback load and also facilitates the algorithm of selecting the best users at the transmitter. The proposed scheme achieves the same ergodic sum capacity growth rate as that of dirty paper coding (DPC) [60] with reduced feedback load and algorithm complexity.

Furthermore, by distributing the processing loads among users communicating in the network, the proposed scheme greatly relaxes the feedback load needed to calculate the optimal precoding vectors, and the complexity of selecting the best user algorithm at the transmitter. This reduction in feedback overhead and complexity as compared to other

[1] This chapter has been partially presented in [92] and [93].

schemes is increased with the number of users and/or transmits antennas and hence, makes the proposed scheme a desired candidate for distributed network management systems.

Four different strategies for user selection, power allocation and precoding are proposed. Each of these strategies is suitable for a specific propagation scenario and channel condition. However, as the structure and algorithms of all of these schemes are identical with minor difference, it is possible to implement all of them at the base station and switch amongst them when necessary. This gives a degree of robustness to the system that can cope with channel impairments and changes.

We consider a broadcast system using a transmitter with M transmit antennas to serve n users, each with N_i antennas (i = 1, …, n). The channel matrix for user i can be represented by an $N_i \times M$ matrix $\mathbf{H}_i$ whose entries are assumed to be Gaussian with zero mean and unit variance. At time instant k, the $M \times 1$ transmit vector $\mathbf{s}$ to different users has a constrained power, i.e., $E\{\mathbf{ss}^H\} \leq P$ and the $N_i \times 1$ received signal vector for user i is $\mathbf{y}_i$, we can then write:

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{s} + \mathbf{n}_i \qquad i = 1, 2, ..., n \tag{4.1}$$

where $\mathbf{n}_i$ is the $N_i \times 1$ additive white Gaussian noise sample vector for user i. Using precoding, the transmitted vector $\mathbf{s}$ is a linear combination of transmitted symbols for each user $\mathbf{s}_i$,

$$\mathbf{s} = \sum_{i=1}^{n} \sqrt{P_i} \mathbf{W}_i \mathbf{s}_i, \quad \sum_{i=1}^{n} P_i = P$$

where $\mathbf{W}_i$, $\mathbf{s}_i$ and P_i are the $M \times N_i$ precoding matrix, transmitted symbols, and allocated power for user i, respectively. For simplicity, in the following discussions, we assume N_i =1, i = 1, …, n. The received signal for the i^{th} user can be represented as

$$y_i = \sqrt{P_i} \mathbf{h}_i \mathbf{w}_i s_i + \sum_{j \neq i} \sqrt{P_j} \mathbf{h}_i \mathbf{w}_j s_j + n_i \qquad i = 1, 2, ..., n \tag{4.2}$$

where the second term is due to the interference from other users. The bold-faced, lower-case symbols in the equation denote vectors corresponding to the case of N_i=1. Our objective is to find the set of precoding vectors to maximize the achievable sum rate, i.e.,

$$R = \max_{\mathbf{w}_i, \sum \|\mathbf{w}_i\|^2 \le P} \sum_{i=1}^{n} \log(1 + \frac{P_i |\mathbf{h}_i \mathbf{w}_i|^2}{1 + \sum_{j=1 \ne i}^{n} P_j |\mathbf{h}_i \mathbf{w}_j|^2}) \tag{4.3}$$

4.2. BC Capacity and Zero-Forcing Precoding

The sum rate capacity of a BC channel achieved by DPC approach has been shown to be [59], [61]:

$$C_{sum} = R_{DPC} = \max_{\sigma_i, \sum_{i=1}^{n} \sigma_i \le P} \log\det(\mathbf{I} + \sum_{i=1}^{n} \sigma_i \mathbf{h}_i^H \mathbf{h}_i) \tag{4.4}$$

which is in fact the capacity of the dual MAC channel. Note that σ_i's are not the same as P_i's in (4.3). In [63], it has been shown that when the number of users (n) is large, the ergodic capacity asymptotically scales like $M \log\log n$. In other words, the ergodic capacity can be written as:

$$\lim_{n \to \infty} E\{C_{sum}\} = E\{R_{DPC}\} = M \log(1 + \frac{P}{M} \log n) \tag{4.5}$$

i.e., in the limit, it can provide a diversity order of log n compare to a single user system. Also, a linear increase in the capacity is achieved by increasing the number of transmit antenna (M). It can be seen from (4.5) that a scheme allocating an average power of $(P/M)\log n$ to each of M best users via M independent paths (subchannels), can provide the same ergodic capacity as the DPC and hence, is asymptotically optimal.

On the other hand, although sub-optimum, zero-forcing scheme can provide M parallel (independent) sub-channels from transmitter to M users by selecting the precoding vectors, $\mathbf{w}_i$'s, such that $\mathbf{h}_i \mathbf{w}_j = 0$ $(i \ne j)$ (i.e., no interference from other users). Since the size of $\mathbf{h}_i$'s is $1 \times M$, there will be at most M precoding vectors that can satisfy the above equations. Therefore, at most M users should be selected among n available users. Let $S \subset \{1,\ldots,n\}$, $|S| \le M$ be a subset of user selected for transmission. By suitable selection of precoding matrices[2], there will be M independent sub-channels and hence the achievable rate can be written as:

[2] The easiest selection of $\mathbf{w}_i$ is such that $\mathbf{W}(S)$ is the pseudo-inverse of $\mathbf{H}(S)$ where $\mathbf{W}(S)$ is a matrix that is composed of the precoding vectors in S and $\mathbf{H}(S)$ is the channel matrix whose columns are composed of channel vectors of the users in S, i.e. $\mathbf{W}(S) = \mathbf{H}^{\dagger}(S)$

$$R_{ZF}(S) = \max_{P_i, \sum_{i \in S} P_i \le P} \sum_{i \in S} \log(1 + \lambda_i P_i) \tag{4.6}$$

where λ_i is the i^{th} sub-channel gain. Furthermore, P_i's can be found via a water-pouring process as $P_i = [v - \lambda_i^{-1}]^+$ where $[x]^+$ =max[0,x] for a scalar x and the constant v is such that the power constraint P is satisfied.

Note that there are different possible selections of user subsets. Achievable sum rate of optimum zero-forcing is defined as [66]:

$$R_{ZF} = \max_{S} R_{ZF}(S) \tag{4.7}$$

where $S \subset \{1, ..., n\}$ $|S| \le M$. The optimal solution may need a lengthy and complicated exhaustive search of all possible user subsets for zero-forcing schemes. Furthermore, required knowledge of instantaneous channel vectors (full channel information) at the transmitter may need a large amount of channel feedback load. This limits the application of optimal zero-forcing method to simple cases with small number of users. In [66], it has been shown indirectly that in the limit of large number of users (n), the zero-forcing beamforming scheme can provide a sum rate equal to that of DPC in (4.5), i.e.,

$$E\{R_{ZF}\} = M \log(1 + \frac{P}{M} \log n) = E\{R_{DPC}\} \tag{4.8}$$

An appealing sub-optimum zero-forcing scheme with simple selection of M best users was discussed in [66]. However, it still needs full channel knowledge at the transmitter with potentially high complexity and feedback load in real wireless applications.

By increasing the number of users in the network, it is more likely that there exists a subset of M users such that a linear precoding scheme (like zero-forcing) can achieve the sum rate capacity of the system. In other words, for a large user population, different sub-optimum zero-forcing algorithms have similar asymptotic performance, and hence, complexity becomes the key selection factor. This chapter mainly focuses on how to reduce the processing and feedback load in selecting the best M users with as much as possible orthogonal channels to maintain the asymptotic gain of log log n over ergodic capacity in (4.8).

4.3. Partial Knowledge Zero-Forcing Scheme

Assume that transmitter and all users have a predetermined known orthonormal basis[3] $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_M)$ of size M (e.g., standard basis for C^M). We will discuss the selection of this orthonormal basis in Section 4.5.

Now, we consider two different strategies: *power* user selection (PUS) and *normalized* user selection (NUS). In the first strategy, at the start of each transmission period, each user calculates the projections of its channel vector on each of the vectors in orthonormal basis ($\mathbf{u}_i$), i.e., for j^{th} user, the projection on i^{th} vector is:

$$\gamma_{ij} = \left\langle \mathbf{h}_j, \mathbf{u}_i \right\rangle^2 = \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 \quad ; \ \forall i:1..M \ \ j:1..n \tag{4.9}$$

While in the second strategy, each user calculates the norm of the projections of its channel vector on each of the vectors in orthonormal basis, i. e.,

$$\gamma_{ij} = \left| \mathbf{h}_j \right|^{-2} \left\langle \mathbf{h}_j, \mathbf{u}_i \right\rangle^2 = \left| \mathbf{h}_j \right|^{-2} \left| \mathbf{h}_j \mathbf{u}_i^H \right|^2 \quad ; \ \forall i:1..M \ \ j:1..n \tag{4.10}$$

Next, in both strategies, each user then sends its maximum γ_{ij} along with its index to the transmitter, i.e.,

$$\gamma_j = \max_i \gamma_{ij} \, ; \, \alpha_j = \arg\max_i \gamma_{ij} \tag{4.11}$$

Transmitter then easily selects the best user over each of the orthonormal basis by finding the maximum γ_j over those users. Assume that the indices of users for which $\alpha_j = i$ are saved in a set $S(i)$. Therefore,

$$\gamma_{i,\max} = \max_{j \in S(i)} \gamma_j \quad \forall i:1...M \tag{4.12}$$

The transmitter selects these M users as the winner users and asks them to send back their channel vectors (in total M vectors of size $M\times 1$). It then selects the precoding vectors. For selection of precoding vectors we also consider two different precoding schemes: *opportunistic* precoding (OP) and *channel-aware* precoding (CAP). In opportunistic precoding, as its name implies, the transmitter does not know the channel vector of selected users and therefore transmit in the direction of orthonormal basis, i.e.,

$$\mathbf{w}_i = \mathbf{u}_i \quad \forall i:1...M \tag{4.13}$$

---

[3] A set $\mathbf{U}$ is orthonormal if $\forall i,j \ \ u_i^H u_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$

where $\mathbf{w}_i$'s are the M precoding vectors associated to each of the selected users. Sending users' signals at the direction of orthonormal basis clearly produces zero interference. Therefore, opportunistic precoding can also be referred as *interference-minimizer* precoding. Maximum sum-rate of the users can then be written as:

$$R_{OP} = \sum_{i=1}^{M}(1+\frac{P}{M}\frac{\left|\mathbf{h}_i\mathbf{u}_i^H\right|^2}{1+\sum_{j=1\neq i}^{M}\left|\mathbf{h}_i\mathbf{u}_j^H\right|^2}) \tag{4.14}$$

In the limit, a large number of the channel vectors of M selected users are in the direction of orthonormal basis with high probability. Therefore, the sum rate can be rewritten as:

$$R_{OP} = \sum_{i=1}^{M}\log(1+\frac{P}{M}\left|\mathbf{h}_i\mathbf{u}_i^H\right|^2) = \sum_{i=1}^{M}\log(1+\frac{P}{M}\left|\mathbf{h}_i\right|^2) \tag{4.15}$$

(4.15) is the direct result of (4.14) and the fact that in opportunistic precoding there is no interference between users. In channel-aware precoding transmitter does know the selected users' channel vectors and selects the precoding vectors as the normalized vector of users' channel vectors:

$$\mathbf{w}_i = \left|\mathbf{h}_i\right|^{-1}\mathbf{h}_i \quad \forall i:1...M \tag{4.16}$$

where the best M users in (4.14) are indexed from 1 to M. In this scheme, we target maximizing received SNR. It is easy to see that by sending at the direction of selected users' channel vectors, we maximize the receive SNR of each user regardless of interference introducing to other users. Therefore, channel-aware precoding can also be called *SNR-maximizer* precoding. Again, the idea behind this scheme is that each channel vector of M selected users is almost in the direction of one of the basis vectors; hence their channel vectors are near-orthogonal to each other. This orthogonality increases with increased number of users in the network. In the best case when the selected channels compose an orthogonal set, the following rate is achievable:

$$R_{CAP} = \max_{P_i}\sum_{i=1}^{M}\log(1+P_i\left|\mathbf{h}_i\right|^2) \text{ s.t. } \sum_i P_i \leq P \tag{4.17}$$

The set of M power loading values P_i's are selected based on waterpouring as $P_i = [\nu - \left|\mathbf{h}_i\right|^{-2}]^+$. Obviously, other users have zero power.

We can therefore, distinguish between four different user selection and precoding strategies: PUS-OP, PUS-CAP, NUS-OP, NUS-CAP. The features of each of these schemes have been summarized in Table 4.1. Regarding fairness, as PUS selects users based on the maximization of projection of their channel vectors over orthonormal bases, it is more likely that users with strong power (i.e., users near base station) are finally selected. Therefore, the issue of fairness is not considered in PUS while it is not the case in NUS scheme. Note that we will elaborate the performance and specifications of each scheme more in Section 4.6 when we present numerical results.

Scheme	channel knowledge of all users at Tx	channel knowledge of selected users at Tx	Power alocation and waterpouring	Fairness	Interference between users
PUS-OP	No	No	No	No	No
PUS-CAP	No	Yes	Yes	No	Yes
NUS-OP	No	No	No	Yes	No
NUS-CAP	No	Yes	Yes	Yes	Yes

Table 4.1: Features of user-selection/precoding strategies

4.4. Asymptotic Performance of Zero-Forcing Scheme

In this section, we investigate the asymptotic performance of NUS user selection schemes. Since the analysis of PUS scheme is also very identical to CAP, for the sake of brevity, we do not discuss it here.

For any two arbitrary vectors **a** and **b**, we define the following orthogonality measure:

$$\varepsilon(\mathbf{a},\mathbf{b}) = 1 - \left|\mathbf{a}^H\mathbf{b}\right|^2 \left|\mathbf{a}\right|^{-2} \left|\mathbf{b}\right|^{-2} \tag{4.18}$$

Clearly, for **a** close to **b**, $\varepsilon(\mathbf{a},\mathbf{b})$ tends to zero. For the best user channel $\mathbf{h}_i$ in the direction of basis vector $\mathbf{u}_i$, we have:

$$\varepsilon_i = \varepsilon(\mathbf{h}_i,\mathbf{u}_i) = \min_j \varepsilon(\mathbf{h}_j,\mathbf{u}_i) = 1 - \max_j \left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 \left|\mathbf{h}_j\right|^{-2} \tag{4.19}$$

***Lemma* 4.1:** For large number of users, the cumulative distribution function (cdf) function of $\varepsilon_i \ \ (i:1..M)$ is linearly increased with n, i.e., $\forall i$: $F_{\varepsilon_i}(x) = \Pr(\varepsilon_i < x) \Box O(nx)$.

Proof: Since $\mathbf{U}$ is an orthonormal set, the random variables $\left|\mathbf{h}_j \mathbf{u}_i^H\right|^2$ are i.i.d. over i and j with $\chi^2(2)$ distribution. As $\left|\mathbf{h}_j\right|^2$ is the sum of square of M Gaussian random variables, it is $\chi^2(2M)$ distributed. It follows that the probability density function (pdf) and cdf of $t_{ij} = \left|\mathbf{h}_j \mathbf{u}_i^H\right|^2 \left|\mathbf{h}_j\right|^{-2}$ are, respectively,

$$f_t(x) = \int_0^\infty y e^{-yx} \cdot \frac{y^{M-1} e^{-y}}{(M-1)!} dy = M(1+x)^{-(M+1)} \tag{4.20}$$

and

$$F_t(x) = 1-(1+x)^{-M} \,. \tag{4.21}$$

Hence, the cdf of ε_i is

$$F_{\varepsilon_i}(x) = P(\varepsilon_i < x) = P(\max_j t_{ij} > 1-x) = 1 - F_t^n(1-x) = 1-\left[1-(2-x)^{-M}\right]^n \tag{4.22}$$

For small x, the right hand side of (4.22) can be approximated as

$$F_{\varepsilon_i}(x) \approx 1-\left[1-2^{-M}(1+2^{1-M}Mx)\right]^n \approx 2^{1-2M} Mnx \Box O(nx) \tag{4.23}$$

for any $\varepsilon_i \ \ (i:1..M)$.

■

Lemma 1 indicates that the probability of ε_i smaller than a specific small value increases linearly with the user population, n, and the best users' channel vectors become more and more orthogonal to the basis directions (vectors). One can also say that by increasing the number of users, it is more probable that there would be M users with each channel vector very close to the direction of a basis. The following Lemma sheds some light on this fact.

***Lemma* 4.2:** For large number of users, the probability that a user has the largest channel vector projection among all users on two or more directions is linearly proportional to $1/n$.

Proof: From Lemma 1, the squared projections, $\left|\mathbf{h}_j \mathbf{u}_i^H\right|^2$, are i.i.d $\chi^2(2)$-distributed variables over i and j. Among them, let Y_n and $X_{M\text{-}1}$ denote the *largest* of the whole population of n users, and the *second largest* of the selected set of M users, respectively. A user has the best channel for two basis vectors if its *second largest* squared projection is larger than the squared projections of all other users for this specific basis vector. The probability of this event is:

$$P(X_{M-1} > Y_n) = 1 - \int_0^{\infty} (1 - F_X(x)) F_X^{M-1}(x) n f_X(x) F_X^{n-1}(x) dx \tag{4.24}$$

where $F_X(x) = 1 - e^{-x}$ and $f_X(x) = e^{-x}$ are cdf and pdf of a $\chi^2(2)$ random variable. Therefore,

$$P(X_{M-1} > Y_n) = 1 - \int_0^{\infty} n(1 - e^{-x})^{n+M-2} e^{-2x} dx \approx \left[1 + \frac{n}{M}\right]^{-1} \tag{4.25}$$

The right hand side of (4.22), behaves like $O(1/n)$ when n is larger than M.

■

The following Lemma guarantees that the proposed NUS-CAP method performs always better than opportunistic zero-forcing precoding (OP) in which M orthonormal vectors are selected randomly at the transmitter and the best users are selected for those random vectors. Transmitter then sends the information to the best M users on those random orthonormal vectors.

***Lemma* 4.3:** The proposed NUS-CAP zero-forcing method with equal power allocation always performs better than NUS-OP zero-forcing precoding.

Proof: Consider $\varepsilon_i \ (i:1..M)$ in (4.19). Assuming equal power allocation, from (4.15) the average sum rate for the defined NUS-OP zero-forcing precoding can be written as:

$$R_{OP} = E\{\sum_{i=1}^{M}(1 + \frac{P}{M}\frac{\left|\mathbf{h}_i\mathbf{u}_i^H\right|^2}{1 + \sum_{j=1 \neq i}^{M}\left|\mathbf{h}_i\mathbf{u}_j^H\right|^2})\} = E\{\sum_{i=1}^{M}(1 + \frac{P}{M}\frac{1-\varepsilon_i}{1+\varepsilon_i})\} \tag{4.26}$$

The last equality results from the fact that $\mathbf{u}_i$'s are orthonormal. Our proposed NUS-CAP scheme, however, is based on the distance of the best users' channel vectors, i.e. $\left|\mathbf{h}_j\mathbf{u}_i^H\right| \ (i:1..M; j:1:M)$ and can be found in terms of $\varepsilon_i \ (i:1..M)$. It is easy to check that on average:

$$\left|\mathbf{h}_i\mathbf{h}_j\right|^2 \leq M^{-1}\varepsilon_i(1-\varepsilon_i) \ \forall i:1..M \tag{4.27}$$

Therefore, considering equal power allocation (no water-pouring), the average sum rate of NUS-CAP scheme is:

$$R_{CAP} = E\{\sum_{i=1}^{M}(1+\frac{P}{M}\frac{\left|\mathbf{h}_i\mathbf{h}_i^H\right|^2}{1+\sum_{j=1\neq i}^{M}\left|\mathbf{h}_j\mathbf{h}_i^H\right|^2})\} = E\{\sum_{i=1}^{M}(1+\frac{P}{M}\frac{1}{1+\varepsilon_i(1-\varepsilon_i)})\} \tag{4.28}$$

From (4.26) and (4.28), as $1-\varepsilon_i/1+\varepsilon_i < [1+\varepsilon_i(1-\varepsilon_i)]^{-1}$, the average sum rate provided by the NUS-CAP scheme is always greater than that of NUS-OP zero-forcing.

■

Note that in the proof of Lemma 3, we considered equal power allocation on basis vectors, while if we apply the water-pouring in (18), NUS-CAP scheme performs much better than random zero-forcing scheme. Now, we prove that our scheme is optimum for large number of users.

***Theorem* 4.1:** In the limit of large number of users, the NUS-CAP partial knowledge zero-forcing precoding approach proposed in Section 4.3 can achieve an average sum rate equal to that of DPC strategy in (4.5), i.e.,

$$E\{R_P\} = M\log(1+\frac{P}{M}\log n) \tag{4.29}$$

Proof: Based on Lemma 3, it is sufficient to show that the discussed NUS-OP zero-forcing precoding scheme can achieve the rate in (4.29). As the proposed NUS-CAP precoding scheme always outperforms NUS-OP zero-forcing scheme, it turns out that it is capable of achieving the rate in (4.29) in the limit of large n.

Based on the results from extreme value theory in [63], $\max_i t_{ij} = \max_i \left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 \left|\mathbf{h}_j\right|^{-2}$ behaves like $\log n + O(\log\log n)$. Therefore, from (4.18) and (4.25), the average sum rate of the NUS-OP zero-forcing precoding for large n

$$\lim_{n\to\infty} R_{rand} = \lim_{n\to\infty} E\{\sum_{i=1}^{M}(1+\frac{P}{M}(\frac{1}{\varepsilon_i}-1)\} = M\log\frac{P}{M}(1+\log n) \tag{4.30}$$

As the NUS-CAP scheme always outperforms NUS-OP zero-forcing precoding, from Lemma 3, it can achieve the sum rate in (4.30).

■

Note that in the proof of *Theorem* 4.1, we assumed that the available power is distributed equally among M precoder vectors (sometimes called directions). In practice, as mentioned in Section 4.4, we apply a standard water-pouring to achieve to optimum power breakdown. However, as the proof shows, using equal powers on all precoder vectors is also asymptotically optimal (when the number of users is large).

4.5. Feedback Load, Complexity and Design Issues

In this section, we briefly discuss the practical considerations and design issues for our proposed schemes. We show that the proposed schemes can be applied with very low amount of feedback load and transmitter complexity compared to that of full channel knowledge schemes. At the end of this section, we briefly point out the selection of basis vectors set $\mathbf{U}$.

4.5.1. Feedback Load

Consider a BC system using a transmitter equipped with M antennas to serve n single-antenna active users. For schemes based on full channel knowledge such as DPC, optimal zero-forcing described in Section 4.3, and the zero-forcing scheme proposed in [66], at any transmission period, $2Mn$ real values should be fed back to the transmitter. On the other hand, the proposed CAP scheme based on partial knowledge scheme initially needs only n integer indices of the best basis vectors and n real best projections of the users' channel vectors and only $2M^2$ for the best M users channel vectors after initialization. Therefore, in total, it needs $n+2M^2$ real and n integer values. Clearly, the proposed scheme requires much lower feedback load, especially at large number of users.

It is also possible to exclude bad users (those are not near-orthogonal to any of the basis vectors) from channel information feedback. Similar to the idea in [63], we can define a threshold value and if a user has its maximum channel vector projection below this threshold, it is considered as a *bad* user.

***Lemma* 4.4:** At the limit of large number of users, for each of the basis vectors, there exists at least one user j with its channel vector $\mathbf{h}_j$ satisfying $\max_{i=1..M} \left|\mathbf{h}_j \mathbf{u}_i^H\right|^2 > \log n$.

Proof: Based on Lemma 2, for large number of users, n, it is not likely that a particular user can be the best one for two or more basis vectors. Hence, maximization over a set of users that have the maximum projection on a specific basis vector is equivalent to maximization over all users for that particular basis vector, i.e.,

$$\max_{j=1..n}\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 = \max_{j\in S(i)}\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 \;\; \forall i = 1..M . \tag{4.31}$$

As $\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2$ are $\chi^2(2)$ distributed random variables, by applying the extreme value theory [63], for any i, $\max_{j\in S(i)}\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2$ and its average behave like log n at large n. Therefore, there is always at least one user with channel satisfying

$$\max_{j\in S(i)}\left|\mathbf{h}_j\mathbf{u}_i^H\right|^2 > \log n = \lambda_{\text{tr}} \;\; \forall i:1..M \tag{4.32}$$

for all basis vectors.

■

Lemma 4 indicates that a good threshold value is $\lambda_{\text{tr}} = \log n$. Using this threshold, the feedback load can be expressed as $n\Pr\{\lambda_i > \lambda_{\text{tr}}\}$ $(i:1..M)$ real and integer numbers. Note that, in practice $\Pr\{\lambda_i > \lambda_{\text{tr}}\}$ is much smaller than 1 and therefore, inserting this threshold value decrease feedback overhead dramatically.

4.5.2. Complexity

The proposed NUS-CAP scheme which is the most complex amongst four schemes just needs n comparisons at the transmitter and M projections and M comparisons at each user while DPC, optimal and suboptimal zero-forcing schemes are much more complex. For example, the zero-forcing algorithm proposed in [66], needs almost $\sum_{i=1}^{M}(n-1)(i+1) > nM(M+1)/2$ projections and nM comparisons at the transmitter to find a near-orthogonal channel vector set. In other words, the proposed NUS-CAP method has a much lower complexity than the other optimal and sub-optimal schemes, and also distributes low processing among users in the system.

4.5.3 Selection of Basis

As mentioned at the beginning of Section 4.4, the set of orthonormal basis $\mathbf{U}$, can be generally any arbitrary orthonormal basis in C^M and as shown in Section 4.5, the proposed schemes can still work fine. However, orthonormal basis can be optimized by a very minor increase in the feedback load from active users to the transmitter.

Let assume that instead of just one predetermined orthonormal basis, we have a number of k bases, $\mathbf{U}^1$, $\mathbf{U}^2$, …, $\mathbf{U}^k$. Both the transmitter and users know these bases. Each user calculates its channel vector projections onto each of these bases vectors and reports the best value along with the indices of this basis and its own basis. This requires one real and *two* integer numbers (instead of one real and one integer). From the reported information, the transmitter selects the orthonormal basis, corresponding to M users with the best orthogonality measures. As an illustrative example, consider M=2 and two orthonormal bases to select. If we select the first one as standard basis for C^2, i.e., $\mathbf{U}^1 = \{(1,0),(0,1)\}$, the other one should be the one whose vectors are as much non-orthogonal as possible to the standard basis, i.e., $\mathbf{U}^2 = \{\frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2}}(1,-1)\}$. Similarly, if we want to select three orthonormal bases, a reasonable choice is $\mathbf{U}^1 = \{(1,0),(0,1)\}$, $\mathbf{U}^2 = \{(a,b),(b,-a)\}$ and $\mathbf{U}^3 = \{(b,a),(a,-b)\}$ where $a = 0.5, b = \sqrt{3/4}$.

The same approach can be applied for larger M and larger number of orthonormal bases. Increasing the number of orthonormal bases can greatly improve the performance of the proposed schemes at the cost of a slight increase in the feedback overhead. More precisely, using k different orthonormal bases, the probability in (4.23) is linearly increasing with almost nk (instead of n) and the probability in (4.25) is linearly proportional to $1/nk$ (instead of $1/n$). It means faster orthogonality rate by increasing the number of users and hence, the proposed schemes become optimal at lower number of users.

4.6. Numerical Results

In this section, we examine the ability of zero-forcing precoder schemes in terms of ergodic capacity by means of simulation. We also compare the complexity and feedback overhead of the proposed precoding schemes to that of DPC and full knowledge based zero-forcing precoders.

We consider a four-antenna transmitter (M=4), for all simulations. Besides, for our schemes we just consider one basis vector set. First, let compare four different schemes in two different scenarios; i.e. in low and high average users' channel power. Figure 4.1

shows the sum rates of different user selection and precoding schemes in two low and high SNR regimes. By SNR regimes we mean the channel SNRs averaged over all available users' channels. As shown, while PUS selection strategy works better at low SNRs, NUS strategy outperforms it at high SNR regime. Therefore, one can conclude that when the average users' SNR is high, it is more reasonable to use NUS selection scheme while in low average SNR, PUS is more beneficial. Also, consider that at very fast fading environment, it would be more desirable to use OP rather than CAP schemes. Based on this discussion, the diagram in Figure 4.2, gives us an idea how to select between different schemes. Note that, because of similarity of structure and algorithms, it is possible to switch between them in a real time system. In other words, when fading is very fast the system uses OP schemes and in slow fading environment switches into CAP schemes. On the other hand, in high quality channel, it works in NUS regime while in poor-conditioned channels it switches into PUS regimes where users' channel power becomes a very important and significant factor.

Next, we simulate the feedback load, complexity and performance of our scheme and compare it with that of optimum and sub-optimum schemes. In Figure 4.3, the feedback load of the schemes for different number of users (n) is illustrated. As shown, our CAP schemes needs much lower amount of feedback compared to full channel knowledge based schemes and this difference in feedback overhead increases with the number of the users in the system. For comparison, it has been compared with the precoding scheme proposed in [63]. At large n, it needs the same amount of channel feedback as that of [66].

Figure 4.4 shows the complexity of different schemes in term of the number of CPU operation needed to finish the optimization tasks. Complexity of NUS-CAP scheme grows linearly with the number of users compared to optimal zero-forcing and DPC that their complexity is exponentially increasing with the number of users. Besides, NUS-CAP scheme needs less CPU clock compared to the full channel knowledge based zero-forcing in [66].

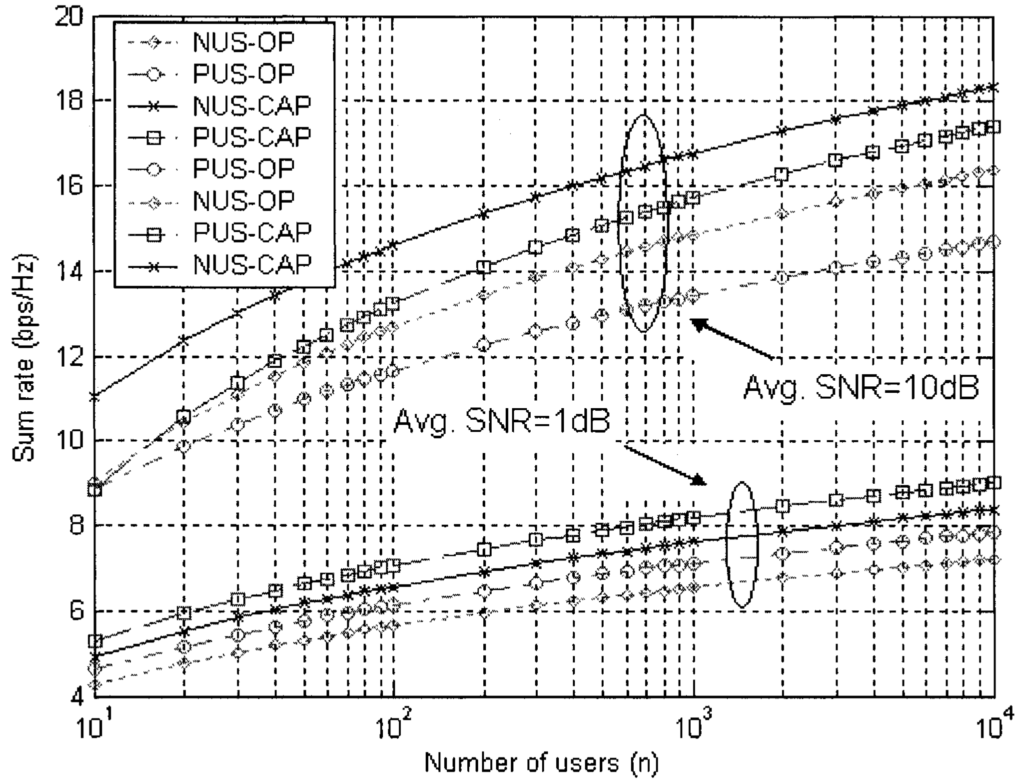


Figure 4.1: Comparison of sum rates of different precoder schemes in different SNR regimes

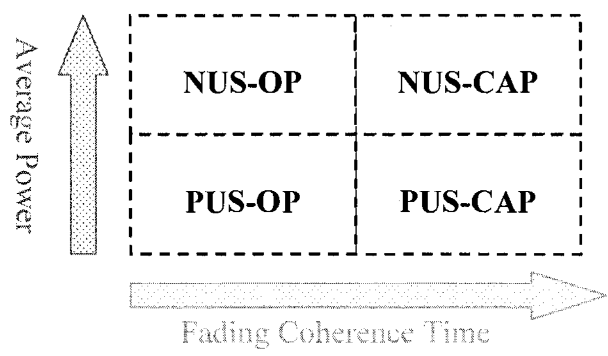


Figure 4.2: Selection guideline based on average power and fading coherence time of the channel

In Figure 4.5, achievable sum rates of different precoding schemes are compared with each other for different number of users. We assume a fixed available power of P=10dB at the transmitter. Different schemes considered are random user selection (M users are selected randomly each time), scheduling zero-forcing discussed in [66], random zero-forcing (OP) scheme proposed in section 4.5, NUS-CAP precoding scheme, full channel zero-forcing scheme in [66], optimal zero forcing described in Section 4.3 and dirty paper coding. Due to high level of complexity, the sum rate of DPC and optimal zero-forcing has been just depicted for small number of users. As shown, NUS-CAP scheme is better than the schemes need partial channel knowledge. Furthermore, the achievable sum rate of our scheme is close to that of full channel knowledge based schemes and the rate of growth is the same as that of optimal schemes at the limit of large number of users. Note that, NUS-CAP scheme works better in high number of users; however, by increasing the number of basis sets it is possible to compensate this drawback in low number of users. As mentioned we have just considered one basis vector set for our proposed schemes. By increasing the number of basis sets, our schemes perform better even when the number of users is small.

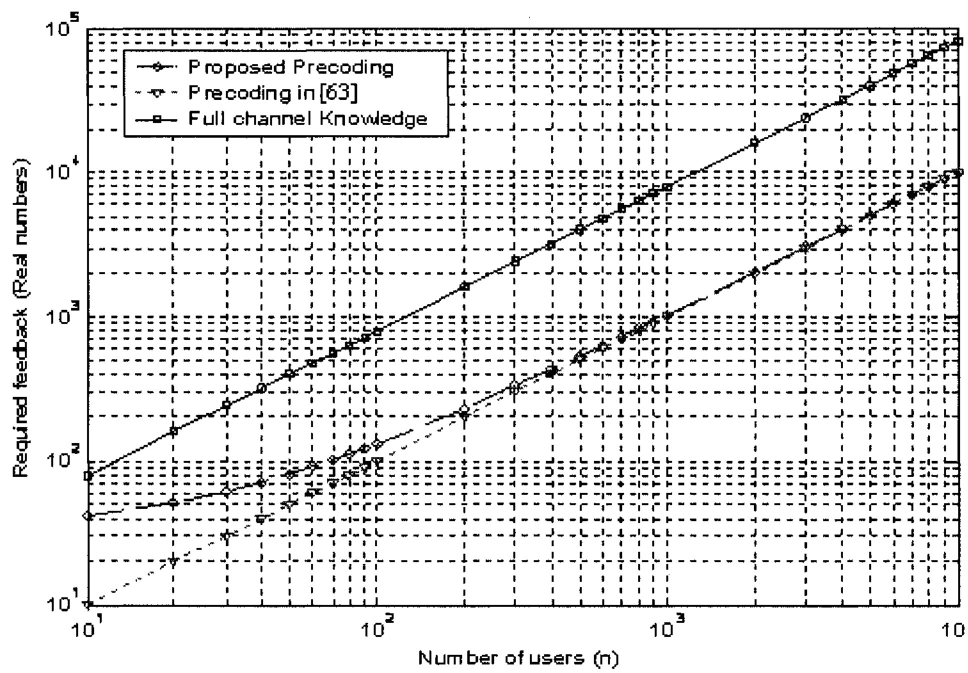


Figure 4.3: Feedback overhead comparison of different schemes with M=4

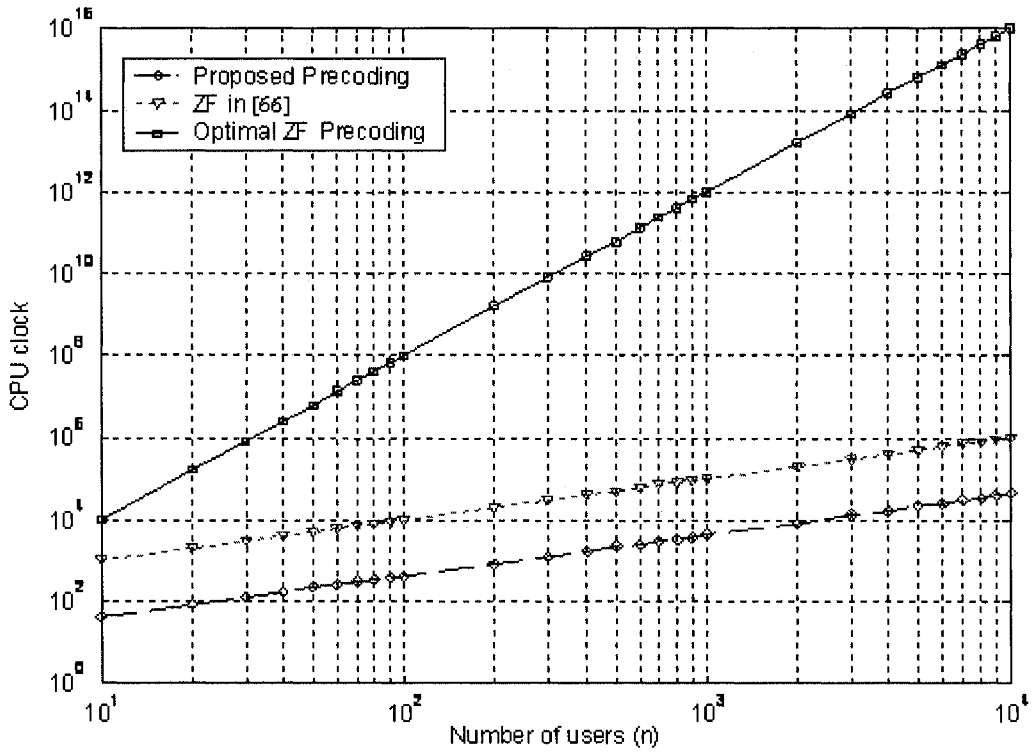


Figure 4.4: Complexity comparison of different schemes with M=4

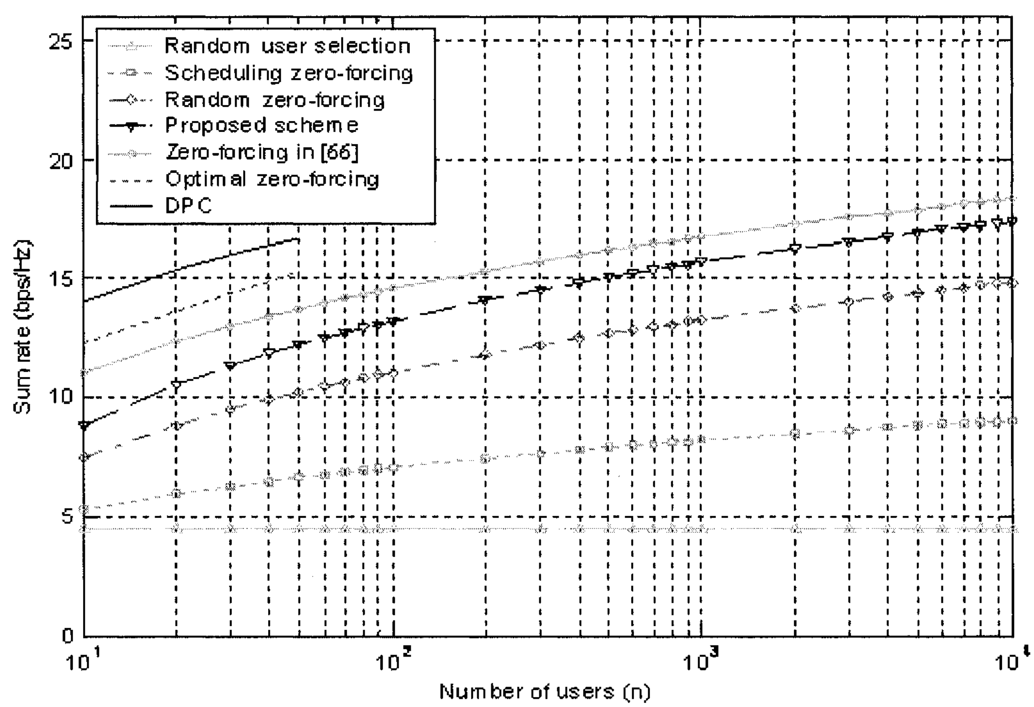


Figure 4.5: Sum rates for different schemes with M=4 and P=10dB

4.7. Chapter Summary

In this chapter, we investigated the design of a partial knowledge based precoder scheme for broadcast MIMO systems. By partial knowledge we mean that the transmitter does not have the instantaneous channel realizations of all active users in the network. This assumption arises because of the fast changing nature of real wireless channels that makes full channel knowledge difficult to obtain at the transmitter.

We proposed different combined user-selection/precoding schemes with no or small amount of channel feedback and low complexity of algorithms, which can achieve the performance of schemes based on full channel knowledge in the limit of large number of users. More precisely, it was shown that the ergodic capacity offered by the proposed schemes can achieve the same growing rate with the number of users (log log n) as that of DPC and optimal zero-forcing precoding. Note that, regardless of the complexity of these optimal algorithms and the time takes the transmitter executes them, they also need a large amount of channel feedback overhead, which makes them and similar approaches difficult to implement in real wireless communication systems.

The ability of proposed schemes in achieving optimality in terms of ergodic capacity growth rate was demonstrated by analysis and simulation. It was shown that, by increasing the number of users, the proposed methods need lower amount of feedback and processing load as compared to others. Moreover, in comparison to each other, each of the proposed schemes provides the best performance in different propagation scenarios and SNR regimes. As the proposed schemes share the same basic structure, it is possible for the base station to switch between them in order to adjust with the physical propagation environment while necessary. This fact also gives a degree of freedom and flexibility in the design of the system and also provides more robustness in compensating for channel changes.

Chapter 5

Relay Selection and Precoding for Cooperative Relay Systems[1]

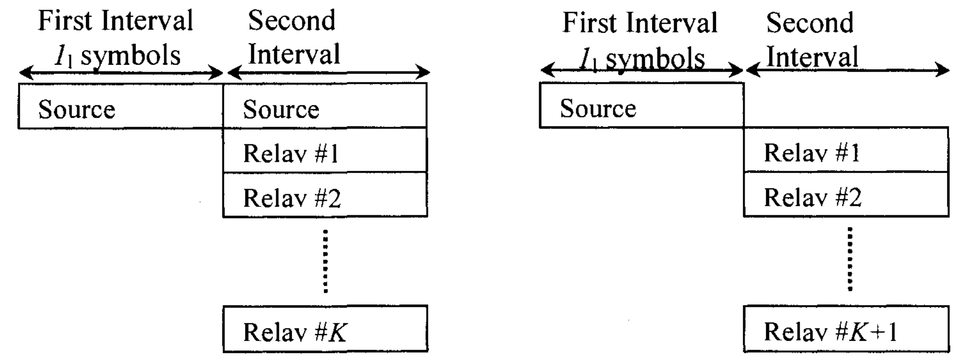
5.1. Relay System Model

We consider a relay system composed of one single-antenna transmit (source) node, a number of single-antenna relay nodes and one multi-antenna receive (destination) node with n antennas. We assume that all terminals are half-duplex; i.e. they can both transmit and receive at different times but just one at a specific time.

Considering this setup, the operation of the system can be described as follows. First, source node transmits a symbol block of length l_1 symbols to the destination as well as the relays. This step takes l_1 symbol blocks to complete. In the second transmission interval (of length l_2), each relay sends a part of the information to the destination. Note that in the second interval, all the relays transmit simultaneously.

We assume block (quasi static) fading channels over blocks of length $l_1 + l_2$ with the channel coefficients are assumed to be Rayleigh distributed. It means that they are fixed during the coherence interval of $l_1 + l_2$ but change from one interval to another. Depending on the nature of the block fading in the channel, quality of source destination channel and number of available relay nodes, source node can cooperate in the second interval or not. Mainly, if the quality (in term of receive SNR) of the source destination channel is high it might be beneficial that source also participates in transmission during the second interval and therefore transmits during the interval of length $l_1 + l_2$. On the contrary, if the direct channel quality is poor and sufficient relay

[1] This chapter has been partially presented in [94].



(a) (b)

Figure 5.1: Transmission frames for AF cooperative system: (a) source cooperates, (b) source remains silent in the second transmit interval

nodes are available in the system it is meaningful that source remains silent during the second interval.

The diagram of transmission of the system during two subsequent intervals of length l_1 and l_2 is illustrated in Figure 5.1 for two different scenarios discussed above. While in the former, source node also collaborates in the second interval; in the latter, it remains silent. Note that the number of relays in two scenarios is different. It is because in the first scenario source node also plays the role of one of the relays in the second interval. In this chapter, we consider the first case and assume the source also cooperates in the second interval, however, most of the derivations and results can be applied to the later case with slight modifications.

Let assume at any transmission interval and based on a specific criterion K relays are selected amongst all available relays. The system model in the first transmit interval can hence be written as:

$$\mathbf{Y}_1 = \mathbf{h}_0.\mathbf{x}^T + \mathbf{N}_1 \tag{5.1}$$

where $\mathbf{x} = [x_1, x_2, ..., x_{l_1}]^T$ is the $l_1 \times 1$ transmit vector, $\mathbf{Y}_1$ and $\mathbf{N}_1$ are the $n \times l_1$ receive and noise matrices in the first transmit interval, respectively and $\mathbf{h}_0$ is the $n \times 1$ source destination channel vector. Elements of channel vector and noise matrix are assumed to be circularly symmetric complex Gaussian random variables. In the second transmit interval:

$$\mathbf{Y}_2 = \mathbf{h}_0 \tilde{\mathbf{x}}_0^T + \sum_{i=1}^{K} \mathbf{h}_i \beta_i (g_i \tilde{\mathbf{x}}_i^T + v_i) + \mathbf{N}_2 \tag{5.2}$$

where $\mathbf{Y}_2$ and $\mathbf{N}_2$ are the $n \times l_2$ receive and noise matrices in the second transmit interval, respectively; g_i and v_i are the channel gain and the noise corresponding to the channel between source node and the i^{th} relay and $\mathbf{h}_i$ is the channel vector corresponding to the channel between i^{th} relay and destination nodes. Moreover, $\tilde{\mathbf{x}}_i$ $(i:0..K)$ are $l_2 \times 1$ vectors each contains a portion of transmit vector $\mathbf{x}$ in (5.1). We assume that all relays (including source node) participate equally in the transmission and each symbol is relayed once. Therefore, $l_2 = l_1/(K+1)$.

In (5.2), β_i is the repetition gain of the i^{th} relay node and assuming an average power limitation of P_i and noise power of N_0 at relay, it can be selected as [35]:

$$\beta_i \le \sqrt{\frac{P_i}{|g_i|^2 P_i + N_0}} \tag{5.3}$$

This selection can satisfy the power criterion at each node with high probability. If there is a total power constraint over all relays, the repetition gains should be jointly optimized to optimize a performance measure while satisfying the power constraint. This problem is however, out of the focus of this chapter. We therefore consider β_i's coefficient to be inside channel vectors $\mathbf{h}_i$'s and omit them in analysis for brevity.

5.2. Relay Selection

In order to develop a transmission scheme based on the model in Section 5.2, the first step is to select the relays. In that, a natural question is what the optimum number of the relays is. In any system, the number of candidates for the relay nodes would be more than the necessary numbers. For example, in an MA system, the number of idle user terminals that can help the specific source node (user) to transmit its signal is very large. The same thing applies to ad-hoc and sensory networks where the density of nodes is normally very high.

To answer this question, we suggest the use of multiplexing-diversity trade-off [68] in relay systems. Using this tool, it is possible to specify the optimum number of relays for any specific rate. To this end, let us first give a formal definition of multiplexing and diversity gain [68].

Definition: A transmission scheme is said to achieve spatial multiplexing gain r and diversity gain d if the data rate:

$$\lim_{SNR \to \infty} \frac{R(SNR)}{\log SNR} = r \tag{5.4}$$

and the average error probability:

$$\lim_{SNR \to \infty} \frac{P_e(SNR)}{\log SNR} = -d \tag{5.5}$$

where $SNR=P/N_0$ is the signal to noise ratio. At any specific signal to noise ratio, these two parameters are related to each other and any increase in multiplexing gain will result in a reduction in diversity and vice versa. Therefore, it is natural to define diversity as a function of multiplexing gain; $d(r)$. The optimum diversity gain $d^*(r)$ is defined as the supremum of diversity achieved by all schemes. It is further possible to define maximum multiplexing gain and diversity order of the system as:

$$d^*_{\max} = d^*(0) \quad r^*_{\max} = \sup\{r : d^*(r) > 0\} \tag{5.6}$$

For example, a MIMO system with m transmit and n receive antenna has the multiplexing-diversity trade-off of [68]:

$$d^*(r) = (m-r)(n-r) \ ; \ 0 \le r \le \min\{m,n\} \tag{5.7}$$

and therefore, for such a system $d^*_{\max} = mn$ and $r_{max} = \min\{m, n\}$.

Now we are ready to discuss the multiplexing-diversity trade-off of a relay system with K relays and system model described in (5.1) and (5.2). The following claim states the trade-off for this system. The formal proof of this claim however seems to be very involved and we just give an intuitive proof.

***Claim* 5.1:** For the NAF relay system whose model is described by (5.1) and (5.2) with $l_2 \ge K+n-1$, the multiplexing-diversity trade-off over transmission length of l_1 + l_2 can be written as:

$$d^*(r) = n(1-r) + nK(1-\alpha r)^+ \ ; \ 0 \le r \le 1 \tag{5.8}$$

where $\alpha = K+1$.

Proof: At any specific rate, diversity order of this system can be defined as the sum of the diversity gains achieved by source node d_s and relay nodes d_r, i.e.:

$$d(r) = d_s(r) + d_r(r) \tag{5.9}$$

But as the source node transmits all the times, it can be viewed as a MIMO system with one transmit and n receive antenna and therefore at a given rate r, $d_s(r) = n(1-r)$. On the other hand, the transmission of the relays in the second interval can be viewed as a multiple access (MA) system composed of K users with similar rates. This is sometimes called *symmetric* MA system. Two separate cases can be recognized:

1) $K < n-1$: In this case, for each user rate one is achievable and the system is always *under-loaded* [71]. Note that to achieve rate one for each user, one should

have $l_2 \geq K+n-1$. Therefore, the diversity order offered by an under-loaded MA system at rate r' can be written as:

$$d_r(r') = \sum_{i=1}^{K} d_i(r') = nK(1-r') \tag{5.10}$$

where $d_i(r')$ is the diversity offered by user i at rate r'. But as the relay nodes transmit only $l_2/(l_1+l_2)$ portion of time and given the system model of Section 5.2:

$$r' \leq \frac{l_2}{l_1+l_2} = \frac{K+n-1}{K(K+n-1)+K+n-1} = \frac{1}{K+1} \tag{5.11}$$

Therefore, maximum achievable rate of relays is $1/(K+1)$. Note that r and r' are not necessarily the same. On the other hand, any increase in the rate of the source node r will result in at least $K+1$ times increase in r' because relays transmit only $1/(K+1)$ portion of times and therefore $r' \geq (K+1)r$. Considering this fact and plugging $d_s(r)$ and $d_r(r)$ from (5.10) into (5.9) will result in the optimal trade-off equation of (5.8).

2) $K \geq n-1$: In this case, it is not possible to achieve the rate one for all relay nodes. Moreover, for $r < n/(K+1)$ the system is *under-loaded* and for higher rates it is *over-loaded.* Nevertheless, as showed in the previous case the rate of the MA system can not exceed $1/(K+1)$. Therefore, the relay system with model described in Section 5.2 has the multiplexing-diversity trade-off of (5.8) and it completes the proof of Claim 1.

■

Note that in the special case of n=K=1, (5.8) will be reduced to (5.12):

$$d^*(r) = (1-r)+(1-2r)^+ \;\; ; \;\; 0 \leq r \leq 1/2 \tag{5.12}$$

which is the same as the result proposed by [75] and [76]. Figure 5.2 shows the diversity multiplexing trade-off of a system composed of K relay nodes whose model described in Section 5.2. As seen, for multiplexing gains larger than $1/(K+1)$ there is no gain using the relays and in fact the NAF relay system with K relays can not provide multiplexing gains more than $1/K+1$ and therefore for higher gains system switches into direct transmission.

Figure 5.3 is more interesting from a practical point of view. It shows that a system composed of K relays not only can not transmit at multiplexing gains higher than $1/(K+1)$ but also in a specific multiplexing gain range, it does not achieve to higher diversity gains compared to a system with K-1 relays. In fact, if $1/2K \leq r \leq 1/(K)$ the NAF system with model described in Section 5.2 composed of K-1 relay nodes is superior to a system with K relays in term of multiplexing-diversity trade-off. This gives us a practical rule to select the optimum number of relay nodes.

The next step after finding the optimal number of relays is to find the optimum relay nodes among a number of available nodes in the system. Normally, the number of the available nodes in the system is much larger than the number of the antennas at the destination node (n). Selection criterion is however very dependent to the transmission scheme used for sending information from source to destination through relay nodes. Therefore, we postpone this discussion to the end of the next section.

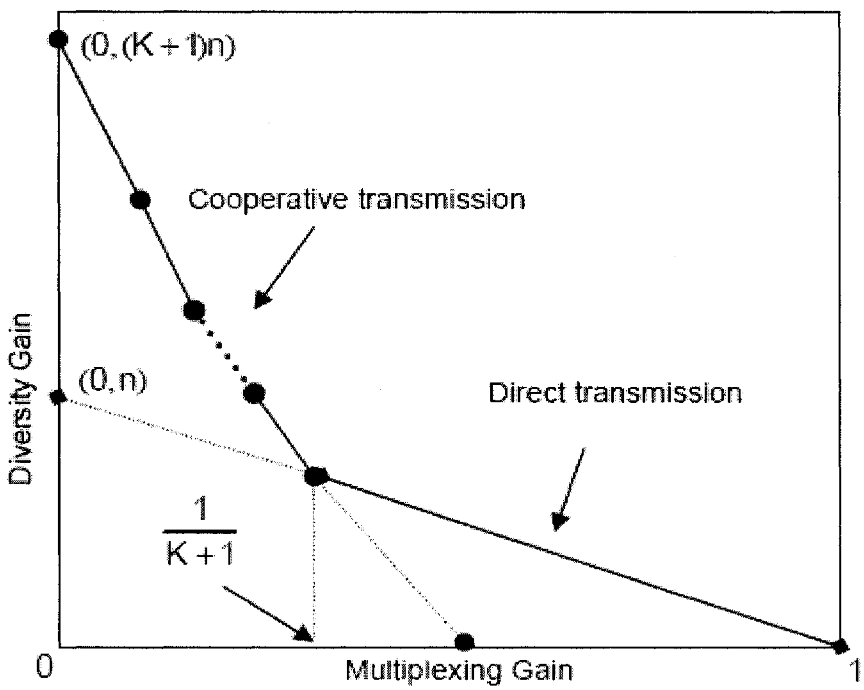


Figure 5.2: Multiplexing-diversity trade-off for a relay system with Krelay nodes and n receive antenna

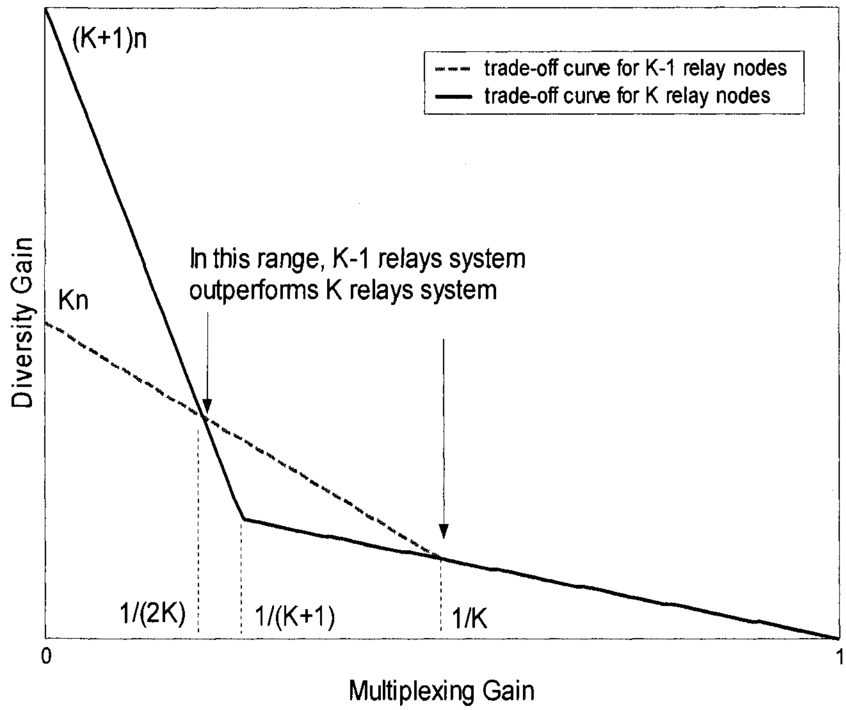


Figure 5.3: Multiplexing-diversity trade-off comparison for two systems composed of *K*-1 and *K* relays

5.3. Cooperative Transmission Scheme

It has been shown that Diagonal BLAST (D-BLAST) transmission with ML detection at the receiver is optimal in terms of diversity-multiplexing trade-off. This gives us an idea to design a transmission scheme for relay systems. The nice characteristic of BLAST system that makes it suitable for cooperative relay transmission is that there is no *cross coding* between transmit antenna elements. Therefore, it is a suitable candidate for transmission in relay systems in which no collaboration or cross coding is possible. The whole idea is that each relay will receive a part of transmitted signal from source and will retransmit the receive data in a specific order.

In D-BLAST, the input data stream is first divided into separate substreams. Next, each stream is transmitted on different antennas in different time slots. The transmission is such that each stream constructs a diagonal in the transmitted symbol matrix. For example, in a 3×3 D-BLAST system, the transmitted signal matrix can be written as:

$$\mathbf{X}=\begin{bmatrix} 0 & 0 & \mathbf{x}_1^{(1)} & \mathbf{x}_1^{(2)} & \cdots \\ 0 & \mathbf{x}_2^{(1)} & \mathbf{x}_2^{(2)} & \mathbf{x}_2^{(3)} & \cdots \\ \mathbf{x}_3^{(1)} & \mathbf{x}_3^{(2)} & \mathbf{x}_3^{(3)} & \mathbf{x}_3^{(4)} & \cdots \end{bmatrix} \in C^{3\times l_2} \tag{5.13}$$

where $\mathbf{x}_i^{(k)}$ denotes the symbol block of length l from k^{th} substream which is transmitted over antenna i. In (5.13), we assumed that the length of the block is l_2 (the length of the second transmit interval in our formulation), the length of each substream in this example is $3l$ and the number of independent substreams in each block turns out to be (l_2/l) - 2. Note that the optimality of D-BLAST in term of multiplexing-diversity trade-off is valid if we ignore the slight increase in throughput due to the zero blocks transmitted at the start of each transmit interval. However, this overhead can only be ignored when l_2 is sufficiently large.

On the other hand, ML detection of D-BLAST becomes very complex specially when the dimension of the channel (number of transmit and receive antennas) or the length of transmission l_2 increase. Therefore, in the context of D-BLAST several suboptimum detection schemes have been proposed to overcome the substantial complexity of ML detection. The problem is that these schemes undermine the optimality of D-BLAST in term of multiplexing-diversity trade-off.

The simplest detection method of D-BLAST is called successive nulling and cancellation. To describe the method briefly, let consider the transmitted matrix **X** in (5.13) as an example. The successive nulling and cancellation receiver, first estimates $\mathbf{x}_3^{(1)}$ and then estimates $\mathbf{x}_2^{(1)}$ by nulling $\mathbf{x}_3^{(2)}$ and considering it as interference. This process will continue over sufficient number of columns of **X** until all symbols of first substream are estimated. The estimates of the first substream are then fed to a joint decoder to detect the first substream. For detection of the second substream, the contribution of the first substream is subtracted from the rest of the signals and each of the substreams will be detected in a specific order. To achieve to the best result, an ordering of substreams is also necessary.

It can be shown that for $\mathbf{x}_i^{(k)}$ blocks with length $l \geq n$, this scheme in a MIMO channel with n transmit and receive antenna, achieves the trade-off equation in (5.14) [71]:

$$d^*(r)=\frac{1}{2}(n-r)(n-r+1) \;\; ; \;\; 0 \leq r \leq n \tag{5.14}$$

which is clearly suboptimum compared to (5.7). Now turning back out attention to the problem of relay transmission, let consider the channel matrix $\mathbf{H}=[\mathbf{h}_K\ \mathbf{h}_{K-1} \dots \mathbf{h}_1\ \mathbf{h}_0]$ in which $\mathbf{h}_i \in C^n$ with independent complex Gaussian entries is an $n\times 1$ vectors corresponding to each of the channel vectors. In that, $\mathbf{h}_0$ represents source-destination channel vector.

Assuming the same ordering as of the indices in **H** (i.e. relay K is first and source last detected), in successive nulling and cancellation method each channel vector $\mathbf{h}_i$ is in fact divided into two parts. The first part is $\mathbf{h}_{i\|}$ which is in the vector space spanned by all other channel vectors with lower indices $\mathbf{h}_j$ $(0 \le j < i)$:

$$\mathbf{h}_{i\|} = \sum_{j=0}^{i-1} \frac{<\mathbf{h}_i, \mathbf{h}_j>}{\left|\mathbf{h}_j\right|^2} \mathbf{h}_j \tag{5.15}$$

where <.,.> represents internal product and $|.|^2$ denotes vector norm. The second part is $\mathbf{h}_{i\perp}$ that is orthogonal to this range space:

$$\mathbf{h}_{i\perp} = \mathbf{h}_i - \mathbf{h}_{i\|} \tag{5.16}$$

Clearly, $\mathbf{h}_{i\perp}$ contributes in the *SNR* in detection of i^{th} substream while $\mathbf{h}_{i\|}$ causes the interference of substream i^{th} over all other substreams whose indices are greater than i. The suboptimality of successive nulling and cancellation method rises from this fact that the detected substreams have *non-zero* interference over remaining substreams. This gives us an insight to the question of relay selection from Section 5.3. To make the successive nulling and cancellation optimum, one should select relay nodes such that $\mathbf{h}_{i\|}$ is close to zero for all relay channel vectors $\mathbf{h}_i$ $(1 \le i \le K)$. Note also that there is also no need to send zero overheads when channel vectors are orthogonal.

Starting from the source-destination channel vector and assuming there are $K'(> K)$ relay nodes in the system, by relay selection in K steps one can build an ε-orthogonal vector set composed of K+1 vectors given that $K<n$ in the following way:

$$\mathbf{h}_i = \arg \max_{j=i+1:K'} \frac{\left|g_j\right|\left|\mathbf{h}_{j\perp}\right|}{\left|g_j\right|\left|\mathbf{h}_j\right|} = \frac{\left|\mathbf{h}_{j\perp}\right|}{\left|\mathbf{h}_j\right|} \qquad i = 1:K \tag{5.17}$$

where g_j is the channel gain of the source to the j^{th} relay node defined in (5.2) and:

$$\varepsilon = \max_{1 \le i \le K} \frac{|\mathbf{h}_{i\|}|^2}{|\mathbf{h}_i|^2} = \frac{|\mathbf{h}_{K\|}|^2}{|\mathbf{h}_K|^2} = 1 - \frac{|\mathbf{h}_{K\perp}|^2}{|\mathbf{h}_K|^2} \tag{5.18}$$

To exclude selection of relays with poor channel gain, we make the selection among the relays whose channel gain $|g_j||\mathbf{h}_j|$ is greater than a specific threshold value. Note that if the vector set of $\{\mathbf{h}_i\,(0 \le i \le K)\}$ is an orthogonal set, $\{\{g_i\mathbf{h}_i\,(0 \le i \le K)\}\}$ is also orthogonal. The selection algorithms such as the one proposed in [17] can be applied to reduce the complexity of (5.17). Here, we give an asymptotic analysis of this user selection scheme when the number of available relay nodes is large.

***Claim* 5.2:** For large number of users, the cdf function of ε in (5.18) is linearly increasing functions of K'; the number of available relay nodes in the system, i.e. $F_\varepsilon(x) = \Pr(\varepsilon < x) \sim O(K'x)$.

Proof: We prove the Claim by finding the cdf function of ε from (5.18). For any vector *i*, the random variables $<\mathbf{h}_i, \mathbf{h}_j>^2$ is $\chi^2(2)$ distributed given that $\mathbf{h}_j$ is a fixed vector. Fixing $\mathbf{h}_i$ (*i*=0:*K*-1) and finding distribution of $|\mathbf{h}_{K\|}|^2$ and $|\mathbf{h}_{K\perp}|^2$, one can see from (5.15) and (5.16) that they are $\chi^2(2K)$ and $\chi^2(2(n-K))$ distributed, respectively. Considering that $|\mathbf{h}_K|^2$ is $\chi^2(2n)$ distributed, it is then easy to calculate the pdf of $t = |\mathbf{h}_{K\perp}|^2 / |\mathbf{h}_K|^2$ as:

$$f_t(x) = \int_0^\infty \frac{y^{K-1}e^{-yx}}{(K-1)!} \cdot \frac{y^{n-1}e^{-y}}{(n-1)!} dy = \binom{K+n-2}{n-1} \frac{1}{(1+x)^{K+n-1}} \tag{5.19}$$

Therefore the cdf function can be written as:

$$F_t(x) = 1 - \frac{C}{(1+x)^{K+n-2}} \tag{5.20}$$

where $C = (K+n-3)!/(n-1)!(K-1)!$. Hence, the cdf of ε in (5.18) considering the relay selection over $K'-K$ nodes in (5.17) can be found as:

$$\begin{aligned} F_\varepsilon(x) &= P(\varepsilon < x) = P(\max_{K'-K} t_j > 1-x) \\ &= 1 - F_t^{K'-K}(1-x) = 1 - \left(1 - \frac{C}{(2-x)^{K+n-2}}\right)^{K'-K} \end{aligned} \tag{5.21}$$

For small *x* the right hand side of (5.21) can be approximated as:

$$F_\varepsilon(x) \approx 1-(1-\frac{C}{2^{K+n-2}}(1+2^{-K-n-1}(K+n-2)x))^{K'-K}$$
$$\approx \frac{(K'-K)C}{2^{K+n-2}}2^{-K-n-1}(K+n-2)x \sim O(K'x) \quad (5.22)$$

and it completes the proof.

■

The result of Claim 2 is important because it shows that by increasing the number of relays, it is more probable to find a set of relays with *near* orthogonal channel vectors to use in our cooperative transmission scheme. Note also that the destination node does not necessarily need to know the source-relay channel gains g_i and the power $|g_i|^2$ is sufficient for relay selection as it does not change the orthogonality conditions.

Considering the above discussions in Sections 5.3 and 5.4, the algorithm of transmission in relay system can be summarized as follows:

1) Assuming that the destination node knows all relay-destination channel vectors as well as source-relay channel powers, for a given rate it can select the optimum number of relays based on the multiplexing-diversity trade-off discussed in Section 5.3 and the optimal relay nodes based on the selection algorithm in Section 5.4.

2) Destination then informs the source node of the number of available relays nodes (K) and source node divides the data into independent substreams of suitable length and transmits the data to the destination and relay nodes during the first transmit interval in Figure 1.

3) In second transmit interval, source and relay nodes will cooperate to send the information already transmitted from source node in the first interval using a distributed BLAST approach with transmit matrix in (5.13).

It is also possible to compensate the non-optimality of the selected channel vector set by applying a precoding approach at the source node. The precoding structure is very close to that of MIMO systems and in the context of cooperative communications can be called *distributive precoding*. The coefficients of the precoder matrix are calculated at the destination node and fed back to the source node. This can buy us some more coding gain at the cost of a slight increase in the complexity of transmission at the source node.

5.4. Numerical Results

We examine the ability of precoder in terms of ergodic capacity by means of simulation. We consider the outage probability as the performance measure of the system. Figure 5.4 shows the outage probability of a system with n=2 antennas and we assume that the relay node is selected among a number of available nodes in the system. We further assume that there is no coding applied at the source node. While finally according to multiplexing-diversity trade-off, just one of the relay nodes is selected as the optimum candidate, Figure 5.4 shows that increasing the number of available relay nodes from 10 to 100 can improve the outage performance of the system. Also, adding the distributive precoder at the transmit side can also buy some gain in term of outage probability. The same simulation can be done in term of probability of error and the same conclusion can be made.

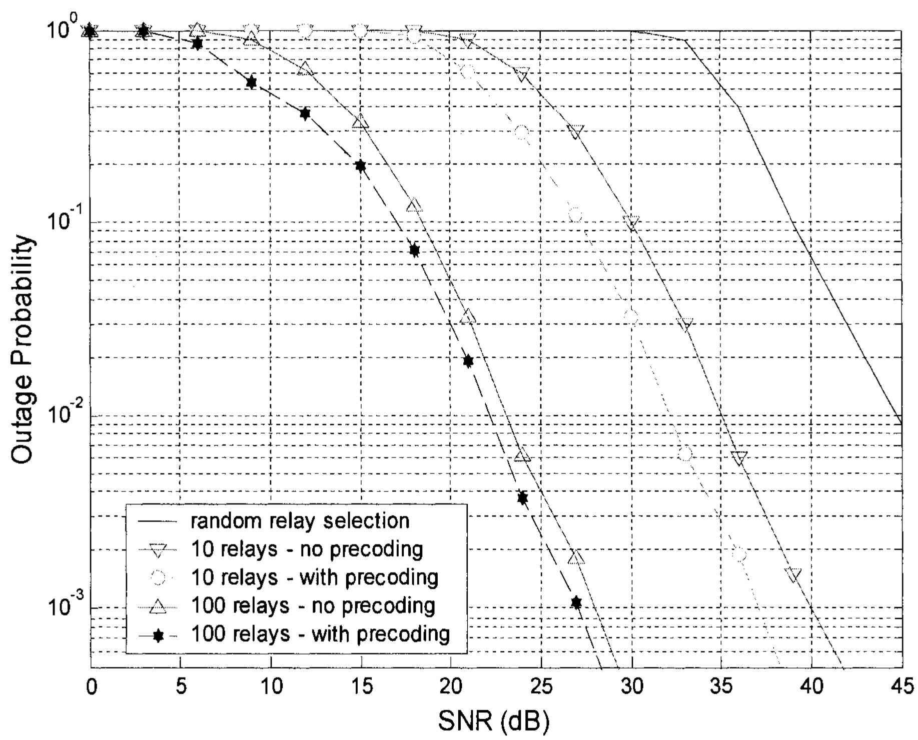


Figure 5.4: Outage probability of system with different number of avilable relays

5.5. Chapter Summary

Cooperative diversity has gained a lot of attention in modern communication systems. It is mainly due to the attractive diversity-multiplexing trade-off offered by such systems compared to direct transmission systems. There are, however, currently many researches focused on transmission schemes that practically benefit from this promising level of multiplexing gain and diversity order. On the other hand, the ability of cooperative systems when multi-antenna terminals are used instead of single antenna has not been investigated extensively. It remains an open problem to specify the optimum number of relays in a multi-antenna cooperative relay system and the way to share the information amongst these relay nodes. In this chapter, our focus was on the design of a system in which the receive node is equipped with multiple antennas. We addressed the problem of optimal number of relays as well as relay selection and showed that the proposed cooperative transmission scheme can provide an appealing multiplexing-diversity trade-off. We next focused on the design of a transmission scheme that can achieve this trade-off. We showed that a distributed BLAST transmission in conjunction with successive nulling and cancellation at destination can achieve optimal trade-off given that the relays are selected based on a specific criterion. Numerical results demonstrated the ability of the proposed method to increase the strength of relay networks in providing reliable communication.

Chapter 6

Optimum Combining and Precoding in Multi-Antenna Cooperative Relay Systems[1]

6.1. Multi-Antenna Relay System Model

In this chapter, we consider a relay system composed of a multi-antenna source (transmit) node, a multi-antenna relay node both with M antennas and one multi-antenna destination (receive) node with N antennas. We assume that all terminals are half-duplex; i.e. they can both transmit and receive at different times but just one at a specific time. We consider single symbol transmission in each time interval; however, it is not difficult to generalize the discussion to the case of multiple symbol transmission.

Considering this setup, the operation of the system can be described as follows. At a specific time instant, source node tends to transmit a symbol x of a pre-determined code book (or constellation) to the destination. It applied a precoding vector $\mathbf{w}_1$ of size $M\times 1$ to this symbol and broadcast it to the destination as well as the relay node. The relay node receives it and in the next time slot based on the specific protocol (AF or DF), multiplies it by another precoding vector $\mathbf{w}_2$ of the same size and resends this precoded version to the destination. Destination then combines the two received signals based on a maximum ration combining strategy.

Based on this, the received signal at the first time slot can be written as:

$$\mathbf{y}_1 = \sqrt{\gamma_1}\mathbf{H}_1\mathbf{w}_1 x + \mathbf{n}_1 \tag{6.1}$$

[1] This chapter has been partially presented in [95].

where $\mathbf{H}_1$ is the $N\times M$ forward channel matrix with normalized circularly symmetric Gaussian random entries, γ_1 is its corresponding SNR, $\mathbf{y}_1$ and $\mathbf{n}_1$ are the receiver and noise vectors of size $N\times1$, respectively. We assumed that the noise is also additive white Gaussian (AWGN). In the second time slot, the signal:

$$\mathbf{y}_2 = \sqrt{\gamma_2}\mathbf{H}_2\mathbf{w}_2\tilde{x} + \mathbf{n}_2 \tag{6.2}$$

is received where $\tilde{x}$ is either a detected version of x at relay node for DF or $\tilde{x} = \sqrt{\gamma_G}\mathbf{G}\mathbf{w}_1 x + \mathbf{n}$ for AF scenario. $\mathbf{G}$ is the $M\times M$ channel matrix between source and relay nodes, $\mathbf{n}$ is the $M\times1$ noise vector at relay and γ_G is its corresponding SNR. $\mathbf{H}_2, \gamma_2$, $\mathbf{y}_2$ and $\mathbf{n}_2$ are defined similar to their equivalents in (6.1).

The destination combines these two signals using two weight vectors $\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$ to construct the received signal. Therefore, we finally have:

$$y = \tilde{\mathbf{w}}_1^H\mathbf{y}_1 + \tilde{\mathbf{w}}_2^H\mathbf{y}_2 \tag{6.3}$$

The decision is made on y to detect the transmit symbol x. Our goal is to find two precoding ($\mathbf{w}_1$ and $\mathbf{w}_2$) and two weight vectors ($\tilde{\mathbf{w}}_1$ and $\tilde{\mathbf{w}}_2$) such that the received SNR is maximized. Here, maximizing SNR will minimize the probability of wrong decision over x. Also, we show that maximizing SNR in this scenario is equivalent to maximizing the instantaneous mutual information and ultimately the system capacity.

Note that, we consider the transmit power at each of the source and destination nodes are limited and it means that we can write:

$$\mathbf{w}_1^H\mathbf{w}_1 = \|\mathbf{w}_1\|^2 \le 1 \qquad \mathbf{w}_2^H\mathbf{w}_2 = \|\mathbf{w}_2\|^2 \le 1 \tag{6.4}$$

6.2. Generalized Maximum Ratio Combining (GMRC)

In this section, we study the optimal combining schemes at the receive side. We start by maximum ratio combining in point-to-point MIMO systems and gradually generalize it to optimal combining scheme for MIMO cooperative relay systems.

6.2.1. Point-to-Point MIMO system (Case A)

Let consider a MIMO system with M transmit and N receive antennas in Figure 6.1(a). The transmitted symbol x is precoded at transmitter by precoder $\mathbf{w}$ and the received vector is combined using vector $\tilde{\mathbf{w}}$ at the receiver. Therefore, the system model can be written as:

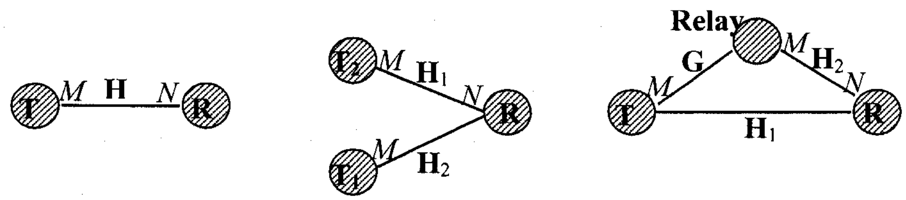


Figure 6.1: Diagrams of different transmission systems: (a) point-to-point MIMO, (b) multipoint-to-point MIMO and (c) relay-assisted MIMO

$$\mathbf{y} = \sqrt{\gamma}\tilde{\mathbf{w}}^H \mathbf{H}\mathbf{w}x + \tilde{\mathbf{w}}\mathbf{n} \tag{6.5}$$

where $\mathbf{H}$, $\mathbf{n}$ and γ are similar to $\mathbf{H}_1$ and $\mathbf{n}_1$ in (6.1). The following Lemma shows that maximum ratio combining is optimum in the sense that it maximizes the received SNR and system capacity.

***Lemma* 6.1:** The optimum combining scheme for the point-to-point precoded MIMO system whose model is stated in (6.5), is maximum ratio combining.

Proof: Writing the received SNR for the system described above, we get:

$$\mathrm{SNR} = \gamma \frac{\mathbf{w}^H \mathbf{H}^H \tilde{\mathbf{w}}\tilde{\mathbf{w}}^H \mathbf{H}\mathbf{w}}{\tilde{\mathbf{w}}^H \tilde{\mathbf{w}}} \tag{6.6}$$

From (6.5), one can directly conclude that to maximize SNR, the two vectors $\tilde{\mathbf{w}}$ and $\mathbf{H}\mathbf{w}$ should be in the same direction and hence one should select $\tilde{\mathbf{w}} = \mathbf{H}\mathbf{w}$. This is the case of maximum ratio combining. For this choice of $\tilde{\mathbf{w}}$, SNR will be derived as $\mathrm{SNR} = \gamma \mathbf{w}^H \mathbf{H}^H \mathbf{H}\mathbf{w}$. Therefore, maximum ratio combining indeed maximizes the receive SNR in precoded MIMO systems. Also, writing the instantaneous mutual information between x and $\tilde{x}$ will result in:

$$I(x;\tilde{x}) = \log(1 + \mathrm{SNR}) = \log(1 + \gamma\, \mathbf{w}^H \mathbf{H}^H \mathbf{H}\mathbf{w}) \tag{6.7}$$

which is clearly the maximum that can be obtained in a precoded MIMO system. Note that as Log is a convex function, maximizing SNR will ultimately result in maximizing mutual information. This completes the proof.

■

6.2.2. Multipoint-to-Point MIMO system (Case *B*)

Now let consider the case when two transmitters send similar information to a receiver but in different time instant. Both transmitters are equipped with M antennas while receiver has N antennas. Figure 6.1(b) illustrates this scenario. It is similar to DF protocol assuming perfect information recovery at the relay node. The system model is the same as (6.1) and (6.2) except that $\tilde{x}$ is replaced by x in (6.2).

***Lemma* 6.2:** The optimum combining scheme for the multipoint-to-point precoded MIMO system described above is a generalized maximum ratio combining (GMRC) scheme in which $\tilde{\mathbf{w}}_1 = \mathbf{H}_1\mathbf{w}_1$ and $\tilde{\mathbf{w}}_2 = \mathbf{H}_2\mathbf{w}_2$.

Proof: Combining (6.1), (6.2) and (6.3), one can write a compound system equation for the above scenario as:

$$y = (\tilde{\mathbf{w}}_1^H \ \tilde{\mathbf{w}}_2^H)\begin{pmatrix} \sqrt{\gamma_1}\mathbf{H}_1 & 0 \\ 0 & \sqrt{\gamma_2}\mathbf{H}_2 \end{pmatrix}\begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix} x + (\tilde{\mathbf{w}}_1^H \ \tilde{\mathbf{w}}_2^H)\begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix} \tag{6.8}$$

This is clearly a MIMO system with $2M$ transmit and $2N$ receive antennas. Because of Lemma 1, the optimum combining scheme for such a system is $\tilde{\mathbf{w}} = \mathbf{H}\mathbf{w}$ in which:

$$\mathbf{H} = \begin{pmatrix} \sqrt{\gamma_1}\mathbf{H}_1 & 0 \\ 0 & \sqrt{\gamma_2}\mathbf{H}_2 \end{pmatrix} y \ ; \ \tilde{\mathbf{w}} = \begin{pmatrix} \tilde{\mathbf{w}}_1 \\ \tilde{\mathbf{w}}_2 \end{pmatrix} \text{ and } \mathbf{w} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{pmatrix} \tag{6.9}$$

Reorganizing the equations, one can derive:

$$\tilde{\mathbf{w}}_1 = \mathbf{H}_1\mathbf{w}_1 \text{ and } \tilde{\mathbf{w}}_2 = \mathbf{H}_2\mathbf{w}_2 \tag{6.10}$$

■

Note that there is a fundamental difference between MRC and GMRC as MRC just considers combining in space domain over elements of receive antennas while GMRC includes combining over both space and time.

In the case of hybrid DF protocol where the relay resends its information only when it can regenerate the information correctly, the result of Lemma 2 is still valid. In general DF protocol, the result should be modified as there is a possibility that the information coming from relay node is already inaccurate. In this case, destination should apply a higher weight on source and a lower weight on relay information.

For simplicity, let consider transmission of BPSK symbols. The result can be generalized to higher order constellations by applying proper approximations. In addition, let us assume that the probability of wrong decision at relay is α. The following Lemma specifies the structure of combiner at the destination node.

***Lemma* 6.3:** The optimum combining vectors for a general DF transmission protocol can be written as:

$$\tilde{\mathbf{w}}_1 = \mathbf{H}_1\mathbf{w}_1 \text{ and } \tilde{\mathbf{w}}_2 = \delta\,\mathbf{H}_2\mathbf{w}_2 \tag{6.11}$$

where δ can be approximated as $\delta \approx 1-2\alpha$ when α is small.

Proof: Let SNR_1 and SNR_2 represent the SNR corresponding to the cases of correct and wrong detection at the relay, respectively, i.e.,

$$\text{SNR}_1 = \frac{(d_1+d_2)(d_1^{'}+d_2^{'})}{d_0}, \quad \text{SNR}_1 = \frac{(d_1-d_2)(d_1^{'}-d_2^{'})}{d_0} \tag{6.12}$$

where $d_1 = \gamma_1 \mathbf{w}_1^H \mathbf{H}_1^H \tilde{\mathbf{w}}_1$, $d_2 = \gamma_2 \mathbf{w}_2^H \mathbf{H}_2^H \tilde{\mathbf{w}}_2$, $d_1^{'} = \gamma_1 \tilde{\mathbf{w}}_1^H \mathbf{H}_1 \mathbf{w}_1$, $d_2^{'} = \gamma_2 \mathbf{w}_2^H \mathbf{H}_2 \mathbf{w}_2$ and $d_0 = \tilde{\mathbf{w}}_1^H \tilde{\mathbf{w}}_1 + \tilde{\mathbf{w}}_2^H \tilde{\mathbf{w}}_2$. Now, the average received SNR can be written as:

$$\begin{aligned}&\text{SNR} = (1-\alpha)\text{SNR}_1 + \alpha\,\text{SNR}_2 = \\ &\frac{\begin{array}{l}\gamma_1 \mathbf{w}_1^H \mathbf{H}_1^H \tilde{\mathbf{w}}_1 \tilde{\mathbf{w}}_1^H \mathbf{H}_1 \mathbf{w}_1 + \gamma_2 \mathbf{w}_2^H \mathbf{H}_2^H \tilde{\mathbf{w}}_2 \tilde{\mathbf{w}}_2^H \mathbf{H}_2 \mathbf{w}_2 + \\ (1-2\alpha)\sqrt{\gamma_1\gamma_2}(\mathbf{w}_1^H \mathbf{H}_1^H \tilde{\mathbf{w}}_1 \tilde{\mathbf{w}}_2^H \mathbf{H}_2 \mathbf{w}_2 + \tilde{\mathbf{w}}_1^H \mathbf{H}_1 \mathbf{w}_1 \tilde{\mathbf{w}}_2^H \mathbf{H}_2 \mathbf{w}_2)\end{array}}{\tilde{\mathbf{w}}_1^H \tilde{\mathbf{w}}_1 + \tilde{\mathbf{w}}_2^H \tilde{\mathbf{w}}_2}\end{aligned} \tag{6.13}$$

From (6.13) and considering a system equation similar to (6.8), one can easily find the corresponding equivalent system equation for this scenario as:

$$y = (\tilde{\mathbf{w}}_1^H \; \tilde{\mathbf{w}}_2^H)\mathbf{H}\begin{pmatrix}\mathbf{w}_1\\ \mathbf{w}_2\end{pmatrix}x + (\tilde{\mathbf{w}}_1^H \; \tilde{\mathbf{w}}_2^H)\begin{pmatrix}\mathbf{n}_1\\ \mathbf{n}_2\end{pmatrix} \tag{6.14}$$

where $\mathbf{H}$ is the channel matrix of the equivalent MIMO system and:

$$\mathbf{H} = \begin{pmatrix}\sqrt{\gamma_1}\mathbf{H}_1 & 0\\ 0 & \sqrt{\gamma_2}(1-2\alpha)\mathbf{H}_2\end{pmatrix} \tag{6.15}$$

Therefore, like Lemma 2, the combining weight vectors can be written as (6.11) and this concludes Lemma 3.

■

Although we derived (6.11) for the case of BPSK transmission, it is also valid for 4-QAM transmission. For higher order modulations, however, it is not straight-forward to derive the weight vectors. Note that one can use other estimations to find coefficient δ or find the optimum weigh ad-hoc. In general, the idea is to reduce the weight of the erroneous signal arrived from relay compared to the case when it does not have error.

6.2.3. AF Relay-Assisted MIMO System (Case *C*)

In this part, we study the case of a relay system with AF protocol. The difficulty with this scenario is that the noise in the second time slot is no longer white. In other word, based on (6.1) and (6.2), the system equation in this case can be rewritten as:

$$y=(\tilde{\mathbf{w}}_1^H\ \tilde{\mathbf{w}}_2^H)\begin{pmatrix}\sqrt{\gamma_1}\mathbf{H}_1 & 0\\ 0 & \sqrt{\gamma_2\gamma_G}\mathbf{H}_2\end{pmatrix}\begin{pmatrix}\mathbf{w}_1\\ \mathbf{W}_2^H\mathbf{G}\mathbf{w}_1\end{pmatrix}x+(\tilde{\mathbf{w}}_1^H\ \tilde{\mathbf{w}}_2^H)\begin{pmatrix}\mathbf{n}_1\\ \mathbf{n}_2+\sqrt{\gamma_2}\mathbf{H}_2\mathbf{W}_2^H\mathbf{n}\end{pmatrix} \quad (6.16)$$

The difference between here and DF protocol is that $\mathbf{W}_2$ is the $M\times M$ precoding matrix (instead of vector) at the relay node. Therefore, the SNR equation can not be written as (6.6) and Lemma 1 and 2 can not be directly applied. Instead, we can define a pre-whitening filter to make the noise in the second interval uncorrelated. Note that, this filter is just used in the second time interval. This idea has been addressed in *Theorem* 6.1.

***Theorem* 6.1:** The optimum combining vectors for an AF relay-assisted MIMO system whose model is described by (6.16) can be written as:

$$\tilde{\mathbf{w}}_1=\mathbf{H}_1\mathbf{w}_1 \text{ and } \tilde{\mathbf{w}}_2=\mathbf{D}^{-1/2}\mathbf{U}^H\mathbf{H}_2\mathbf{W}_2^H\mathbf{G}\mathbf{w}_1 \quad (6.17)$$

where $\mathbf{D}$ and $\mathbf{U}$ are the diagonal and unitary matrices resulted from eigenvalue decomposition of autocorrelation matrix of noise matrix $\mathbf{n}_2+\sqrt{\gamma_2}\mathbf{H}_2\mathbf{W}_2^H\mathbf{n}$.

Proof: Assuming $\mathbf{n}$ and $\mathbf{n}_2$ are uncorrelated, the autocorrelation function of the noise in the second time interval can be written as:

$$\begin{aligned}\Lambda&=E\{(\mathbf{n}_2+\sqrt{\gamma_2}\mathbf{H}_2\mathbf{W}_2^H\mathbf{n})(\mathbf{n}_2+\sqrt{\gamma_2}\mathbf{H}_2\mathbf{W}_2^H\mathbf{n})^H\}\\&=\mathbf{I}_M+\gamma_2\mathbf{H}_2\mathbf{W}_2^H\mathbf{W}_2\mathbf{H}_2^H\end{aligned} \quad (6.18)$$

where $\mathbf{I}_M$ denoted for identity matrix of size M. Applying an eigenvalue decomposition over Λ, one can write $\Lambda=\mathbf{U}\mathbf{D}\mathbf{U}^H$ where $\mathbf{U}$ and $\mathbf{D}$ are $N\times N$ unitary and diagonal matrices, respectively.

Defining the pre-whitening filter as $\mathbf{W}_P=\mathbf{D}^{-1/2}\mathbf{U}^H$ and applying it to the received signal in the second time slot will ensure us that the output noise is white. With that, one can apply the results of Lemma 2 directly to derive the structure of combining vectors. The system model can be considered as (6.14) with:

$$\mathbf{H}=\begin{pmatrix}\sqrt{\gamma_1}\mathbf{H}_1 & 0\\ 0 & \sqrt{\gamma_2\gamma_G}\mathbf{D}^{-1/2}\mathbf{U}^H\mathbf{H}_2\end{pmatrix} \quad (6.19)$$

From (6.16) and (6.19), the structure of combining vectors in (6.17) will be directly concluded.

■

In other words, the optimum weight vector in the second time interval is a combination of an MRC vector and a pre-whitening filter. Note also that, by applying pre-whitening filter, the output noise will be white and therefore one can apply (6.7) to calculate instantaneous mutual information. On the other hand, as the combining weight vectors in (6.17) maximize the SNR, these also maximize the instantaneous mutual information and ultimately the system capacity.

6.3. Precoding for Relay-Assisted MIMO Systems

In the previous section, we elaborate the structure of optimum receiver based on MRC approach. In this section, we focus on the optimum transmit and relay structures. In other words, out goal here is to investigate the design of precoding vectors, $\mathbf{w}_1$ and $\mathbf{w}_2$, in (6.1) and (6.2).

We start by simple case of point-to-point MIMO transmission and then generalize the results to the case of relay-assisted MIMO systems.

***Lemma* 6.4:** The optimum precoding vector that maximizes the SNR of an MRC based MIMO system can be written as:

$$\mathbf{w} = \mathbf{u}_{\max}(\mathbf{H}^H\mathbf{H}) \tag{6.20}$$

where $\mathbf{u}_{\max}$ stands for the eigenvector of $\mathbf{H}$ corresponding the maximum eigenvalue.

Proof: The proof is straight forward and comes from the calculated SNR equation in (6.6) after applying MRC weight vector, i.e. $\text{SNR} = \gamma\mathbf{w}^H\mathbf{H}^H\mathbf{H}\mathbf{w}$. To maximize this SNR subject to power constraint over $\mathbf{w}$ similar to (6.4), one should take $\mathbf{w}$ in the direction of the eigenvector of Hermitian matrix $\mathbf{H}^H\mathbf{H}$ associated with $\lambda_{\max}$ where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{H}^H\mathbf{H}$.

■

With this precoding vector and considering the receive combining vector of $\tilde{\mathbf{w}} = \mathbf{H}\mathbf{w}$, the receive SNR of the system can be written as $\text{SNR}=\gamma\,\lambda_{\max}$. This result can be generalized to the case of GMRC directly by Lemma 5.

***Lemma* 6.5:** For a multipoint-to-point MIMO system with GMRC combining vectors at receiver, the optimum precoding vectors are:

$$\mathbf{w}_1 = \mathbf{u}_{\max}(\mathbf{H}_1^H\mathbf{H}_1) \text{ and } \mathbf{w}_2 = \mathbf{u}_{\max}(\mathbf{H}_2^H\mathbf{H}_2) \tag{6.21}$$

Proof: The proof follows from Lemmas 2 and 4.

■

The same conclusion is also valid for the case of DF protocol for relay-assisted MIMO systems. Now, we are ready to revisit the problem of AF protocol. The difficulty with this case is that the optimization of precoding matrix at the relay node, $\mathbf{W}_2$, is not independent of optimization of precoding vector at the source node, $\mathbf{w}_1$. This is because the receive SNR equation resulted in the second transmit interval is:

$$\mathrm{SNR}_2 = \sqrt{\gamma_2 \gamma_G}\, \mathbf{w}_1^H \mathbf{G}^H \mathbf{W}_2 \mathbf{H}_2^H \mathbf{H}_2 \mathbf{W}_2^H \mathbf{G} \mathbf{w}_1 \tag{6.22}$$

and therefore, is a function of both $\mathbf{w}_1$ and $\mathbf{W}_2$. (6.22) comes from plugging (6.17) into original SNR equation in (6.6). Although direct maximization (6.22) for $\mathbf{W}_2$ would be difficult, one can make a clever guess that if we select $\mathbf{W}_2$ such that SNR over source-relay link, $\mathbf{G}$, this can ultimately result in the maximization of the SNR in (6.22) over the entire link from source to destination. Therefore, let assume $\mathbf{W}_2$ can be written as:

$$\mathbf{W}_2 = \mathbf{G}\mathbf{w}_1 \hat{\mathbf{w}}_2^H \tag{6.23}$$

where the part $\mathbf{G}\mathbf{w}_1$ is responsible for maximizing the SNR at relay node while $\hat{\mathbf{w}}_2$ is an independent vector reserved for further optimization of precoding matrix at the relay node. Moreover, for maximizing the SNR over the relay-destination link, using Lemma 4, it is easy to check that $\hat{\mathbf{w}}_2$ turns out to be:

$$\hat{\mathbf{w}}_2 = \mathbf{u}_{\max}(\mathbf{H}_2) \tag{6.24}$$

Now, $\mathbf{w}_1$, the precoding vector at the source node is the only remaining parameter to be selected. However, $\mathbf{w}_1$ appears in both SNR equations in the first and the second time slots, i.e. SNR_1 and SNR_2. Therefore, to optimize $\mathbf{w}_1$, the SNR equation similar to (6.16) should be considered which makes the optimization problem very difficult if not impossible to solve. To solve this problem, we focus on each of the SNR equations, separately. Since in a communication systems, there are usually a number of available relay terminals (rather than just one), we, ultimately, propose the use of relay selection approach to maximize the overall SNR of the system. The following claim is the main result of this section.

***Claim* 6.1:** Amongst all K candidate relay nodes, the best relay node is selected as:

$$i = \arg\max_{i=1..K} \langle \mathbf{u}_{\max}(\mathbf{H}_1), \mathbf{u}_{\max}(\mathbf{G}) \rangle\, \gamma_G \gamma_2 \lambda_{\max}(\mathbf{H}_2^H \mathbf{H}_2) \tag{6.25}$$

if source-destination link is stronger that the source-relay and relay-destination links, or:

$$i = \arg\max_{i=1..K} \langle \mathbf{u}_{\max}(\mathbf{H}_1), \mathbf{u}_{\max}(\mathbf{G}) \rangle \lambda_{\max}(\mathbf{H}_1^H \mathbf{H}_1) + \gamma_G \gamma_2 \lambda_{\max}(\mathbf{H}_2^H \mathbf{H}_2) \tag{6.26}$$

if source-relay and relay-destination links are stronger than source-destination link. Note that $<,>$ stands for inner product.

This selection maximized the SNR of the relay-assisted AF system with high probability. In other words, Theorem 2 articulated that the best relay is the one with strong source-relay and relay-destination links, i.e. large γ_G and γ_2, and when the eigenvector corresponding to the maximum eigenvalue of source-relay link matrix is the closest to $\mathbf{u}_{\max}(\mathbf{H}_1)$.

Proof: Let first, take the receive SNR in the first transmit interval:

$$\mathrm{SNR}_1 = \gamma_1 \mathbf{w}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{w}_1 \tag{6.27}$$

Based on Lemma 6.4, this is clearly maximized if one selects $\mathbf{w}_1 = \mathbf{u}_{\max}(\mathbf{H}_1)$. On the other hand, in the second transmit interval, by plugging $\mathbf{W}_2$ from (6.23) into (6.22), it turns out that the source precoding vector will be responsible for the source-relay portion of the SNR_2. Therefore, to maximize this portion, one should select $\mathbf{w}_1 = \mathbf{u}_{\max}(\mathbf{G})$. These two equations for $\mathbf{w}_1$ are definitely in contrast with each other. The best scenario is that $\mathbf{u}_{\max}(\mathbf{H}_1) = \mathbf{u}_{\max}(\mathbf{G})$. In this case, based on Lemma 4, the overall receive SNR of the system can be written as:

$$\mathrm{SNR} = \lambda_{\max}(\mathbf{H}_1^H \mathbf{H}_1)\ (\gamma_1 + \gamma_G \gamma_2 \lambda_{\max}(\mathbf{H}_2^H \mathbf{H}_2)) \tag{6.28}$$

Now, let assume that $\mathbf{u}_{\max}(\mathbf{H}_1) \neq \mathbf{u}_{\max}(\mathbf{G})$ but there are K available relay nodes in the system. The best relay should be selected such that the overall receive SNR is maximized. The performance degradation appears as a factor of $\langle \mathbf{u}_{\max}(\mathbf{H}_1), \mathbf{u}_{\max}(\mathbf{G}) \rangle$ in the SNR equation. Two extreme cases can be considered.

First case is that the source-destination link is very strong. In this case one should choose the precoding vector in the direction of $\mathbf{u}_{\max}(\mathbf{H}_1)$. Therefore, the SNR loss will be in the second transmit interval (compared to the optimum case in (6.28)) and overall SNR can be expressed as.

$$\mathrm{SNR} = \lambda_{\max}(\mathbf{H}_1^H \mathbf{H}_1)\ \gamma_1 + \langle \mathbf{u}_{\max}(\mathbf{H}_1), \mathbf{u}_{\max}(\mathbf{G}) \rangle\ \gamma_G \gamma_2 \lambda_{\max}(\mathbf{G}^H \mathbf{G})\ \lambda_{\max}(\mathbf{H}_2^H \mathbf{H}_2)) \tag{6.29}$$

Therefore, the best relay node is the one that maximizes the right hand side in (6.25).

Second case is that the relay links are much stronger compared to direct link. The natural selection in this case is to select the source precoding vector in the direction of $\mathbf{u}_{\max}(\mathbf{G})$. In this case, the SNR loss will be in the second transmit interval. The overall receive SNR can also be written as:

$$\text{SNR} = \langle \mathbf{u}_{\max}(\mathbf{H}_1), \mathbf{u}_{\max}(\mathbf{G}) \rangle \lambda_{\max}(\mathbf{H}_1^H \mathbf{H}_1)\ \gamma_1 + \gamma_G \gamma_2 \lambda_{\max}(\mathbf{G}^H \mathbf{G})\ \lambda_{\max}(\mathbf{H}_2^H \mathbf{H}_2)) \quad (6.30)$$

which maximizing this SNR over all possible K relay nodes will result in (6.26) and this concludes the proof of Claim.

■

6.4. Numerical Results

In this section, we study the performance of the proposed MRC-based precoding technique by means of simulation. We consider all source, relay and destination nodes are equipped with two antennas, i.e. M=N=2. We simulate received SNR and average mutual information over these systems.

Figure 6.2 shows received SNR for different setups in Figure 6.1. For the sake of comparison we also show the performance of precoding based equal gain combining in point-to-point MIMO transmission. First we observe that MRC-based precoding (case A) is superior to precoding with equal gain combining in term of received SNR by about 3 dB. On the other hand, in the case of multipoint-to-point transmission (case B), precoding based on MRC can provide us with an additional gain compared to the case of point-to-point transmission. This is mainly due to the availability of an additional path between second transmit and receive node which provides a higher average SNR at the receiver (especially when the link from first transmitter to receive node is poor). Finally, we see that the relay-assisted amplify-and-forward precoding system (case C) provides us with a received SNR comparable to that of ideal multipoint-to-point system. The slight decrease in the received SNR is due to the noise enhancement in the relay node. The same conclusion is also valid for the case of decode-and-forward system as there is a decrease in the received SNR due to the probability of wrong decision at the relay node.

Figure 6.3 also shows that the average mutual information for all the above scenarios. The same conclusions can be drawn for averaged mutual information, i.e. relay assisted MIMO precoding can provide us an average mutual information close to the case when two transmitter send identical information to the relay node.

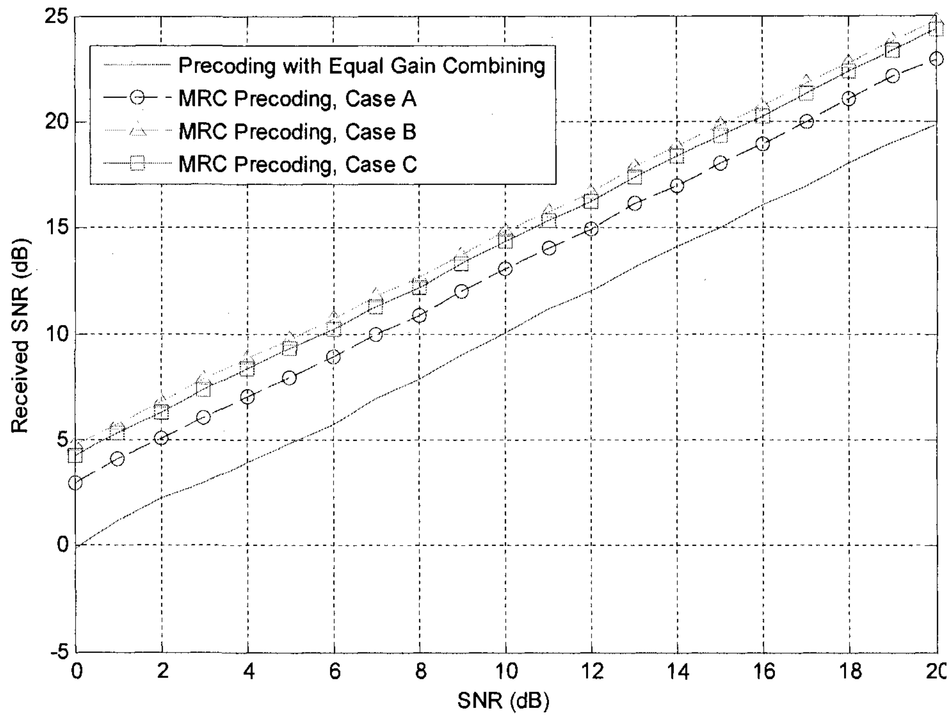


Figure 6.2: Received SNR in different scenarios

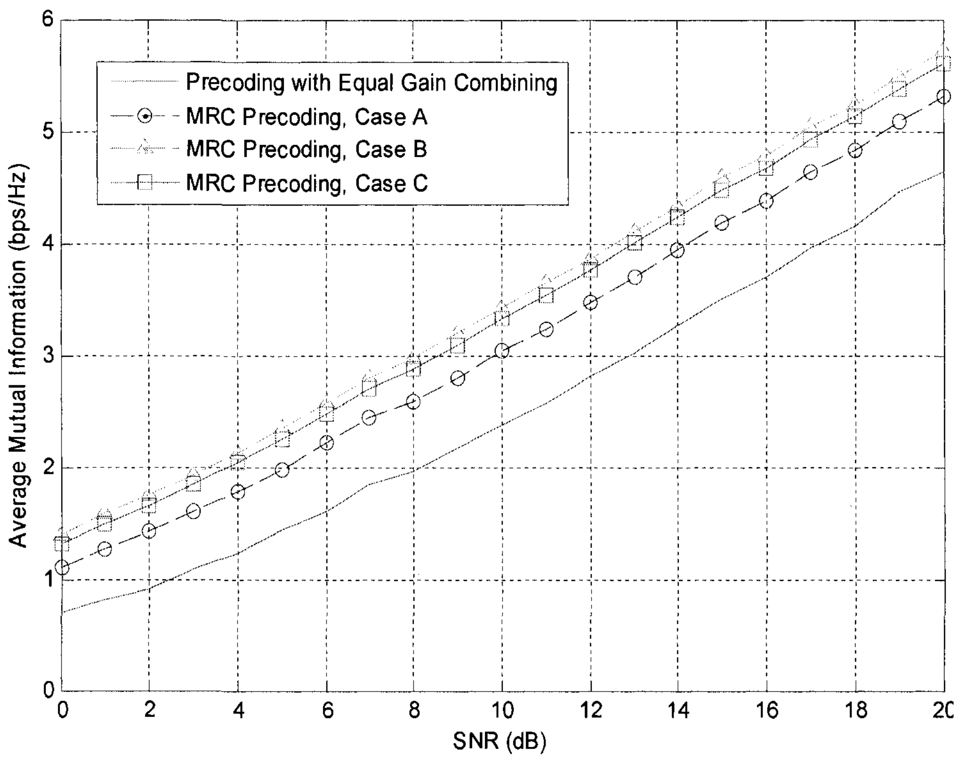


Figure 6.3: Average mutual information in different scenarios

6.5. Chapter Summary

In this chapter, we studied optimal linear transmit and receive strategies for a variety of MIMO systems. Our focus was mainly on relay-assisted MIMO systems. These systems are very attractive from both theoretical and practical point of view. However, there are still many open questions concerning transmission and reception schemes in the field. We built a framework based on well-known MRC scheme at receive (destination) side. As we demonstrated, the optimal combining strategies for both DF and AF relaying protocols can be constructed based on the concept of MRC.

We first derived the optimal linear receive structure for these systems. Next, based on the structure of optimum receivers, we investigated the optimal linear precoding vectors for source and relay nodes. Our results, showed that for the optimum receiver, the optimal transmit precoding strategy for DF protocol is to end the information in the direction of the eigenvector of forward and relay channel matrices associated with strongest eigenvalues. This simple result is, however, not valid in the case of AF relaying protocol. Instead, for AF protocol, we proposed the use of relay selection scheme to facilitate the design of precoders at the source and relay nodes. Finally, different numerical examples proved that the proposed optimal transmission and reception techniques are indeed effective and provide a meaningful gain in term of received SNR and system capacity while maintaining the complexity very low due to linearity.

Chapter 7
Conclusions

7.1. Summary and Concluding Remarks

In this research, the key idea was the use of partial channel information in the design of low-complexity transmission schemes for MIMO and cooperative communications. More specifically, we investigated the problem of designing linear precoders based on partial channel information for MIMO systems in different scenarios. As stated, in MIMO communications one possible way to achieve capacity or performance (bit error rate) equal or close to that of single-input single-output (SISO) channel is to decouple the channel into some independent subchannels by applying linear precoding. Linear precoding is an approach that makes use of the available channel information at the transmitter to remove the inter-relation between MIMO links and hence construct independent parallel channels. While precoding techniques usually assume perfect knowledge of instantaneous channel response at the transmitter, we focused on the designs based on partial channel information due to its practicality in fast fading environments. More precisely, we proposed to investigate channel statistical properties rather than the actual channel time-varying space-frequency responses and use them as partial channel knowledge at the transmitter for developing MIMO precoding schemes in single-user, multi-user and cooperative communications. Our approach in this research was analogical (deductive) meaning that the thesis flow started by considering optimal precoding design for the simple case of point-to-point transmission. While in the big picture, this can be considered as a per-link optimization of transmission scheme. Next, we considered the case of precoding design for multi-user broadcast systems. This is a more general problem as it targets optimization over several links rather than single channel. Therefore, this scheme can be applied on top of the per-link optimization of previous step. The last step toward generalization of our discussion was to consider the case of cooperative relay networks. We argued that in our context, cooperative communications can be considered as the combination of broadcast and multi-access transmissions. Hence, the discussed cooperative communication is composed of two multi-user transmission phases. This key observation shed lights on the analogical relationship between steps three and two. Therefore, an approach like the one discussed in the second step for multi-user systems can be applied to each of these phases while in a more general perspective, it is meaningful to optimize the transmission over two transmission phases in cooperative communications. Based on this roadmap, we divided the structure of the thesis into five main chapters.

In Chapters 2 and 3, we tackled the first step to design precoding techniques that are optimal in the sense of per-link performance. We came up with two precoder designs for point-to-point MIMO systems with underlying frequency-flat and frequency-selective channels based on partial channel information. While we considered three different criteria, namely PEP, MMSE and ergodic capacity, for designing the precoder in the case of frequency-flat channels, it was assumed that we only know the spatial (transmit and receive) correlation matrices at the transmit side. It turned out that in all three cases, the structures of the precoding matrices are the same and they are composed of two unitary and diagonal matrices. They, hence, can be viewed as eigen-beamformers with beams refer to the eigen structure of transmit correlation matrix. The difference is in the entries of diagonal matrices which specify the power to be allocated to each eigen-mode. While selection of spatial correlation matrices appeared a reasonable choice for frequency-flat channels, we showed that in the case of frequency-selective channels, adding another level of partial channel information at transmitter at the cost of slight increase in feedback load is indeed helpful. More precisely, in the case of frequency-selective channel, in addition to spatial correlation matrices, we also added the knowledge of path correlation matrix at the transmit side. Although, the structure of precoder in this case is more complex compared to the case of frequency-flat channel, we showed that one can come up with optimal solution by eigen-decomposition of spatial and path correlation matrices. Nevertheless, the solution can not be generally interpreted as eigen-beamforming over correlation matrices.

Although the derivations in chapters 2 and 3 are appealing for point-to-point (single-link) transmission, it could not provide us with any big picture insight into the problem of transmission optimization in a multi-link network. Therefore, in the second step and in Chapter 4, we switched into the problem of precoder design for point-to-multipoint MIMO systems. More specifically, our target was to design user selection and precoding schemes for a multi-antenna transmit unit given that we just have a small knowledge about users' channels at transmit side. We showed that having a very limited knowledge about channel, transmitter can select the users and allocates power to them in an efficient way. The proposed scheme was a combination of user selection and precoding. A group of best users is selected based on the closeness of their channel vectors to a pre-defined orthonormal vector set without the need of channel information at the transmitter. The second stage, however, assumed availability of channel information of selected users at the base station. This is however a small feedback load as the number of selected users is far less than the actual number of available users. Note that, assumption of partial channel information is a key assumption here and this becomes very important in real broadcast systems where there are a lot of active users available and it is not feasible to acquire their channel information at the transmitter.

In the third step and in Chapter 5, we focused on optimum transmission in cooperative communication systems. Collaborative transmission is an important class of transmission strategies in today communications. It will result in higher diversity degrees and also can provide us with interference mitigation. Diversity and interference mitigation can result in a higher performance and throughput. Cooperative diversity concept, therefore, promises a power-efficient solution for future wireless communications systems to achieve broader coverage and to mitigate channel impairments without the need to use large power at the transmitter. It is, therefore, important to examine the potential behind cooperative communications. We proposed the use of diversity-multiplexing trade-off as a means to evaluate the capabilities of such systems. More specifically, we, first, concentrated on the theoretical question of multiplexing-diversity trade-off that a relay system can offer assuming multi-antenna source (transmit) and destination (receive) and single-antenna relay nodes. The result was very interesting even from a practical point of view because it can give us an idea for selecting the minimum number of relays that should be used in order to optimize diversity-multiplexing trade-off. It turns out that at some point, adding more relay nodes is not beneficial from diversity-multiplexing point of view.

This result is significant yet does not suggest any practical transmission scheme to use in cooperative communication. We, therefore, proposed to apply a distributed BLAST transmission scheme over selected relays in Chapter 5. It has been proved that a BLAST transmission with ML detection can offer the optimum diversity-multiplexing trade-off offered by MIMO channels. Therefore, BLAST is a suitable candidate for applying in the case of relay networks. We argued that ML detection is not indeed practical. Instead, we proposed a near-optimum relay selection scheme in conjunction with successive interference cancellation. We showed that assuming a suitable relay selection scheme such as the one proposed in Chapter 4, it is possible to achieve a considerable performance gain compared to point-to-point transmission without adding large processing overhead at relay nodes.

For the last milestone of this thesis and in Chapter 6, we designed a practical and easy-to-implement maximum ratio combining (MRC) precoding scheme for cooperative relay networks. Compared to Chapter 5, this chapter provided a practical framework for the case when relay nodes also equipped with multi-antenna rather than single-antenna. We proposed suitable transmission and reception strategies for both amplify-and-forward (AF) and decode-and-forward (DF) MIMO relay systems to improve the performance of the system. We showed that by simple linear precoding at combining techniques at source, relay and destination nodes a substantial gain in terms of performance can be achieved.

In general, our study showed that partial knowledge based precoding is a simple but practical technique that is applicable to any transmission system and propagation scenarios. The selection of type of partial knowledge and target performance measure is of vital importance for its effectiveness. In practice, successful implementation of partial knowledge based precoding in a wireless network is dependent to many parameters such as target performance measure (throughput, performance, etc.), propagation environment (fast or slow fading, frequency-flat or selective, etc.), acceptable feedback load, tolerable complexity at each network node (centralized or distributed) that should be carefully addressed in the design issues.

7.2 Suggestions for Further Studies

Our main focus in this thesis was on the use of partial channel information in the design of transmission techniques for MIMO communications. The term "Partial channel information" is a general term and covers a wide range of channel parameters and statistics. We, however, touched some certain classes of partial channel information such as spatial and path correlation matrices. Advantages and disadvantages of different partial knowledge based transmission techniques in terms of performance, complexity and feedback load are laid in the type of the partial knowledge they use. It is therefore of great importance to select the type of channel information to be acquired at the transmitter carefully; especially when the system becomes more complex (i.e. moving from simple point-to-point transmission to point-to-multipoint and cooperative systems). To continue this research in the future, we propose a comprehensive study on other types of partial channel information that can be acquired and utilized at transmit side to increase the overall system performance.

On the other hand, throughout this study we did not touch other performance measures such as outage probability. Therefore, a study on different types of channel information and comparing their advantages and disadvantages over several system performance measures becomes of considerable importance. To make the later point clearer, let consider the example of spatial correlation matrices. It is well known that spatial correlation degrades the capacity (throughput) of a MIMO system significantly. While on the other hand, its effect on the bit error rate of the system is less pronounced. Therefore, for transmission schemes that target throughput such as spatial multiplexing, it is vital to have spatial correlation information at the transmit side while it is not the case in a space-time coded system that targets higher performances rather than capacity. Hence, the effectiveness of selected partial knowledge defers from one performance measure to another and one should pick up proper channel parameter or statistic to improve any specific performance measure.

A through study on these two areas can ultimately shed light on proper selection of partial channel information in different propagation scenarios and/or for different performance measures.

In the case of point-to-multipoint transmission, we discussed a zero-forcing scheme with small complexity and feedback load that is asymptotically optimum in term of sum rate. Indeed, there are other transmission schemes such as opportunistic beamforming that benefit from high user population to achieve optimality in term of sum rate (or other performance measures). Although very appealing when there is considerable number of user terminals available in the system, these schemes are not suitable for low populated systems. A significant study is to look for transmission schemes that are near-optimum while the number of users is low. On the other hand, the proposed scheme requires the selected users to feedback their instantaneous channel information to the base station. It would not be practical in fast fading channels especially in a multi-user system when the number of selected users is high. Therefore, another interesting subject of study in this area is to design codebooks or quantization techniques to feedback the selected users' channel information to the base station in order to reduce feedback load.

In the context of cooperative communications, we first briefly addressed the problem of multiplexing-diversity trade-off for a specific transmission strategy in single-antenna relay networks. It is currently one of active areas of research in information theory. Solving the problem of multiplexing-diversity in a more general and rigorous framework when all terminals are equipped with multi-antenna is an open problem needs to be addressed. Also, from a theoretical point of view, there are still several open problems in finding the inherent capacity, diversity and multiplexing gain offered by these systems especially when terminals are equipped with multi-antenna. While these questions are of great importance from a theoretical point of view, the problem of transmission in multi-antenna AF and DF relay networks is an interesting area of study for future work. We addressed the problem of optimal precoding for AF and DF strategies while in the case of AF we did not come up with a pure precoding approach and actually combined precoding with relay selection. A non-trivial study is to design precoding schemes for AF systems without the need of relay selection. There are also very noteworthy problems on the design of partial knowledge based precoding for cooperative systems. For instance, from a practical point of view, an important yet open problem in this field is the transmission designs that need minimum channel feedback at the transmit node. Note that for example in a cellular system where some mobile terminals help specific user terminal (source) to transmit its information, it is not very convenient to acquire all relay channel information. On the other hand, it is far from practical to inform the source user terminal of all channel information. Therefore, it is highly desirable to devise cooperative transmission schemes that need no or minimum relay channels information.

We also did not touch the problem of *centralized* versus *distributed* signal processing in cooperative communications. This problem is however more related to networking aspect of cooperative systems. By *centralized* we mean that a base station makes all the decisions on relay selection, power allocation and precoding. While *distributed* refers to the case that all nodes participate in decision making. While *distributed* approach would enjoy less complexity and feedback load over *centralized* strategy, its performance is limited due to limited network information at each node. An interesting analysis is to determine whether distributing signal processing load through the whole network is helpful or not. If yes, the question is how transmit and relay nodes can share the load of processing amongst each other.

References:

[1] H. Sampath, P. Stoica, A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.,* vol. 49, no. 12, pp. 2198-2206, Dec. 2001.

[2] H. Sampath, "Linear precoder and decoder designs for multiple-input-multiple-output wireless systems", *PhD Dissertation,* Stanford University, Stanford, CA, Apr. 2001.

[3] H. Sampath and A. Paulraj, "Linear precoding for space-time Coded systems with known fading correlations," *35th Asilomar Conf. Sig., Syst., Comp.*, vol. 1, pp. 246-251, Pacific Grove, CA, Nov. 4-7, 2001.

[4] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: models and performance prediction", *IEEE Trans. Commun.,* vol. 50, no. 12, pp. 1926-1934, Dec. 2002.

[5] E. Yoon, J. Hansen, and A. Paulraj, "Space-frequency precoding for an OFDM based system exploiting frequency and path correlation," *IEEE GLOBECOM 2004*, vol. 2, pp. 859-863, Vancouver, BC, Sep. 24-28, 2004.

[6] H. Sampath and A. Paulraj, "Linear precoding for space-time coded systems with known fading correlations," *IEEE Commun. Letters,* vol. 6, no. 6, pp. 239-241, June 2002.

[7] E. A. Jorswieck and H. Boche, "Optimal transmission strategies and impact of correlation in multiantenna systems with different types of channel state information," *IEEE Trans. Sig. Proc.,* vol. 52, no. 12, pp. 3440-3453, Dec. 2004.

[8] S. A. Jafar, S. Vishwanath, and A. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," *IEEE Int. Conf. Commun. 2001*, vol. 7, pp. 2266–2270, St. Petersburg, June 11-14, 2001.

[9] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Inform. Theory,* vol. 48, no. 3, pp. 611-627, Mar. 2002.

[10] Y. Zhao, R. Adve and T. J. Lim, "Precoding of orthogonal STBC with channel covariance feedback for minimum error probability," *IEEE PIMRC 2004*, vol. 1, pp. 503-507, Barcelona, Spain, Sept. 5-8, 2004.

[11] A. Hjorungnes, D. Gesbert, "Minimum exact SER precoding of orthogonal space-time block codes for correlated MIMO channels," *IEEE GLOBECOM 2004*, vol. 1, pp. 111–115, Dallas, TX, Nov. 29- Dec. 3, 2004.

[12] A. Hjorungnes, D. Gesbert, J. Akhtar, "Precoding of space-time block coded signals for joint transmit-receive correlated MIMO channels", submitted to *IEEE Trans. Wireless Commun.*, http://www.unik.no/~arehj/publications/letterversion.pdf.

[13] M. Kiessling, J. Speidel, I. Viering, and M. Reinhardt, "A closed-form bound on correlated MIMO channel capacity", *IEEE Vehic. Technol. Conf., Fall 2002*, vol. 2, pp. 859-863, Vancouver, BC, Sep. 24-28, 2002.

[14] E. A. Jorswieck and H. Boche, "Channel capacity and capacity-range of beamforming in MIMO wireless systems under correlated fading with covariance feedback," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1543-1553, Sep. 2004.

[15] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.

[16] T. A. Cover and J. M. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[17] C. Aitken, *Determinants and Matrices*. Edinburgh/London: Oliver and Boyd, 1964.

[18] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[19] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal design for space-time linear precoders and decoders," *IEEE Trans. Sig. Proc.*, vol. 50, no. 5, pp. 1051- 1064, May 2002..

[20] D. Chizhik, G. J. Foschini, M. J. Gans, R. A. Valenzuela, "Keyholes, correlations, and capacities of multielement transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 361-368, Apr. 2002.

[21] H. R. Bahrami, T. Le-Ngoc, A. M. Nasri, S. H Jamali, "Precoder design based on correlation matrices for MIMO systems," *IEEE Int. Conf. Commun. 2005,* May 2005, Seoul, Korea.

[22] H. R. Bahrami and T. Le-Ngoc, "Precoder design based on correlation matrices for MIMO systems", *IEEE Trans. Wireless Communications,* vol. 5, no. 12, pp. 3579-3587, Dec. 2006.

[23] R. A. Horn, and C. R. Johnson, *Matrix Analysis*. New York: Cambridge Press, 1985.

[24] D. A. Gore, R. U. Nabar, and A. J. Paulraj, "Selecting an optimal Set of transmit antennas for a low rank matrix channel", *IEEE Int. Conf. Acoust., Speech and Sig. Proc. 2000*, Istanbul, Turkey, June 2000.

[25] R. U. Nabar, D. A. Gore, and A. J. Paulraj, "Optimal selection and use of transmit antennas in wireless systems", *Int. Conf. Telecommunications*, Acapulco, Mexico, May, 2000.

[26] A. F. Molisch, "MIMO systems with antenna selection - An overview", *Radio and Wireless Conf., 2003. RAWCON 03*, pp 167-170, Aug. 2003.

[27] A. F. Molisch, M. Z. Win, and J. H. Winters, "Capacity of MIMO systems with antenna selection", *IEEE Int. Conf. Commun. 2001*, vol. 2, pp. 570-574, June 2001.

[28] M. Gharavi-Alkhansari, and A. B. Gershman, "Fast antenna subset selection in MIMO systems", *IEEE Trans. Sig. Proc.*, vol. 52, no. 2, pp. 339-347, Feb. 2004.

[29] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems", *IEEE Int. Conf. Acoust., Speech and Sig. Proc.*, vol. 3, pp. 2857-2860, May 2002.

[30] Y. S. Choi, A. F. Molisch, M. Z. Win, and J. H. Winters, "Fast algorithms for antenna selection in MIMO systems", *IEEE Vehic. Technol. Conf., Fall 2003*, Orlando, FL, Oct. 2003.

[31] H. Zhang and H. Dai, "Fast transmit antenna selection algorithms for MIMO systems with fading correlation," *IEEE Vehic. Technol. Conf., Fall 2004*, Los Angeles, CA, Sept. 2004.

[32] D. A. Gore, R. W. Health, Jr. and A. J. Paulraj, "Statistical antenna selection for spatial multiplexing systems," *IEEE Int. Conf. Commun. 2002,* vol. 1, pp. 450 -454, New York City, 28 April- 2, May, 2002.

[33] R. Narasimhan, "Spatial multiplexing with transmit antenna and constellation selection for correlated MIMO fading channels", *IEEE Trans. Sig. Proc.,* vol. 51, no. 11, pp. 2829-2838, Nov 2003.

[34] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 572–584, Sept.1979.

[35] J. N. Laneman, D. N. C. Tse, G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior", *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3062-3080, December 2004.

[36] Rohit U. Nabar, Helmut Bölcskei, Felix W. Kneubuhler, "Fading relay channels: Performance limits and space-time signal design," *IEEE J. Selected Areas in Commun.*, vol. 22, no. 6, pp.1099-1109, Aug 2004.

[37] A. Stefanov and E. Erkip, "Cooperative coding for wireless networks," *IEEE Trans. Commun.*, vol. 52, pp.1470-1476, Sept. 2004.

[38] J. C. H. Lin, A. Stefanov, "Coded cooperation for OFDM systems", *WirelessCom 2005*, Sheraton Maui Resort, Maui, Hawaii, June 13-16, 2005.

[39] A. Sendonaris, E. Erkip and B. Aazhang, "User cooperation diversity, Parts I & II," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927-1948, Nov.2003.

[40] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw-Hill, 2001.

[41] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[42] D. Chizhik, F. Rashid-Farrokhi, J. Ling, and A. Lozano, "Effect of antenna separation on the capacity of BLAST in correlated channels," *IEEE Comm. Letters*, vol. 4, no. 11, pp. 337-339, Nov. 2000.

[43] K. I. Pedersen, J. B. Andersen, J. P. Lermoal and P. Mogensen, "A stochastic multiple-input-multiple-output radio channel model for evaluation of space-time coding algorithms," *IEEE Vehic. Technol. Conf., Fall 2000*, vol. 2, pp. 893-897, Boston, MA, Sep. 24-28, 2000.

[44] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[45] E. Bonek, H. Ozcelik, M. Herdin, W. Weichselberger, and J. Wallace, "Deficiencies of a popular stochastic MIMO radio channel model," *Int. Symp. on Wireless Personal Multimedia Communications, WPMC 2003*, Yokosuka, Japan, Oct. 19-22, 2003.

[46] J. Muirhead, *Aspects of Multivariate Statistical Theory.* New York: Wiley, 1982.

[47] L. Schumacher, K. I. Pedersen, P. E. Mogensen, "From antenna spacing to theoretical capacities – Guidelines for simulating MIMO systems," *IEEE PIMRC 2002*, vol. 2, pp. 587-592, Portugal, Sep. 2002.

[48] J. Salz and J.H. Winters, "Effect of fading correlation on adaptive arrays in digital wireless communications," *IEEE Vehic. Technol. Conf., Spring 1993*, vol.3, pp.1768–1774, NJ, May 1993.

[49] K. Zhang, Z. Song, Y. L. Guan, "Multicarrier channels with non-identical subcarrier fading distributions – Analysis and Adaptation," *IEEE Commun. Letters,* vol. 8, no. 6, pp. 368-370, June 2004.

[50] H. Ozcelik, N. Czink, E. Bonek, "What makes a good MIMO channel model?," *IEEE Vehic. Technol. Conf., Spring 2005*, vol. 1, pp. 156-160, Stockholm, Sweden, May 30-June 1, 2005.

[51] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communications," *IEEE Trans. Commun.,* vol. 46, no. 3, pp. 357-366, Mar. 1998.

[52] K. Werner, M. Jansson, P. Stoica, "On estimation of correlation matrices with Kronecker product structure," submitted to *IEEE Trans. Sig. Proc.*, 2006.

[53] I. E. Telatar, "Capacity of multi-antenna Gaussian channel," *Tech. Rep. Bell Labs*, 1995.

[54] V. Kafedziski, "Capacity of frequency selective fading channels with side information," *32nd Asilomar Conf. Signals, Systems and Computers*, vol. 2, pp. 1758–1762, 1998.

[55] K. Liu, T. Kadous, A. M. Sayeed, "Orthogonal time–frequency signaling over doubly dispersive channels," *IEEE Trans. Inform. Theory,* vol. 50, no. 1, pp. 2583-2603, Nov. 2004.

[56] R. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge: Cambridge University Press, 2004.

[57] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[58] J. Akhtar and D. Gesbert, "Spatial multiplexing over correlated MIMO channels with a closed-form precoder," *IEEE Trans. Wireless Commun.*, no. 4, vol. 5, pp.2400–2409, Sep. 2005.

[59] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[60] M. Costa, "Writing on dirty paper," *IEEE. Trans. Inform. Theory*, vol 29, no. 3, pp. 439–441, May 1983.

[61] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates and sum rate capacity of Gaussian MIMO broadcast channel," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[62] A. Goldsmith, S. A. Jafar,N. Jindal, and S.Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 5, pp. 684–702, Jun. 2003.

[63] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channels with partial side information," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 506–522, Feb. 2005.

[64] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and downlink-uplink duality," *IEEE Trans. Inform. theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[65] W. Yu and J. M. Cioffi, "Sum capacity of vector Gaussian broadcast channels, *IEEE Trans. Inform. Theory*, vol. 50, no. 9., pp. 1875-1892, Sep. 2004.

[66] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," *IEEE Int. Conf. Commun. 2005*, vol. 1, pp. 542-546, Seoul, May 2005.

[67] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *37th Asilomar Conf. on Signals, Systems and Computers*, vol. 1, pp. 958-962, Pacific Grove, CA, Nov. 2003.

[68] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Info. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[69] P. Gupta and P. R Kumar, "The capacity of wireless networks," *IEEE Trans. Info. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[70] A. Sendonaris, E. Erkip, and B. Aaazhang, "User cooperation diversity – Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939-1948, Nov. 2003.

[71] D. N. C. Tse, P. Viswanath, and L. Zheng; "Diversity-multiplexing tradeoff in multiple access channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 1859–1874, Sep. 2004.

[72] J. N. Laneman and G. W. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.

[73] A. Bletsas, A. Khisti, D. P. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659-672, Mar. 2006.

[74] A. Bletsas, *Intelligent antenna sharing in cooperative diversity wireless* networks, Ph.D. dissertation, MIT, 2005.

[75] K. Azarian, H. El Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4152–4172, Dec. 2005.

[76] K. Azarian, *Outage limitted cooperativechannels: Protocoles and analysis*, Ph.D. dissertation, The Ohio State University, 2006.

[77] M. Yuksel and E. Erkip, "Cooperative wireless systems: A diversity multiplexing tradeoff perspective," *IEEE Trans. Infrom. Theory, submitted.*

[78] M. Gastpar and M. Vetterli, "On the asymptotic capacity of Gaussian relay netwroks," *Int. Symp. Inform. Theory*, p. 195, Lausanne, Switzerland, Jul. 2002.

[79] A. F. Dana and B. Hassibi, "On the power efficiency of sensory and ad hoc wireless networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 2890–2914, Jul. 2006.

[80] F. H. P. Fitzek and M. D. Katz, *Cooperation in Wireless Networks: Principles and Applications*, Springer, 2006.

[81] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: The relay case," *IEEE INFOCOM*, vol. 3, New York, NY, June 2002, pp. 1577-1586.

[82] A. Sendonaris, E. Erkip, and B. Aaazhang, "User cooperation diversity – Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939-1948, Nov. 2003.

[83] H. Bölcskei, R. U. Nabar, O. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no.6, pp. 1433–1444, June 2006.

[84] S. Oveis Gharan, A. Bayesteh, and A. K. Khandani, "Asymptotic analysis of amplify and forward relaying in a parallel MIMO relay network," Submitted to *IEEE Trans. Inform. Theory*, March 2007.

[85] A. Wittneben and B. Rankov, "Impact of cooperative relays on the capacity of rank-deficient MIMO channels," *12th IST Summin on Mobile Wireless Communications*, Aveiro, Portugal, June 2003, pp. 421-425.

[86] H. R. Bahrami and T. Le-Ngoc, "Eigen structures of MIMO fading channel correlation matrices and optimum linear precoding designs for maximum ergodic capacity," *EURASIP J. Adv. Sig. Proc.*, vol. 2007, Article ID 29749, 9 pages, 2007.

[87] H. R. Bahrami, T. Le-Ngoc, "MMSE-based MIMO precoder using partial channel knowledge," *IEEE International Conf. on Wireless and Mobile Computing, Networking and Communications 2005, WIMOB05,* vol. 1, pp. 211-216, Montreal, Aug. 22-24, 2005.

[88] H. R. Bahrami, T. Le-Ngoc, "MIMO precoding for frequency-selective fading channels based on spatial and path correlation," *IEEE Trans. Vehicular Technology, to appear.*

[89] H. R. Bahrami, T. Le-Ngoc, "MIMO precoder design based on spatial and path correlation information for frequency-selective channels," *IEEE GLOBECOM 2006,* San Francisco, Nov. 2006.

[90] H. R. Bahrami, T. Le-Ngoc, "MIMO precoders using spatial and path correlations for multipath fading channels," *IEEE Int. Conf. Commun. 2007, ICC07,* pp. 4351-4356, Glasgow, June 24-28, 2007.

[91] H. R. Bahrami, T. Le-Ngoc, "MIMO precoding structures for frequency-flat and frequency-selective fading channels," *IEEE International Conf. on Communications and Electronics, 2006,* pp. 193-197, Hanoi, Oct. 10-11, 2006.

[92] H. R. Bahrami, T. Le-Ngoc, "A capacity achieving precoding scheme based on partial channel information for broadcast MIMO systems," *IEEE Wireless Communications and Networking Conf. 2007, WCNC07,* pp. 1045-1050, Hong Kong, Mar. 11-15, 2007.

[93] H. R. Bahrami, T. Le-Ngoc, "User selection and precoding based on partial channel information for broadcast MIMO systems," *IEEE J. on Select. Areas Commun., under review.*

[94] H. R. Bahrami, T. Le-Ngoc, "Relay selection and distributed BLAST in MIMO relay systems," *IEEE GLOBECOM 2007,* pp. 3442-3446, Washington DC, Nov. 26-30, 2007.

[95] H. R. Bahrami, T. Le-Ngoc, "Maximum ratio combining precoding for multi-antenna relay systems," *IEEE Int. Conf. Commun. 2008, ICC08,* Beijing, May 19-23, 2008.

[96] J. Kermoal, L. Schumacher, K. Pederson, P. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE J. Select. Areas Commun.,* vol. 20, no. 6, pp. 1211-1226, Aug. 2002.

[97] K. Yu, M. Bengtsson, B. Ottertsten, D. McNamara, P. Karlsson, and M. Beach, "Second order statistics of NLOS indoor MIMO channels based on 5.2 GHz measurements," *IEEE GLOBECOM 2001,* vol. 1, pp. 156-160, Vancouver, BC, Nov. 22-25, 2001.

[98] D. Bliss, A. Chan, and N. Chang, "MIMO wireless communication channel phenomenology," *IEEE Trans. Antenna Propagat.*, vol. 52, no. 8, pp. 2073-2082, Aug. 2004.

[99] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H. Vincent Poor, *MIMO Wireless Communications*. Cambridge, UK: Cambridge University Press, 2007.

[100] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632-2639, Sep. 2001.

[101] S. Venkatesan, s. Simon, and R. Valenzuela, "Capacity of a Gaussian MIMO channel with nonzero mean," *IEEE Vehic. Technol. conf., Fall 2003*, vol. 3, pp. 1767-1771, Oct. 6-9, 2003.

[102] E. Jorswieck, A. Sezgin, H. Boche and E. Costa, "Optimal transmit strategies in MIMO Ricean channels with MMSE receiver," *IEEE Vehic. Technol. conf., Fall 2004*, vol. 5, pp. 3787-1791, Sep. 26-29, 2004.

[103] M. Vu and A. Paulraj, " Optimal linear precoders for MIMO wireless correlated channels with non-zero mean in space-time coded systems," *IEEE Trans. Sig. Proc.*, vol. 54, no. 6, pp. 2318-2322, Jun. 2006.

[104] S. Zhou and G. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel mean feedback," *IEEE Trans. Sig. Proc.*, vol. 50, no. 10, pp. 2599-2613, Oct. 2002.

[105] S. Zhou and G. Giannakis, "Optimal transmitter eigen-beamforming and space-time block coding based on channel correlations," *IEEE Trans. Inform. Theory*, vol. 49, no. 7, pp. 1673-1690, Jul. 2003.

[106] M. Vu and A. Paulraj, "Capacity optimization of Rician correlated MIMO wireless channels," *39th Asilomar Conf. Sig. Sys. and Comp.*, pp. 133-138, Oct. 18-Nov 1, 2005.

[107] A. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Sig. Proc.*, vol. 50, no. 10, pp. 2563-2579, Oct. 2002.

[108] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90-100, Jan. 2006.

[109] M. Vu and A. Paulraj, "Optimum space-time transmission for a high *K* factor wireless channel with partial channel knowledge," *Wiley J. Wireless Commun. Mobile Computing*, vol. 4, pp. 807-816, Nov. 2004.

[110] R. W. Heath Jr. and A. Paulraj, "Switching between diversity and multiplexing in MIMO systems," *IEEE Trans. Commun.*, vol. 53, no. 6, pp. 962-968, Jun. 2005.

[111] R. Fischer, C. Stierstorfer, and J. Huber, "Precoding for point-to-multipoint transmission over MIMO ISI channels," *Zurich Sem. Commun.*, pp. 208-211, Zurich, Feb. 2004.

[112] A. Bayesteh and Amir K. Khandani, "How much feedback is required in MIMO broadcast channels?," submitted to *IEEE Trans. Inform. Theory*, August 2007.

[113] A. Bayesteh and Amir K. Khandani, "On the user selection for MIMO broadcast channels", submitted to *IEEE Trans. Inform. Theory*, Revised August 2007.

[114] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs. Technical Journal,* vol. 1, no. 2, pp. 41-59, Oct. 1996.