ON ADAPTIVE TRANSMISSION, SIGNAL DETECTION AND CHANNEL

ESTIMATION FOR MULTIPLE ANTENNA SYSTEMS

A Dissertation

by

YONGZHE XIE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2004

Major Subject: Electrical Engineering

ON ADAPTIVE TRANSMISSION, SIGNAL DETECTION AND CHANNEL

ESTIMATION FOR MULTIPLE ANTENNA SYSTEMS

A Dissertation

by

YONGZHE XIE

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

_____
Costas N. Georghiades
(Chair of Committee)

_____
Krishna R. Narayanan
(Member)

_____
Garng M. Huang
(Member)

_____
Graham D. Allen
(Member)

_____
Chanan Singh
(Head of Department)

August 2004

Major Subject: Electrical Engineering

ABSTRACT

On Adaptive Transmission, Signal Detection and Channel Estimation for Multiple
Antenna Systems. (August 2004)
Yongzhe Xie, B. S., Shanghai Jiaotong University;
M. Eng, The National University of Singapore
Chair of Advisory Committee: Dr. Costas N. Georghiades

This research concerns analysis of system capacity, development of adaptive transmission schemes with known channel state information at the transmitter (CSIT) and design of new signal detection and channel estimation schemes with low complexity in some multiple antenna systems. We first analyze the sum-rate capacity of the downlink of a cellular system with multiple transmit antennas and multiple receive antennas assuming perfect CSIT. We evaluate the ergodic sum-rate capacity and show how the sum-rate capacity increases as the number of users and the number of receive antennas increases. We develop upper and lower bounds on the sum-rate capacity and study various adaptive MIMO schemes to achieve, or approach, the sum-rate capacity. Next, we study the minimum outage probability transmission schemes in a multiple-input-single-output (MISO) flat fading channel assuming partial CSIT. Considering two special cases: the mean feedback and the covariance feedback, we derive the optimum spatial transmission directions and show that the associated optimum power allocation scheme, which minimizes the outage probability, is closely related to the target rate and the accuracy of the CSIT. Since CSIT is obtained at the cost of feedback bandwidth, we also consider optimal allocation of bandwidth between the data channel and the feedback channel in order to maximize the average throughput of the data channel in MISO, flat fading, frequency division duplex (FDD) systems.

We show that beamforming based on feedback CSI can achieve an average rate larger than the capacity without CSIT under a wide range of mobility conditions. We next study a SAGE-aided List-BLAST detection scheme for MIMO systems which can achieve performance close to that of the maximum-likelihood detector with low complexity. Finally, we apply the EM and SAGE algorithms in channel estimation for OFDM systems with multiple transmit antennas and compare them with a recently proposed least-squares based estimation algorithm. The EM and SAGE algorithms partition the problem of estimating a multi-input channel into independent channel estimation for each transmit-receive antenna pair, therefore avoiding the matrix inversion encountered in the joint least-squares estimation.

To Mom, Dad and Lan

ACKNOWLEDGMENTS

I am deeply indebted to my advisor, Prof. Costas N. Georghiades, who has provided technical guidance, enlightening ideas and insights throughout my Ph.D program. He also provided generous financial support during the last four year. His encouragement, trust and understanding helped me endure many hard times.

I am very thankful to Dr. Kamyar Rohani of Motorola Labs, who unfortunately passed away in the summer of 2003. Kamyar initiated and supervised our Motorola UPR project on adaptive transmission for multiple antenna systems, which forms most of this dissertation. He provided many useful technical references and insightful suggestions. His dedication to work will always inspire me. I am also grateful to Eric Schorman who succeeded Kamyar in supervising the UPR project. I would like to thank Motorola for funding me during the last three years of my Ph.D. studies.

I am very thankful to Prof. Krisha R. Narayanan, Prof. Xiaodong Wang and Prof. Scott Miller for all the excellent courses I took from them. Prof. Don Allen's course on Linear Algebra and Matrix Theory also helped me in my study. I would like to thank Prof. A. Araposthathis for help in proving an important Lemma in Chapter IV. I would like to thank Mr. Qiang Li for very helpful discussions regarding Chapter VI. Dr. Wei Yu also helped me much by answering some of my questions regarding his papers on MIMO Broadcast channels.

Discussions with many of my other colleagues, Jing (Tiffany) Li, Wurong Yu, Angelos D. Liveris, Beminihennadige Janath Peiris, Jun Zheng, Wenyan He, Samul Chen, Doan Ngoc Dung, Vivek Gulati, Zigang Yang, Ben Lu, Yan Liu, Yan Wang, Yu (Frank) Zhang, have also been helpful.

I am indebted to my wife, Lan Qiang, for all her support and understanding during my Ph.D study. I dedicate this work to Lan, my parents, Ruihe Xie and

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE   Page

CHAPTER I

INTRODUCTION

The use of multiple transmit/receive antennas has emerged as a promising solution for high data rate communication over wireless channels. The resulting multiple antenna system can provide crucial spatial diversity and additional "degrees of freedom" which, if appropriately exploited, can yield significant capacity gains [1, 2].

This work sets two goals in the research of multiple antenna wireless systems. One is to analyze the capacity of some multiple antenna systems and develop adaptive transmission schemes with emphasis on exploring channel side information at the transmitter. The other is to introduce new signal detection and channel estimation schemes with low complexity. More specifically, the dissertation has studied the ergodic sum-rate capacity of a multiple input and multiple output (MIMO) broadcast system and some candidate adaptive transmission schemes assuming perfect channel state information at the transmitter side (CSIT), minimum outage probability transmission in a multiple input and single output (MISO) fading channel with partial CSIT, optimal bandwidth allocation in a FDD system, and efficient EM-type signal detection and channel estimation algorithms for multiple antenna fading channels with application to cellular systems.

A. Dissertation Outline

Chapter II introduces some background knowledge on capacity analysis, and transmission and detection schemes for multiple antenna systems.

Chapter III analyzes the sum-rate capacity of the downlink of a cellular sys-

---

This dissertation follows the style of *IEEE Transactions on Automatic Control.*

tem with multiple transmit antennas and multiple receive antennas assuming perfect CSIT. Modelling the downlink as a flat fading multiple-input-multiple-output (MIMO) broadcast channel (MIMO-BC), we evaluate the ergodic sum-rate capacity using the duality between a MIMO multiple access channel (MIMO-MAC) and a MIMO-BC. We show how the sum-rate capacity increases as the number of users and the number of receive antennas increase. We also develop upper and lower bounds on the sum-rate capacity and study various adaptive MIMO schemes to achieve or approach the sum-rate capacity. Sub-optimal transmission schemes, such as ranked known interference cancellation based on channel matrix triangulation and zero-forcing beamforming based on channel matrix inversion are shown to be able to achieve close to capacity performance.

In Chapter IV, we consider transmission schemes assuming partial CSIT, since perfect channel state information can be too optimistic in practice. We derive the minimum outage probability transmission schemes in a multiple-input-single-output (MISO) flat fading channel for two special cases: the mean feedback case where the CSIT and the actual channel state are jointly Gaussian, and the covariance feedback case where only the spatial covariance matrix of the channel states is known at the transmitter. In the case of mean feedback, the optimal transmission strategy is proven to be transmitting several independent data streams in the direction of the channel mean vector and its orthogonal directions. In contrast to the case of maximizing the ergodic capacity, the optimum power allocation scheme which minimizes outage probability is closely related to the target rate. For both mean and covariance feedback, we show that it is more desirable to spread the power over all transmission directions than beamforming to a single direction for sufficiently small target rates.

In Chapter V, we study the joint optimization of the forward data channel and the feedback channel in terms of bandwidth allocation in order to maximize the average

throughput of the data channel in a MISO frequency division duplex (FDD) system. In FDD systems, CSI is usually estimated by the receiver and then fed back to the transmitter through a reliable link, which inevitably requires additional bandwidth and power. If one views bandwidth and power as common resources that can be shared by the data and feedback channels, the question is whether the increased capacity is worth the penalty paid for it. We consider two models of the partial CSIT: the noisy CSIT which is jointly Gaussian distributed with the actual channel state, and the quantized CSIT. In the first model, we use distortion rate theory to relate the CSIT accuracy to the feedback bandwidth. In the second model, we derive a lower bound on the achievable rate of the data channel based on the ensemble of a set of random quantization codebooks. We show that in the MISO flat fading channel case, beamforming based on feedback CSI can achieve an average rate larger than the capacity without CSIT, under a wide range of mobility conditions.

Chapter VI proposes a Space Alternating Generalized Expectation-Maximization (SAGE) aided List-BLAST detection scheme, which can achieve performance close to that of the maximum-likelihood detector with low computational complexity. The SAGE algorithm searches for the ML solution iteratively by resolving the interference among signals from different transmit antennas. To improve the probability of convergence to the ML solution, multiple initial points are used. The List-BLAST algorithm, which exhausts the constellation points in the first layers of a BLAST detection scheme, is shown to be an excellent way to generate initial points. The complexity of the proposed detection scheme is compared with that of the sphere detection scheme, and it is shown to have a number of implementation advantages.

In Chapter VII, we study channel estimation for an orthogonal frequency division multiplexing (OFDM) system with multiple transmit antennas in a frequency selective fading channel. We propose the EM and the SAGE iterative channel esti-

mation algorithms and compare them with a recently proposed least-squares based estimation algorithm. We study the convergence properties of the proposed schemes, the overall system performance and implementation issues through both theoretical analysis and simulation. At each iteration and for every OFDM link, the EM-type algorithms partition the problem of estimating a multi-input channel into independent channel estimation for each transmit-receive antenna pair, therefore avoiding the matrix inversion encountered in the joint least-squares estimation. We also show that the convergence rate for both algorithms is unrelated to the channel delay profile and decreases when the length of the channel or the number of transmit antennas increases.

Finally, we conclude the dissertation with a summary on the major contributions in Chapter VIII.

B. A Note on Notation

Throughout the dissertation, if not otherwise specified in each chapter, we use the following general rules in notation.

We use boldface and lower case letters to denote vectors and boldface and upper-case letters to denote matrices. Superscripts T, $*$ and H denote transpose, conjugate and transpose conjugate of a matrix or a vector, respectively; $\mathbf{A}^{-1}$, $\text{tr}(\mathbf{A})$ and $|\mathbf{A}|$ denote the inverse, trace and determinant of matrix $\mathbf{A}$, respectively; $\mathbf{I}_n$ denotes the identity matrix of dimension n; when there is no ambiguity on the dimension, $\mathbf{I}$ is used to denote the identity matrix; $\mathbf{A}[i,j]$ denotes the $[i,j]^{th}$ entry of matrix $\mathbf{A}$; a_i denotes the ith entry of vector $\mathbf{a}$.

$\mathbf{E}(\cdot)$ is the expectation operator; $\bar{\mathbf{a}}$ will also be used to denote the mean of $\mathbf{a}$. Symbol $\overset{\triangle}{=}$ is used for definition. Both the scalar or the vector Gaussian distribution

is denoted as $\mathcal{N}(\alpha, \Sigma)$ with α denoting the scalar or vector mean and Σ denoting the variance or the covariance matrix. $f(a|b)$ or $p(a|b)$ is used to denote the conditional PDF of the random variable a given b. $\mathbf{E}(a|b)$ is used to denote conditional mean. Since we do not need to distinguish between a random variable and its value by using different notations in this dissertation, the variable on the right side of the conditional symbol | always denotes the actual value of the corresponding random variable if not otherwise specified.

Given a sequence $a_1, a_2, \cdots, a_n$ of positive numbers, we say that a positive number b_n is of the order of $O(a_n)$ as $n \to \infty$ if $\frac{a_n}{b_n}$ is bounded by some constant.

CHAPTER II

A BRIEF OVERVIEW OF MULTIPLE ANTENNA SYSTEMS

Multiple antenna systems were first used at the receiver side to provide multiple independent spatial copies of the received signal to combat fading in wireless communication systems. The recent interest is mainly in the use of multiple transmit antennas because some important applications limit the use of multiple receiver antennas. For example, it is hard to implement two independent antennas on a small mobile device. If multiple antennas are used on both the transmitter and receiver side in a rich scattering wireless channel, the capacity of such a system with channel known at the receive side can increase linearly as the minimum of the number of transmit and receive antennas increases [1][2]. This discovery has triggered enormous research interests in multiple antenna systems in recent years. In this chapter, we will briefly introduce some capacity results, well known transmission and detection schemes of multiple antenna systems, which are closely related to the rest of the chapters.

A.   Channel Models and Capacity Analysis

1.   Single User MIMO, MISO, SIMO Channels

Consider the point-to-point communication over a rich-scattering frequency non-selective wireless channel with N_t transmit antennas and N_r receive antennas. The system in each channel use can be modelled as follows:

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

where $\mathbf{H}$ is a $N_r \times N_t$ matrix denoting the channel, with each element of the matrix modelled as i.i.d. zero mean, circularly symmetric, complex Gaussian with normalized

variance. If $N_r = 1$, the channel is usually referred to as multiple input single output (MISO) channel . Similarly, we can define a single input and multiple out channel (SIMO) when $N_t = 1$, and a multiple input and multiple out (MIMO) when $N_t \neq 1$ and $N_r \neq 1$. $\mathbf{w} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I})$ is a $N_r \times 1$ vector denoting the circularly symmetric complex Gaussian noise corrupting the different receivers. $\mathbf{y}$ is the received signal vector of dimension $N_r \times 1$. $\mathbf{x}$ denotes the $N_t \times 1$ complex transmit signal (column) vector. Let $\mathbf{S} \triangleq \mathbf{E}[\mathbf{x}\mathbf{x}^H]$. The transmitter is constrained in total power as:

$$\mathrm{tr}(\mathbf{S}) = P. \tag{2.2}$$

a.   Ergodic Capacity

The Ergodic capacity is the maximum average achievable rate of a channel with zero error probability. The ergodic capacity of multiple antenna systems with two different assumptions is summarized below:

- *CSI perfectly known only at the receiver.* In this case, the average mutual information $\mathbf{I}(\mathbf{x}; \mathbf{y}|\mathbf{H})$ between the input and output given $\mathbf{H}$ is maximized when $\mathbf{x}$ is complex Gaussian distributed and can be computed as

$$\mathbf{I}(\mathbf{x}; \mathbf{y}|\mathbf{H}) = \log \left| \mathbf{I}_{N_r} + \mathbf{HSH}^H \right| \tag{2.3}$$

  The ergodic capacity is maximized when $\mathbf{S} = \frac{P}{N_t}\mathbf{I}_{N_t}$ [1].

- *CSI perfectly known at both the transmitter and receiver.* Let the singular value decomposition (SVD) of $\mathbf{H}$ be $\mathbf{H} = \mathbf{UDV}^H$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\mathbf{D}$ is a diagonal matrix. Since $\mathbf{U}$ and $\mathbf{V}$ are available at both the transmitter and receiver, the channel as shown in equation (2.1) can be diagonalized by pre-filtering (multiply) $\mathbf{x}$ by matrix $\mathbf{V}$, and post-filtering (multiply) the received vector $\mathbf{y}$ by $\mathbf{U}^H$. Since orthogonal transformation does not change

the distribution of $\mathbf{x}$ and $\mathbf{W}$, the MIMO channel is transformed into a parallel set of $N = \min\{N_t, N_r\}$ Gaussian scalar channels, whose capacity can be achieved by the water-filling power allocation scheme over space and time [3].

b.   Outage Capacity

In some scenarios of wireless communications, due to delay limits, channel cannot be assumed to be ergodic during the transmission of a code word. For example, the well known quasi-static fading channel model assumes channel remains constant within a transmission block, but changes independently from block to block. In this case, the mutual information expression can be treated as random entities, giving rise to capacity-versus-outage considerations [4].

The channel outage probability is simply defined as

$$\epsilon_o = \text{Prob}(\mathbf{I}(\mathbf{x}; \mathbf{y}) < R_t), \qquad (2.4)$$

where R_t is the target rate. We can also define the outage capacity C_ϵ as the maximum achievable rate at the given target outage probability. For example, $C_{1\%} = 3$ (bit) means that three bits per channel-use can be achieved with a probability of 99%.

2.   Sum-Capacity of Multiple User Systems

We consider here two kinds of multiple user systems: the multiple access channel, where multiple transmitters (or users) communicate to a single receiver and the broadcast channel where a single transmitter communicates to multiple receivers (or users). In a cellular system, the multiple access channel corresponds to the uplink (from mobile to base) and the broadcast channel corresponds to the downlink.

A K user MIMO-MAC channel can be modelled as

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_k + \mathbf{w}, \tag{2.5}$$

where $\mathbf{H}_k$ denotes the matrix channel between the kth transmitter and the receiver. The transmitted signal vector $\mathbf{x}_k$ of the kth user usually has an individual power constraint as

$$\mathrm{tr}(\mathbf{S}_k) \leq P_k \tag{2.6}$$

where P_k is the available transmission power of transmitter k.

The MIMO-BC channel can be modelled as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{w}_k, \quad 1 \leq k \leq K, \tag{2.7}$$

where $\mathbf{y}_k$ is the received signal of the kth user; $\mathbf{H}_k$ denotes the matrix channel between the transmitter and the kth receiver; $\mathbf{w}_k$ denotes the white Gaussian noise at the kth receiver. The transmit power constraint can be expressed as

$$\sum_{k=1}^{K} \mathrm{tr}(\mathbf{S}_k) \leq P, \tag{2.8}$$

where P is the available total transmission power.

In a multiple user system, one usually defines the capacity region to be the closure of the set of achievable rate vectors $(R_1, R_2, \ldots R_K)$, where R_k denotes the rate of the kth user [5]. Besides the capacity region, sum-rate capacity, defined as the maximum achievable sum-rate, $\sum_{k=1}^{K} R_k$, is often used to measure the total throughput of a multiple user system.

a.   Gaussian MIMO MAC

Let $\mathbf{X}_k^c$ denote the set of all user's transmit vectors except $\mathbf{x}_k$. The capacity region of a MAC channel is the closure of the convex hull of all rate vectors $(R_1, R_2, \ldots R_K)$ satisfying

$$R_k \;\leq\; \mathbf{I}(\mathbf{x}_k; \mathbf{y}|\mathbf{X}_k^c), \text{ for any } k, \tag{2.9}$$

$$\sum_{k=1}^{K} R_k \;\leq\; \mathbf{I}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K; \mathbf{y}) \tag{2.10}$$

for some input distribution satisfying the power constraints. For the case of a MIMO-MAC channel, the capacity region is shown to be [6]

$$\mathbf{R} = \bigcup_{\mathrm{tr}(\mathbf{S_k}) \leq P_k, \mathbf{S}_k \succeq 0} B(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_K), \tag{2.11}$$

where $\mathbf{S}_k \succeq 0$ means that $\mathbf{S}_k$ should be positive semi-definite. $B(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_K)$ is defined as the set of $(R_1, R_2, \ldots R_K)$ achieved by a given choice of power allocation scheme $(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_K)$, which can be expressed as

$$R_k \;\leq\; \log_2(|\mathbf{H}_k \mathbf{S}_k \mathbf{H}_k^H + \mathbf{I}|); \text{ for any } k, \tag{2.12}$$

$$\sum_{k=1}^{K} R_k \;\leq\; \log_2 \left(\left| \sum_{k=1}^{K} \mathbf{H}_k \mathbf{S}_k \mathbf{H}_k^H + \mathbf{I} \right| \right). \tag{2.13}$$

The last equation also shows the sum-rate for the given power allocation scheme, which can be maximized over all possible choices of power allocation schemes to obtain the sum-rate capacity.

b.   Gaussian MIMO BC

Compared to the multiple access channel, the broadcast channel is less understood. Only the capacity region of a small class of broadcast channels, called degraded broadcast channels, is known [5]. The most simple type of degraded broadcast channel

is formed by two-user scalar AWGN channels, where one receiver (corresponding to the "good" user) experiences a Gaussian noise with less variance than that of the other user's receiver (the "bad" user). The border of the capacity region in this case can be achieved by cancellation at the receivers; the "bad user" always treats the encoded information for the good user as Gaussian interference; the "good user" can always decode the "bad" user's information first and then cancel its effect and decode its own information.

However, a Gaussian MIMO-BC channel is usually not degraded; thus, its capacity region is unknown. The capacity region and the sum-rate capacity of MIMO-BC and MIMO-MAC have been shown to be closely related. A more detailed treatment of this topic will be given in Chapter III, where the ergodic sum-rate capacity of a fading MIMO-BC is derived and analyzed based on this relation.

B.   Transmission and Detection Schemes for Multiple Antenna Systems

1.   Space-time Coding

The concept of space-time coding was first proposed by Tarokh et al. to improve data rate and reliability of communications over fading channels using multiple transmit antennas [7]. By carefully designing the codewords, potential spatial diversity provided by multiple transmit antennas can be achieved. For example, in a slow and frequency non-selective Rayleigh fading channel, performance is shown to be determined by matrices constructed from pairs of distinct code sequences. The minimum rank among these matrices quantifies the diversity gain, while the minimum determinant of these matrices quantifies the coding gain. Based on these criteria, space-time trellis codes have been designed to achieve 2-3 dB away from the outage capacity. The decoding of space-time trellis codes requires a maximum-likelihood (ML) sequence de-

tection scheme, whose complexity increases exponentially as the the number of trellis states increases.

Another well known type of space-time codes is the orthogonal space-time block codes which have a very simple ML decoding scheme such as the Alamouti's scheme for a system with two transmit antennas [8][9]. Although the orthogonality can simplify detection, it usually results in capacity loss except in some special cases [10]. Some recently proposed block codes can achieve close to capacity rate while maintaining a relatively simple decoding structure [10].

Space time coding techniques usually assume no CSIT. Reliable communication is achieved by careful design of the structure of the code sequences. In this thesis, we are mainly focused on techniques utilizing either full or partial CSIT.

2. ML Detection and Sphere Decoding

Assume that transmit signal $\mathbf{x}$ in the channel model of equation (2.1) is composed of uncoded QAM or QPSK signals. We assume perfectly known $\mathbf{H}$ at the receiver side. The maximum-likelihood (ML) detector can be expressed as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \Omega^{N_t}} ||\mathbf{y} - \mathbf{H}\mathbf{x}||^2 \tag{2.14}$$

where Ω^{N_t} denotes the set of constellation points in the complex N_t-dimensional space. Since exhaustive search for the ML solution over the whole set of Ω^{N_t} is too complex to be implementable, sphere decoding can be used to reduce complexity. Equation (2.14) can be shown to be equivalent to the following:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \Omega^{N_t}} (\mathbf{x} - \mathbf{x}_{ls})^H \mathbf{R}^H \mathbf{R} (\mathbf{x} - \mathbf{x}_{ls}), \tag{2.15}$$

where $\mathbf{x}_{ls} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \mathbf{y}$ is the least-square or zero-forcing estimate of $\mathbf{x}$ assuming $\mathbf{x}$ is continuous; $\mathbf{R}$ is the upper triangular matrix in the QR decomposition of $\mathbf{H} = \mathbf{Q}\mathbf{R}$.

To solve (2.15), the sphere decoder avoids the exhaustive search by considering only those points satisfying $(\mathbf{x}-\mathbf{x}_{ls})^H \mathbf{R}^H \mathbf{R}(\mathbf{x}-\mathbf{x}_{ls}) \leq r^2$. This search can be implemented efficiently by exploiting the triangular structure of $\mathbf{R}$ as shown in [11, 12].

3. ZF and MMSE Detector

Assume $Nr \geq N_t$. Both the zero-forcing (ZF) and the minimum mean square error (MMSE) detectors perform linear transformation over the received signal $\mathbf{y}$ as

$$\mathbf{y}' = \mathbf{B}^H \mathbf{y} = \mathbf{B}^H \mathbf{H} \mathbf{x} + \mathbf{B}^H \mathbf{w} \tag{2.16}$$

ZF uses $B = \mathbf{H}(\mathbf{H}^H\mathbf{H})^{-1}$; MMSE uses $B = \mathbf{H}(\mathbf{H}^H\mathbf{H} + \frac{\mathbf{I}}{\text{SNR}})^{-1}$. Symbol-by-symbol detection is then performed on $\mathbf{y}'$ to detect each element of $\mathbf{x}$. Note that since noise becomes correlated after the transformation, symbol-by-symbol detection, although very simple, is not optimal.

4. BLAST

Different from linear detectors such as ZF and MMSE, the Bell Lab Layered Space-Time (BLAST) scheme [13] is based on nulling and cancelling as introduced below. Denoting the QR decomposition of $\mathbf{H} = \mathbf{QR}$, we can perform a linear transformation on the received signal as $\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$; the system can be expressed as

$$\mathbf{y}' = \mathbf{R}\mathbf{x} + \mathbf{w}', \tag{2.17}$$

where $\mathbf{w}' = \mathbf{Q}^H\mathbf{w}$ has the same distribution as $\mathbf{w}$ since $\mathbf{Q}$ is unitary. In the triangulized model above, each row denotes a different encoding/decoding layer with the kth layer interfered only by layers with indexes larger than k. Considering the N_t^{th} row (layer) of (2.17), which denotes an underlying scalar channel, one can first detect x_{N_t}; assuming $\hat{x}_{N_t}$ is correct, the interference of $\mathbf{R}[N_t - 1, N_t]\hat{x}_{N_t}$ can be subtracted

from layer $N_t - 1$ and $\hat{x}_{N_t - 1}$ can be detected as in a scalar channel. Similarly, layer $N_t - 2, N_t - 3, \cdots, 1$ can be detected in order. In practice, nulling and cancelling is conducted in a certain order. One usually hopes to first detect the strongest channel in order to minimize error propagation [14].

5. Beamforming for MISO Channels

The MISO channel is a very important type of channel in wireless communication systems, and in particular cellular systems due to the fact that multiple receive antennas are hard to implement in a mobile device due to the limited space constraint. Two types of transmit diversity schemes are standardized in the current third generation cellular systems [15]. The closed loop diversity or beamforming, requires CSIT; the open loop diversity, including selection diversity and space time block codes, etc. does not require CSIT.

Consider the channel model of (2.1), where $\mathbf{H} = \mathbf{h}$ is a $1 \times N_t$ vector. Assuming perfect CSIT at the transmitter, the beamforming scheme simply transmits a single data stream, which is weighted by a vector $\frac{\mathbf{h}}{||\mathbf{h}||}$ and then transmitted over different N_t antennas. It can be easily shown that this schemes achieves the capacity of the MISO channel.

6. Transmitter Side Pre-filtering: Zero-Forcing Beamforming and Ranked Known Interference

In a single user system, if both the transmitter and the receiver have perfect CSI, singular value decomposition suggests a natural adaptive transmission scheme that can achieve capacity. Both transmitter and receiver antennas need to co-operate in order to implement the multiplications of $\mathbf{V}$ and $\mathbf{U}^H$ for diagonalizing the $\mathbf{H}$ [3]. However, in a multi-user broadcast channel, since receivers belonging to different users

cannot co-operate, only transmitter side pre-filtering can be used. We introduce below two pre-filtering techniques, namely Zero-Forcing Beamforming (ZFB) and Ranked Known Interference (RKI) [1] which can be viewed as the dual of ZF and BLAST MIMO detectors, respectively [16].

Consider a broadcast channel model similar to (3.1) with N_t transmit antennas at the base and $K^{'}$ users, each with a single receive antenna $(N_t = 1)$. In the case of $N_r \neq 1$, we can view $K^{'} = N_r K$, and still apply the same technique. Let $X = \mathbf{B}\mathbf{v}$, where $\mathbf{B}$ denotes the pre-coding filter and $\mathbf{v}$ is a $K^{'} \times 1$ vector with the $k^{'th}$ element denoting the information signal intended for User $k^{'}$.

In ZFB, $\mathbf{B} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$, so that the system is reduced to $K^{'}$ independent parallel Gaussian channels whose power gain can be shown to be [16]

$$b_{k'} = 1/(\mathbf{H}\mathbf{H}^H)^{-1}[k^{'}, k^{'}]. \tag{2.18}$$

Note that ZFB requires $K^{'} \leq N_t$ for the pseudo-inverse to be available.

Let $m = \text{rank}(\mathbf{H})$. Consider a QR-type decomposition $\mathbf{H} = \mathbf{G}\mathbf{Q}$, where $\mathbf{G} \in C^{K^{'} \times m}$ is a lower triangular matrix and $\mathbf{Q} \in C^{m \times N_t}$ has orthonormal rows. In RKI, $\mathbf{B} = \mathbf{Q}^H$ and the channel becomes a set of m scalar sub-channels with interference as follows:

$$y^{'}_k = \mathbf{G}[k^{'}, k^{'}]v_{k'} + \sum_{j<k^{'}} \mathbf{G}[k^{'}, j]v_j + w_{k'}, \quad k^{'} = 1, 2, \ldots, m. \tag{2.19}$$

We denote $d_{k'} = |g_{k',k'}|^2$ as the power gain of the $k^{'th}$ sub-channel to be used later. Since $\mathbf{v}$ and $\mathbf{G}$ are known at the transmitter, the interference in each channel is non-causally known at the transmitter; therefore, it can be pre-subtracted before transmission using the "dirty paper" type coding schemes [17, 18]. Since the ordering

---

[1]RKI is renamed as "zero-forcing dirty paper coding" in [16]

of the users affects the total achievable rate, the scheme is referred to as "ranked known interference". Note that in this scheme, the base can at most communicate with N_t mobiles at a given instant, as in ZFB.

CHAPTER III

ON THE SUM-RATE CAPACITY OF MIMO FADING BROADCAST
CHANNELS

A.   Introduction

A challenge in the design of cellular systems originates from the sharing of a common transmission medium by multiple users. On one hand, the system capacity of current generation cellular systems is limited by intra-cell and inter-cell interference, motivating techniques aimed at mitigating or suppressing multi-user interference. On the other hand, if some knowledge of user channels is available at the transmitter, adaptive transmission techniques, such as optimal resource allocation and interference pre-subtraction schemes, can be employed to exploit *multi-user diversity* and *avoid* multi-user interference, which can greatly improve overall system capacity.

Consider a single cell and assume interference from other cells is modelled as Gaussian noise for mathematical convenience. Additionally, it is assumed the fading states of all the mobiles are known at the transmitter and all the receivers. Then, due to the presence of multi-user interference, the optimal power control scheme that maximizes the sum-rate of all the users in the cell for both uplink and downlink should consider the fading states of all the user channels. For the case of a single antenna at both the base and the mobile, the authors in [19] show that the maximum ergodic sum-rate capacity for uplink is achieved by a "water-filling" scheme across the mobile users. In other words, at any instance, the base need communicate only with the mobile enjoying the best received signal-to-noise-ratio (SNR). Similar results hold true for the downlink channel. If multiple transmit antennas are available at the base, adaptive antenna array techniques [15] can be employed to maximize the received

effective SNR of a mobile. The base can still communicate with the single user with the best effective SNR. If a mobile also has multiple receive antennas, a multiple-input-multiple-output (MIMO) data link can be established between the mobile and the base. At a given time, the base can communicate with the user whose MIMO link has the largest potential rate. This scheme is referred to as single-user-MIMO (SU-MIMO) in later discussions.

Recently, the authors in [6] showed that the optimal power control scheme that achieves the ergodic sum-rate capacity of a fading MIMO multiple-access-channel (MIMO-MAC) is one that may allow multiple mobile users to communicate with the base. In particular, up to $\frac{1}{2}N(N+1)$ mobile users can communicate with the base at a given instant, where N denotes the number of receive antennas at the base. Can this result be directly extended to the downlink MIMO-broadcast channel (MIMO-BC)? To answer this question, the sum-rate capacity of a non-degraded Gaussian MIMO-BC needs to be evaluated first, whose capacity region is not known [5]. In [16], an interference pre-subtraction strategy using "dirty paper" type coding [17, 18] was proposed and shown to achieve the sum-rate capacity in the case of two transmit antennas and two users each with one receive antenna. Ref. [20] extended the work of [16] to the more general case of arbitrary number of users and antennas, and showed that the optimal precoding structure corresponds to a decision feedback equalizer that decomposes the broadcast channel into a series of single-user channels with interference pre-subtracted at the transmitter. Ref. [21] established a duality between the "dirty paper" achievable region for the MIMO-BC and the capacity region of the MIMO-MAC channel. The authors also showed that the sum-rate capacity of a Gaussian MIMO-MAC is the same as that of the Gaussian MIMO-BC with equal total power constraint, which greatly simplifies the evaluation of the sum-capacity of the MIMO-BC. The duality concept and sum-rate capacity of the MIMO-BC was

also independently studied in [22]. Applying the duality theory, Ref. [6]'s results could be extended to the fading MIMO-BC, i.e., the optimal power control scheme should allow the base to communicate with more than one mobiles simultaneously. Therefore, SU-MIMO may not be optimal.

In this chapter, we use convex optimization techniques to solve the optimal power allocation problem, evaluate the sum-rate capacity, and derive upper and lower bounds on the sum-rate capacity of the fading MIMO-BC. We show that the sum-rate capacity of the fading MIMO-BC with perfect channel state information at the transmitter (CSIT) increases with the number of users K, but at an asymptotically very low rate [23].

In practice, the optimal solution requires large computation and can be hard to implement. Therefore, we study three sub-optimal multiple transmit antenna schemes, SU-MIMO, ranked known interference (RKI) [16] and zero-forcing beam-forming (ZFB) [16, 24], in terms of achievable sum-rate and rate-loss compared to the optimal scheme [23]. Independent work on the topic has also appeared recently in [25, 26].

Note that all results in this chapter are based on the assumption that the number of transmit antennas N_t satisfies $N_t \ll K$, which is practically reasonable for cellular systems. For wireless LAN applications, a recent paper [25] studied the sum-rate capacity when the number of receive antennas is $N_r = 1$ and K grows to infinity in a fixed ratio with N_t ($\frac{N_t}{K} = \beta > 1$), and evaluated the rate of linear growth of the sum-rate capacity. Finally, we note again that perfect CSIT is a key assumption in our model. In practice, a time division duplex (TDD) system under slow mobility conditions could be a good approximation of the model assumed in this correspondence, because in this case the channel states for the downlink could be estimated accurately from the uplink.

This chapter is organized as follows: In Section B we discuss the sum-rate capacity of the fading MIMO-BC and derive an upper-bound to it. In Section C we introduce three transmit pre-processing schemes based on RKI, ZFB and SU-MIMO to exploit multi-user diversity. We also derive the performance of a sub-optimal RKI scheme, which can serve as a lower-bound on the sum-rate capacity. Section D includes simulation results and Section E concludes.

B. Sum-Rate Capacity of Fading MIMO-BC

Consider a discrete-time fading MIMO-BC with N_t transmit antennas at the base and K mobile users each with N_r receive antennas. Let $\mathbf{x} \in \mathbf{C}^{N_t \times 1}$ denote the transmitted vector, $\mathbf{H}_k[j, i]$ the i.i.d. zero-mean flat fading channel gain between transmit antenna i ($1 \leq i \leq N_t$) and receive antenna j ($1 \leq j \leq N_r$) for User k ($1 \leq k \leq K$) and $\mathbf{w}_k \in \mathbf{C}^{N_r \times 1}$ the white Gaussian noise vector with $\mathbf{w}_k \sim N(\mathbf{0}, \mathbf{I})$. Let $\mathbf{y}_k \in \mathbf{C}^{N_r \times 1}$ denote the received vector of the kth user. We have

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{3.1}$$

where $\mathbf{y} = [\mathbf{y}_1^T \ \mathbf{y}_2^T \ \cdots \mathbf{y}_K^T]^T$, $\mathbf{H} = [\mathbf{H}_1^T \ \mathbf{H}_2^T \ \cdots \mathbf{H}_K^T]^T$ and $\mathbf{w} = [\mathbf{w}_1^T \ \mathbf{w}_2^T \ \cdots \mathbf{w}_K^T]^T$, respectively.

As shown in [21], the sum-rate capacity of a Gaussian MIMO-BC is equal to the sum-rate capacity of its dual Gaussian MIMO-MAC channel under the same total power constraint at the transmitter side. This dual MIMO-MAC channel has N_t receive antennas at the base-station and K users each with N_r transmit antennas, with the channel gain between transmit antenna j of User k and receive antenna i equal to $\mathbf{H}_k^*[j, i]$. Given all users' CSI available at both the transmitter and the receiver side, the sum-rate capacity of the dual Gaussian MIMO-MAC based on fixed

channel state $\mathbf{H}$ is [27]

$$C_{sum}^{MAC}(\mathbf{H}) = \max_{\mathbf{S}_k} \log \left(\left| \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{S}_k \mathbf{H}_k + \mathbf{I} \right| \right), \tag{3.2}$$

where $\mathbf{S}_k$ is the covariance matrix of the transmitted complex Gaussian signal vector of User k, subject to the sum power constraint $\sum_{k=1}^{K} \text{tr}(\mathbf{S}_k) \leq P$. According to the duality result of [21], we can evaluate the ergodic sum-rate-capacity of the fading MIMO-BC as

$$C_{sum}^{BC} = \mathbf{E}_{\mathbf{H}} \left\{ C_{sum}^{MAC}(\mathbf{H}) \right\}, \tag{3.3}$$

where the expectation is with respect to the joint channel distribution of $\mathbf{H}$. Since the value of $\mathbf{H}$ is known at the transmitter, C_{sum}^{BC} is achieved by choosing the optimal $\mathbf{S}_k$ for each channel state. The problem of maximizing the ergodic sum-rate capacity of the fading MIMO-BC can be formulated as

$$C_{sum}^{BC} = \max_{\mathbf{S}_k(\mathbf{H})} \mathbf{E}_{\mathbf{H}} \left[\log \left(\left| \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{S}_k(\mathbf{H}) \mathbf{H}_k + \mathbf{I} \right| \right) \right] \tag{3.4}$$

subject to:

$$\sum_{k=1}^{K} \text{tr}(\mathbf{S}_k(\mathbf{H})) \leq P \tag{3.5}$$

$$\mathbf{S}_k(\mathbf{H}) \succeq 0, \quad \text{for} \quad k = 1, 2, ..., K. \tag{3.6}$$

where P is the total available power at the base-station. Note that $\mathbf{A} \succeq 0$ means $\mathbf{A}$ is a positive semi-definite matrix. Here, instead of using a long-term average power constraint $(\mathbf{E}[\sum_{k=1}^{K} \text{tr}(\mathbf{S}_k(\mathbf{H}))] \leq P)$ [28], we use a short-term power constraint, which is a more practical assumption in the cellular downlink. Due to the sum power constraint, this problem is different from the one encountered in computing the sum-rate

capacity of the fading MIMO-MAC [6, 27], where each user has an individual power constraint $(\text{tr}(\mathbf{S}_k(\mathbf{H})) \leq P_k$ for all $k)$. A recent paper [29] extended the iterative water-filling algorithm of [27] to be used in solving the sum power constrained optimization. Although the proposed algorithm is shown to converge in the simulations, a rigorous proof on the convergence and the efficiency of the algorithm in the general case is yet not available. Note that the constrained optimization problem $((3.4)$ - $(3.6))$ is convex with the objective function containing the determinant of a complex Hermitian matrix. This type of problem can be numerically solved by the interior point method of [30]. However, we need to transform the complex matrix optimization problem into an equivalent real format before using the method in [30], which can only deal with real matrices. This process is shown in Appendix A.

Here, we develop an upper bound which is more informative. To simplify notation, we denote $\mathbf{S}_k(\mathbf{H})$ by $\mathbf{S}_k$. Let $\mathbf{h}_k^i$ denote the ith column of $\mathbf{H}_k$ and $\xi_i = \sum_{k=1}^{K} \mathbf{h}_k^{i}{}^H \mathbf{S}_k \mathbf{h}_k^i$, for $i = 1, 2, \cdots, N_t$. Letting $\Psi = \sum_{k=1}^{K} \mathbf{H}_k^H \mathbf{S}_k \mathbf{H}_k + \mathbf{I}$, we have

$$\Psi = \begin{pmatrix} \xi_1 + 1 & \cdots & \cdots & \cdots \\ \cdots & \xi_2 + 1 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \cdots & \cdots & \cdots & \xi_{N_t} + 1 \end{pmatrix}. \tag{3.7}$$

Let λ_k^j denote the jth eigenvalue of $\mathbf{S}_k$ and $\lambda_k^{max} = \max_j \lambda_k^j$. Letting $h_i^{max} = \max_k ||\mathbf{h}_k^i||^2$, we have

$$\xi_i \ \leq \ \sum_{k=1}^{K} ||\mathbf{h}_k^i||^2 \lambda_k^{max}, \text{ for } i = 1, 2, \cdots, N_t, \tag{3.8}$$

$$\leq \ h_i^{max} \sum_{k=1}^{K} \lambda_k^{max}, \text{ for } i = 1, 2, \cdots, N_t, \tag{3.9}$$

where $|| \cdot ||$ denotes Euclidean norm and the first inequality above is in view of the

Rayleigh-Ritz theorem [31]. $||\mathbf{h}_k^i||^2$ is actually the effective channel power gain if maximum ratio combining is employed at the receiver side for the link between User k and base antenna i. The power constraint is equivalent to $\sum_{k=1}^{K} \sum_{j=1}^{N_r} \lambda_k^j = P$ (note that $\lambda_k^j \geq 0$ for any k and j), which suggests

$$\sum_{k=1}^{K} \lambda_k^{max} \leq P. \tag{3.10}$$

We have the following series of bounds:

$$
\begin{aligned}
C_{sum}^{BC} &\leq \mathbf{E}\left\{\log\left[\prod_{i=1}^{N_t} (\xi_i + 1)\right]\right\} \tag{3.11} \\
&\leq \mathbf{E}\left\{\log\left[\prod_{i=1}^{N_t} \left(h_i^{max} \sum_{k=1}^{K} \lambda_k^{max} + 1\right)\right]\right\} \tag{3.12} \\
&\leq N_t \mathbf{E}\left[\log\left(1 + h_1^{max} P\right)\right], \tag{3.13}
\end{aligned}
$$

where the first inequality is in view of Ψ being positive definite and Hadamard's inequality, the second is in view of (3.9) and the third because the h_i^{max} are identically distributed random variables and (3.10). We note that the bound in equation (3.11) without expectation is true for every fading state, and therefore also true when the expectation is taken.

Remarks:

- $||\mathbf{h}_k^i||^2$ $(k = 1, 2, \ldots, K)$ are i.i.d. random variables having central Chi-square distribution with $2N_r$ degrees of freedom, denoted as $\chi_{2N_r}^2$. The corresponding probability density and cumulative density functions are

$$f(z) = z^{N_r - 1} e^{-z} / (N_r - 1)! \tag{3.14}$$

and

$$F(z) = 1 - e^{-z} \sum_{i=0}^{N_r-1} \frac{z^i}{i!}, \tag{3.15}$$

respectively. The asymptotic cumulative distribution function of h_1^{max} ($K \to \infty$), which is the maximum of K i.i.d. $\chi_{2N_r}^2$ distributed random variables, can be evaluated according to [32, 33] as

$$F(z) = \exp\left[-e^{-(z-l_K)}\right], \tag{3.16}$$

where l_K can be computed by solving the following equation

$$e^{-l_K}\left(\sum_{i=0}^{N_r-1} \frac{l_K^i}{i!}\right) = \frac{1}{K}. \tag{3.17}$$

Since $l_K > 0$, we have $\sum_{i=0}^{N_r-1} \frac{l_K^i}{i!} \geq 1$ with equality iff $N_r = 1$. Therefore, equation (3.17) suggests that $l_K \geq \ln(K)$, i.e., the channel gain h_1^{max} grows, on average, at least as $\ln(K)$. Moreover, a larger N_r results in a larger l_K. However, for any fixed N_r, we always have $\lim_{K\to\infty} \frac{\ln(K)}{l_K} = \lim_{K\to\infty}[1-\ln(\sum_{i=0}^{N_r-1} \frac{l_K^i}{i!})/l_K] = 1$. Therefore, as $K \to \infty$, l_K increases as $\ln(K)$ independent of N_r.

To see how C_{sum}^{BC} can be affected by K, we use Jensen's inequality to further bound the right-hand side of (3.13):

$$\begin{aligned} C_{sum}^{BC} &\leq N_t \log\left[1 + \mathbf{E}\left(h_1^{max}\right)P\right] & (3.18) \\ &= N_t \log\left[1 + P[\gamma + l_K + E_i(1, e^{l_K})]\right] & (3.19) \\ &\approx N_t \log\left[1 + P(\gamma + l_K)\right], & (3.20) \end{aligned}$$

where

$$E_i(n, x) = \int_1^\infty e^{-xt}/t^n dt \tag{3.21}$$

is the exponential integral and $\gamma = 0.577215...$ is Euler's constant. The approximation in (3.20) is quite good since

$$E_i(1, e^{l_K}) \quad < \quad \int_1^\infty e^{-e^{l_K}t}dt \tag{3.22}$$

$$\leq \quad \int_1^\infty e^{-Kt}dt = \frac{e^{-K}}{K}, \ (\text{using } l_K \geq \ln(K)) \tag{3.23}$$

even for a moderate value of $K = 5$, $\frac{e^{-K}}{K} = 0.0013 \ll \gamma$. Therefore, for large K, we can ignore $E_i(1, e^{l_K})$. We also note that the RHS of (3.18) is also a good approximation of the RHS of (3.13) when K is large as shown in the simulation results. This is because the asymptotic distribution of h_1^{max} is highly concentrated around l_K, so Jensen's inequality is fairly tight.

- The upper bound of (3.18) suggests that the sum-rate capacity of a fading MIMO-BC with perfect CSI available at the transmitter is mainly limited by N_t when $N_t \ll K$. The upper bound increases log-likely with N_r, which can be roughly concluded by the fact that $\mathbf{E}[||\mathbf{h}_k^i||^2] = N_r$. This result is not surprising since the sum-rate-capacity of the MIMO-BC is bounded by the capacity of the $N_t \times (N_r K)$ single user MIMO channel where receivers can cooperate. Then, according to [1], the ergodic capacity of the single user MIMO channel can increase linearly only with $\min\{N_t, N_r K\} = N_t$ due to $N_t \ll K$. However, we note that as $K \to \infty$ (3.18) is an asymptotically tighter bound than the capacity of the cooperative MIMO system at high SNR. This is because the capacity of the single user MIMO system increases log-likely at high SNR as K increases due to receiver cooperation. In contrast, the proposed upper bound of (3.20) increases log-likely with l_k at high SNR, which in turn increases as $\ln(K)$ as $K \to \infty$. It is well known that the sum-rate capacity increases with increasing number of users due to "multi-user diversity" when perfect CSIT is available

[19, 33]. However, the upper bound suggests that the increase is rather slow for large K (at most of the order of $\log(\ln(K))$ rate).

C. Preprocessing to Achieve Multi-user Diversity

As shown in Appendix A, although computing the optimum power allocation for the dual multiple access channel is possible, it is very complex for large N_t, N_r and K. We also need to transform the optimal solution into the optimal power allocation scheme for the downlink broadcast channel by use of the duality relations [21]. Moreover, all the computations must be repeated for different channel states. Even if the optimum power allocation is available, implementing it is sometimes non-trivial. According to a result on optimal transmission for fading MIMO-MAC in [6] and the duality result in [21], the optimum transmission scheme that maximizes the ergodic sum-rate capacity of a MIMO-BC could allow the base to transmit up to $N_t(N_t + 1)/2$ users simultaneously at a given instant. It is not clear how simultaneous transmission to more than N_t users by interference pre-cancellation can be implemented efficiently. Thus, the need for practical, albeit suboptimal, schemes.

In this section, we consider using two transmitter-side preprocessing schemes, RKI and ZFB, to transmit to a selected group of N_t out of K users to achieve multiuser diversity for the cellular downlink in Subsection 1. In Subsection 2 we derive the achievable sum-rate of RKI using another selection scheme called group user selection whose achievable sum-rate is easy to compute and can serve as a good lower bound to the sum-rate capacity (as shown by simulations).

1.   Multiuser Diversity in a Fading MIMO-BC

We assume a TDMA-like setting where the channel for each user remains constant in each time slot, but varies independently in different time slots. In a given time slot, the channel gains associated with different transmit and receive antenna pairs are modeled as i.i.d. complex Gaussian random variables. The time-slot is assumed long enough that the associated capacity of the instantaneous channel is achievable. Perfect CSI for all users is assumed available at the base-station.

We consider three different schemes: ZFB, RKI and SU-MIMO. In ZFB and RKI, each mobile user has a single receive antenna. In a given time slot, the base selects the $K' = N_t$ users out of K users who have the largest sum-rate capacity to communicate with. In the SU-MIMO case, each user has N_t receive antennas. In a given time slot, the base selects the MIMO-link $(N_t \times N_t)$ which has the largest capacity out of all the users' MIMO links. In all the schemes, the optimum "water-filling" power allocation is used across the associated parallel sub-channels.

Denoting the eigenvalues of matrix $\mathbf{HH}^H$ by $\{\lambda_k\}$, the power gain associated with parallel sub-channels for SU-MIMO, RKI and ZFB are $\{\lambda_k\}, \{d_k\}$ and $\{b_k\}$ $(k = 1, 2, \ldots, N_t)$, respectively. To obtain the optimal solution, we could exhaustively search through all possible combinations of users for the best. This method is referred to as exhaustive ranking. The number of choices required to be considered in the exhaustive ranking is $C_K^{N_t}$ for ZFB, $N_t! C_K^{N_t}$ for RKI and K for SU-MIMO. As N_t increases, the complexity for the RKI and ZFB schemes is high for a relatively large K (say, $K = 40$). Therefore, in RKI and ZFB, we consider a sub-optimal, but much simpler user selection criterion, which requires considering only about $N_t K$ choices. The scheme is as follows: first, we select from K users the one with the largest d_1; we then fix d_1 and select from the remaining $(K - 1)$ users the one causing the largest

d_2. We repeat until all N_t users are selected. This scheme is referred to as successive best user selection in later discussions.

2.   A Lower Bound on the Sum-Rate Capacity

The achievable sum-rate by SU-MIMO can be evaluated as the average value of the maximum of K independent single user MIMO capacities. However, it is hard to theoretically evaluate the achievable sum-rate of RKI and ZFB using either exhaustive ranking or successive best user selection. This is because both user selection criteria introduce dependence among the associated parallel sub-channels, whose joint distribution is difficult to compute.

We consider the following user selection criterion in RKI so that the resulting achievable sum-rate can be analytically computed. The sum-rate capacity of the resulting sub-optimal scheme is then a lower-bound on the sum-rate capacity. Denote the k'-th row of $\mathbf{H}$ (see Equation (3.1)) as $\mathbf{h}_{k'}$, where $1 \leq k' \leq N_r K$. $\mathbf{h}_{k'}$ is therefore the vector channel between the transmitter and the k'-th receive antenna. We first randomly divide all $N_r K$ such vectors into N_t non-overlapping groups. Each group has ϕ_i vectors, such that $\sum_{i=1}^{N_t} \phi_i = N_r K$. Next, we select from group 1 the vector that maximizes the first scalar sub-channel power gain d_1. Given the $(i-1)$ selected vectors, we then select from group i the vector that maximizes d_i successively until all N_t vectors are selected. Obviously, d_i is independent for different i. We can further show that d_i has the same distribution as the maximum of ϕ_i i.i.d $\chi^2_{2(N_t-i+1)}$ random variables [16, 34], whose CDF is easy to derive. Therefore, the achievable rate of such a system with uniform power allocation becomes a computable lower bound to the

sum-rate capacity as shown below:

$$C^{BC}_{sum} > R_{RKI-LB} = \sum_{i=1}^{N_t} \mathbf{E} \left[\log \left(1 + \frac{P}{N_t} d_i \right) \right]. \tag{3.24}$$

In practice, one can simply select $\phi_i = \lfloor \frac{KN_r}{N_t} \rfloor$ for $1 < i \leq N_t$, and $\phi_1 = KN_r - (N_t - 1)\lfloor \frac{KN_r}{N_t} \rfloor$. The corresponding achievable sum-rate is denoted as RKI-LB in Figures 2 and 3 in the simulation section that follows.

Using Jensen's inequality as a good approximation rather than an inequality[1] and similar derivations in Section B, as $K \to \infty$, we have

$$R_{RKI-LB} \approx \sum_{i=1}^{N_t} \log \left[1 + \frac{P}{N_t} (\gamma + l^i_{K'}) \right], \tag{3.25}$$

where $l^i_{K'}$ can be computed by solving the following equation for x,

$$e^{-x} \left(\sum_{j=0}^{N_t-i} \frac{x^j}{j!} \right) = \frac{1}{\lfloor \frac{KN_r}{N_t} \rfloor}. \tag{3.26}$$

Following the same arguments as in Section B, the lower bound increases approximately log-likely with $l^i_{K'}$, which increases at the order of $\ln(\lfloor \frac{KN_r}{N_t} \rfloor)$ as $K \to \infty$ independent of i and N_t.

D. Simulation Results

We conducted simulations to evaluate the sum-rate capacity and the derived upper bounds of the fading MIMO-BC, and comparatively studied the proposed practical schemes in Subsection 1. In all the results, we define SNR as the received SNR at each receive antenna. Since we have normalized both channel gain and noise, SNR = P.

---

[1]Since we are considering a lower bound on the sum-rate capacity, we need an inequality in the other direction. The simplified expression obtained is thus not a lower bound, but an approximation.

Fig. 1 plots the derived upper-bounds and the actual channel sum-rate capacity for $N_t = 2$. For all SNR levels evaluated, doubling of N_r could only result in a capacity increase of less than 1 bit per transmit antenna. For both $N_r = 1$ and $N_r = 2$, the upper bound is about 0.7 bits per transmit antenna above the actual capacity. Although not very tight, the upper bounds estimate the trend of the capacity increase quite well.

Fig. 2 compares the average achievable sum-rate per transmit antenna for the three schemes in the case of $N_t = 2$. We also plotted the performance of a scheme termed single-user beamforming (SU-BF), which transmits only one data stream to the user enjoying the highest SNR using the closed-loop transmit diversity scheme [15] in each time slot. We note again that each mobile is assumed to have two receive antennas in the case of SU-MIMO, but only one receive antenna in all the other cases. In both ZFB and RKI, exhaustive ranking is used. For both SNR levels, SU-BF performs the worst, and only achieves a rate slightly large than half of that of RKI or ZFB at SNR = 30 dB. When the number of users is small, SU-MIMO achieves the largest sum-rate. This is because SU-MIMO allows both transmitter side and receiver side co-processing; the corresponding parallel sub-channels with power gain $\{\lambda_k\}$ have larger capacity than those with $\{d_k\}$ and $\{b_k\}$; therefore SU-MIMO performs better than RKI and ZFB which do not allow receiver side co-processing. However, for a large number of users K, both RKI and ZFB outperform SU-MIMO due to a larger multi-user diversity gain, although only a single receive antenna is used. This can be simply explained by the fact that in RKI and ZFB we have significantly more selections than that in SU-MIMO. Note that in the SU-MIMO scheme, the N_r receive antennas of a user have to be selected together. In contrast, in ZFB and RKI, selection of users to communicate is more "flexible"; the associated N_t parallel sub-channels can be all "good" with higher probability.
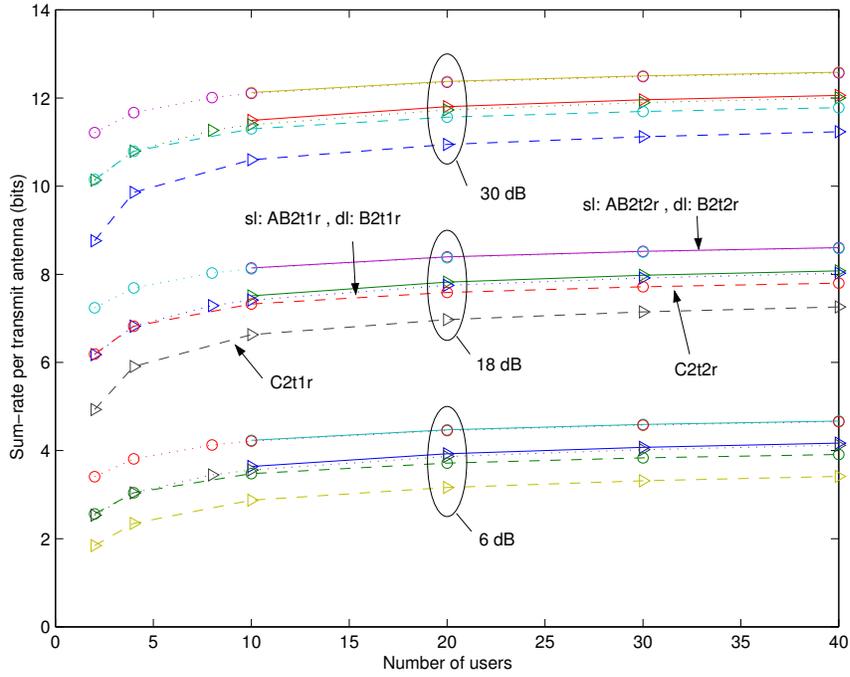
Fig. 1. Upper-bounds on sum-rate capacity for $N_t = 2$. ("Bxtyr","ABxtyr" and "Cxtyr" denote capacity bound (equation (3.13)), asymptotic bound (equation (3.20)), and sum-rate capacity with $N_t = x$ and $N_r = y$, respectively; "sl" and "dl" denote solid line and dotted line, respectively.
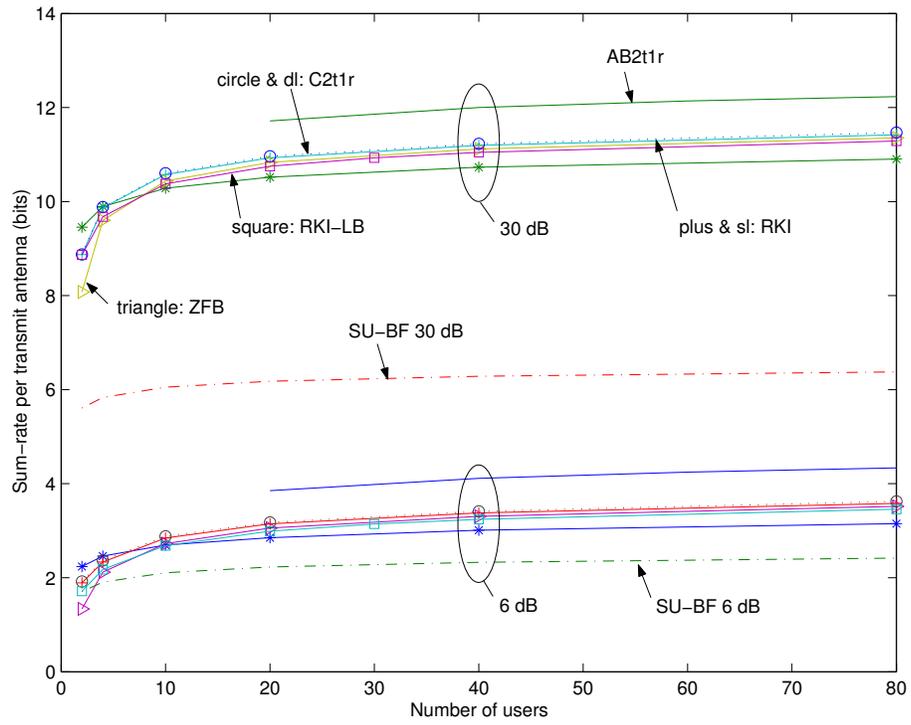
Fig. 2. Achievable sum-rate per transmit antenna ($N_t = 2$).

Another interesting fact is that the performance gap between RKI and ZFB is very small under our assumption of $N_t \ll K$, because in this case the selected N_t user channels tend to be orthogonal to each other, so the associated power gains d_k and b_k tend to be close. The sum-rate capacity (associated with $N_r = 1$) is almost achieved by both schemes at large K, which is in contrast to the case studied by [16, 25] where K grows to infinity in a fixed ratio with N_t ($\frac{N_t}{K} = \beta > 1$). In that case, ZFB only achieves a large fraction ($70\% - 80\%$) of the sum-capacity. Therefore, our setting of a fixed N_t and a much larger K favors the use of ZFB over RKI in a practical cellular system, since the latter requires nested lattice codes, whose encoding and decoding is not easy, to implement interference pre-subtraction [18]. Comparing the achievable sum-rate of ZFB corresponding to $K = 20$ and $K = 40$, we conclude that doubling the number of receive antennas for each user will not significantly increase the achievable sum-rate of ZFB if the number of users is relatively large. However, if each user is equipped with at least N_t receive antennas, zero-forcing receivers can be used instead to decompose each user's MIMO channel into N_t scalar sub-channels as shown in [35]. Note that in this case, only power gains of $K N_t$ scalar channels are fed back to the base for user selection, while ZFB requires the feedback of CSI of K $N_t \times N_t$ MIMO channels in a frequency division duplex (FDD) system. The lower bound (Equation (3.24)) discussed in Subsection 2 with $N_r = 1$, denoted as RKI-LB, is shown to have an asymptotically lower rate than both RKI and ZFB, but is still very tight in this case.

Fig. 3 compares the average achievable sum-rate per transmit antenna for the case of $N_t = 4$. Similar results are observed as in Fig. 2. Note that the upper bound is much looser as compared with the case of $N_t = 2$. The required number of users for both ZFB and RKI to outperform SU-MIMO is much larger than the case of $N_t = 2$ at the same SNR. It is also more obvious that a larger number of users is required for
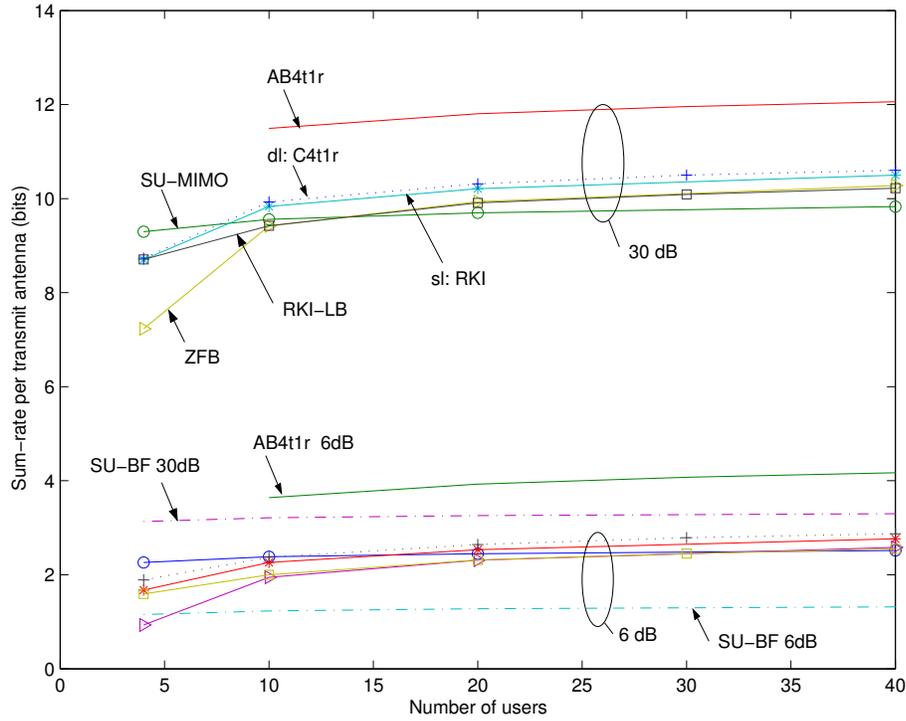
Fig. 3. Achievable sum-rate per transmit antenna ($N_t = 4$).

ZFB and RKI to outperform SU-MIMO at lower SNR than at high SNR.

Fig. 4 shows the performance of ZFB and RKI using the successive best user selection for $N_t = 2$ (denoted as SUB-ZFB and SUB-RKI, respectively). In both cases, the loss due to simplified selection is negligible compared to exhaustive ranking.

E. Conclusion

We have evaluated the sum-rate capacity of the flat fading MIMO-BC with perfect CSI at the transmitter, both through exact numerical computation and derived upper and lower bounds. We show that as the number of users $K \to \infty$, the upper bound increases log-likely with a parameter which only increases on the order of $\ln(K)$. Therefore, the increase of sum-rate capacity due to multi-user diversity cannot be above this rate asymptotically. We have also comparatively studied by simulation
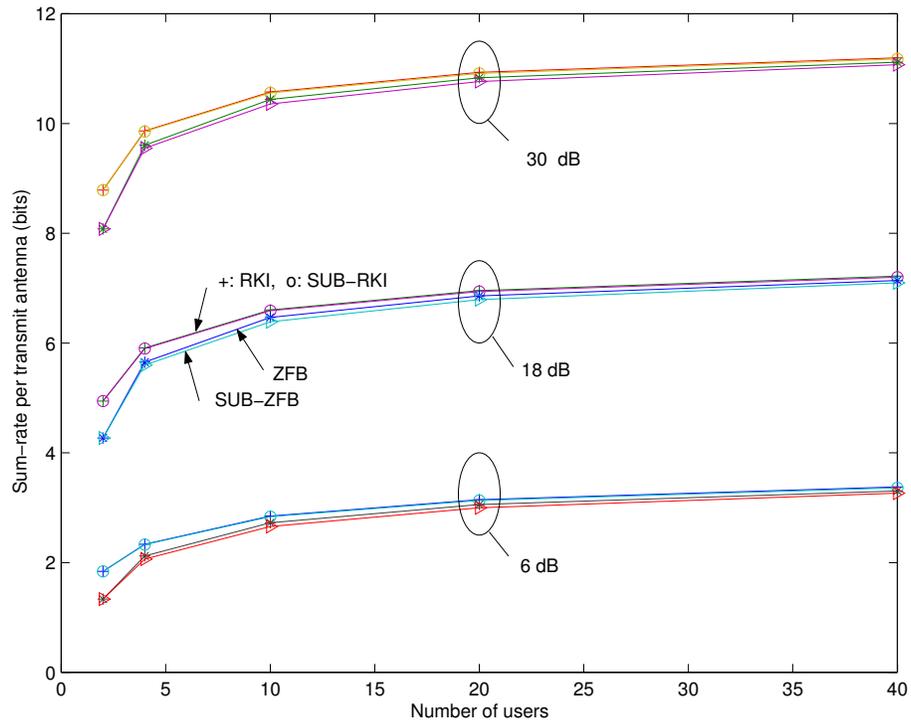
Fig. 4. Achievable sum-rate per transmit antenna for different user selection criteria $(N_t = 2)$.

three different MIMO transmission schemes, RKI and ZFB which transmit to a selected set of N_t users each with one receive antenna, and SU-MIMO which transmits to the "best" single user with N_t receive antennas. For $N_t \ll K$, ZFB and RKI are shown to be able to achieve higher capacity than SU-MIMO due to larger multi-user diversity caused by more flexible user selection; both RKI and ZFB are shown to be able to achieve rates close to the sum-rate capacity.

CHAPTER IV

MINIMUM OUTAGE PROBABILITY TRANSMISSION WITH IMPERFECT

FEEDBACK FOR MISO FADING CHANNELS[*]

A. Introduction

In wireless communications, in addition to the ergodic capacity which characterizes the long-term average achievable rate limit of a fading channel, information outage capacity [4] is also used since practical codeword lengths are limited due to delay constraints. Perfect channel state information at the transmitter (CSIT) has been shown to significantly improve the channel outage capacity for both single and multiple antenna systems [28, 36]. However, if only *imperfect* CSIT is available, is the optimum transmit strategy changed? How is the outage capacity affected? These questions are usually hard to answer partly due to the difficulty in evaluating the distribution of the instantaneous mutual information.

Some recent papers proposed some partial CSIT models in the case of multiple transmit antennas and a single receive antenna and studied optimum transmission in terms of maximizing the ergodic capacity using these models [37, 38]. Here we are interested in the models of [38]. It is assumed that the receiver has perfect channel state information, and feedback some channel information to the transmitter. Based on the feedback, the transmitter models the channel as shown in the following two cases:

- Mean Feedback: The channel distribution is modelled at the transmitter as $\mathbf{h} \sim \mathcal{N}(\bar{\mathbf{h}}, \sigma_h^2 \mathbf{I})$, where the mean $\bar{\mathbf{h}}$ could be interpreted as an estimate or prediction

of the channel based on feedback and σ_h^2 as the variance of the estimation or prediction error. This is the case of slow fading.

- Covariance Feedback: The channel distribution is modelled as $\mathbf{h} \sim \mathcal{N}(\mathbf{0}, \Sigma)$. This models very fast fading, in which the feedback channel fails to provide an accurate estimate of the current channel value. However, Σ, determined by the relative geometry of the propagation paths, changes slowly compared to the fading, thus can be tracked by feedback. Moreover, Σ is practically the same for both the uplink and the downlink channels in FDD systems and, therefore, can be estimated from uplink data, obviating the need for feedback [39].

The solution in both cases is determined by solving simple numerical optimization problems. When there is a moderate disparity between the strengths of different paths from the transmitter to the receiver, it is nearly optimal to employ the simple beamforming strategy that concentrates all the transmit power in the strongest direction indicated by the feedback. This problem is further studied by [40, 41] for some special MIMO channels in the case of covariance feedback. The condition for beamforming to be optimal is studied in [41, 42, 43]. In this chapter, we study the optimal transmission strategies in terms of minimum outage probability for the case of mean feedback. In particular, we prove that the optimal transmission directions in this case are the same as in maximizing the ergodic capacity. Note that a recent paper [44] also studied the same problem and provided a proof, which, however, seems to be wrong (See Appendix II of [44], in particular, the derivation associated with Equation (65) and (66)). We also provide some supplemental results on minimum outage probability transmission in the case of covariance feedback other than those presented in [41]. Note that for fast fading, it may not be meaningful in the covariance feedback case to consider channel outage capacity as defined in [4] and in this chapter

when a codeword may experience different channel states. However, our results for this case can be used to illustrate how knowledge of Σ at the transmitter can affect the outage capacity of a block fading channel [45].

B. Mean Feedback

Consider the discrete model of a multiple-input-single-output (MISO) fading channel [38],

$$y = \mathbf{x}^H \mathbf{h} + n, \tag{4.1}$$

where $n \sim \mathcal{N}(0, \sigma_n^2)$ is circularly symmetric complex Gaussian noise with variance $\frac{1}{2}\sigma_n^2$ per dimension. Let N_t denote the number of transmit antennas. $\mathbf{h}$ is an $N_t \times 1$ complex Gaussian channel vector modelled as $\mathbf{h} \sim \mathcal{N}(\bar{\mathbf{h}}, \sigma_h^2 \mathbf{I})$ and $\mathbf{x}$ denotes the complex transmit signal vector. Define $\mathbf{S} \triangleq \mathbf{E}[\mathbf{x}\mathbf{x}^H]$. For a given $\mathbf{S}$, since the optimum input distribution of $\mathbf{x}$ that maximizes the conditional mutual information $\mathbf{I}(\mathbf{x}; y|\mathbf{h} = h)$ for any fixed channel realization h is zero mean complex Gaussian, the problem of minimal outage transmission can be formulated as

$$\min_{\mathbf{S}} \epsilon(R) = \min_{\mathbf{S}} \Pr\left[\log_2\left(\frac{P}{\sigma_n^2}\mathbf{h}^H \mathbf{S} \mathbf{h} + 1\right) < R\right] \tag{4.2}$$

$$= \min_{\mathbf{S}} \Pr\left[\mathbf{h}^H \mathbf{S} \mathbf{h} < t\right] \tag{4.3}$$

$$\text{subject to:} \quad \text{tr}(\mathbf{S}) = 1, \tag{4.4}$$

where $t \triangleq \frac{\sigma_n^2}{P}(2^R - 1)$; R and P denote the target rate and available power, respectively. Note that here we only consider short-term power control [28] so that P is not a function of time. Since $\mathbf{S}$ is positive semi-definite, we have the eigenvalue decomposition (EVD) $\mathbf{S} = \mathbf{U}\mathbf{D}\mathbf{U}^H$, where $\mathbf{D} = \text{diag}\{d_1, d_2, \cdots, d_{N_t}\}$ is a diagonal matrix with $d_i \geq 0$ indicating the power allocated to transmission directions indicated by

the corresponding column vectors of unitary matrix $\mathbf{U}$. Letting $\nu = \mathbf{U}^H \mathbf{h}$, we have $\nu \sim \mathcal{N}(\mathbf{U}^H \bar{\mathbf{h}}, \sigma_h^2 \mathbf{I})$. The problem can then be expressed as

$$\min_{\mathbf{U},\mathbf{D}} \epsilon(t) = \min_{\mathbf{U},\mathbf{D}} \Pr\left[\nu^H \mathbf{D} \nu < t\right] \tag{4.5}$$

$$= \min_{\{d_i\},\{\bar{\nu}_i\}} \Pr\left[\sum_{i=1}^{N_t} d_i |\nu_i|^2 < t\right] \tag{4.6}$$

$$\text{subject to:} \quad \sum_{i=1}^{N_t} d_i = 1 \tag{4.7}$$

$$\text{and} \quad \sum_{i=1}^{N_t} |\bar{\nu}_i|^2 = \xi \stackrel{\triangle}{=} \|\bar{\mathbf{h}}\|^2, \tag{4.8}$$

where $\nu_i \sim N(\bar{\nu}_i, \sigma_h^2)$ are independent circularly symmetric complex Gaussian random variables. Therefore, $|\nu_i|^2$ is non-central Chi-square distributed with 2 degrees of freedom and a non-centrality parameter $|\bar{\nu}_i|^2$. To minimize $\epsilon(t)$, we need to find the optimal values of $|\bar{\nu}_i|$ and d_i for $i = 1, 2, \cdots, N_t$, subject to their individual constraints. We first fix the power allocation to each transmit antenna by fixing the d_i and consider the optimization with respect to $|\bar{\nu}_i|$. The following Lemma helps provide a solution.

Lemma 1. *Given two real and independent random variables $X \sim \mathcal{N}(\bar{x}, 1)$, $Y \sim \mathcal{N}(\bar{y}, 1)$, and $\gamma > 1$, then*

$$\min_{\bar{x},\bar{y}} Pr(X^2 + \gamma^{-2} Y^2 < q^2)$$

subject to $\bar{x}^2 + \bar{y}^2 = m^2$ is attained at $\bar{x} = m$, $\bar{y} = 0$ for any q.

The proof is given in the Appendix B. Based on this lemma, we have the following proposition:

Proposition 1. *Without loss of generality, assume $d_1 > d_i$ for $2 \leq i \leq N_t$. Then the optimal solution of the problem in (4.6) is achieved by $|\bar{\nu}_1|^2 = \xi$, $|\bar{\nu}_i| = 0$ for all*

$i \neq 1$.

Proof. Since the distribution of $|\nu_i|$ is not dependent on $\theta \overset{\triangle}{=} \arctan\left(\frac{\text{Im}\{\bar{\nu}_i\}}{\text{Re}\{\bar{\nu}_i\}}\right)$, without loss of generality, we can let $\nu_i = a_i + jb_i$ for all $1 \leq i \leq N_t$, where $a_i \sim N(|\bar{\nu}_i|, \frac{1}{2}\sigma_h^2)$ and $b_i \sim N(0, \frac{1}{2}\sigma_h^2)$ are two independent real Gaussian random variables. Therefore, the optimization over $\bar{\nu}_i$ can be equivalently transferred to the optimization over $\{\bar{a}_i\}$.

Denote the optimal solution as $\{\bar{a}_i^{opt}\}$ $(i = 1, 2, \cdots N_t)$ and the associated minimum outage probability as ϵ_{min}. Assume there exists some $k, (k \neq 1)$, such that $\bar{a}_k^{opt} > 0$. Construct $Z_1 = d_1 a_1^2 + d_k a_k^2 = d_1(a_1^2 + \frac{d_k}{d_1} a_k^2)$ and $Z_2 = \sum_{i=2, i \neq k}^{N_t} d_i a_i^2 + \sum_{i=1}^{N_t} d_i b_i^2$. Let the cumulative distribution function (CDF) of Z_1 and the probability density function (PDF) of Z_2 be $F_{Z_1}(x)$ and $f_{Z_2}(x)$, respectively. Applying Lemma 1 to $F_{Z_1}(x)$, we have the following inequality

$$\epsilon_{min} = \Pr[Z_1 + Z_2 < t] = \int_0^t F_{Z_1}(t-x)f_{Z_2}(x)dx > \int_0^t F'_{Z_1}(t-x)f_{Z_2}(x)dx \quad (4.9)$$

where F'_{Z_1} is the CDF associated with another solution $\{\bar{a}'_i\}$ $(i = 1, 2, \cdots N_t)$, in which $\bar{a}'_k = 0$, $\bar{a}'_1 = \sqrt{(\bar{a}_1^{opt})^2 + (\bar{a}_k^{opt})^2}$ and $\bar{a}'_i = \bar{a}_i$ for all $i \neq 1, k$. Since the solution of $\{\bar{a}'_i\}$ achieves lower outage probability than ϵ_{min}, $\{\bar{a}_i^{opt}\}$ is not optimal, which contradicts our assumption. Thus, we have $\bar{a}_i^{opt} = 0$ for any $i \neq 1$. $\square$

When there are more than only d_1 which have equal largest value, it can be easily shown that the choice of $\bar{a}_i^2$ can be arbitrary within this set, as long as their sum is equal to ξ. Therefore, allocating ξ to a single $\bar{a}_i^2$ is still optimal, which indicates that the optimal choice of $\bar{\nu}$ is $[\beta, 0, \cdots, 0]^H$, where β is a complex scalar such that $|\beta|^2 = \xi$. Without loss of generality, let $\beta = \sqrt{\xi}$. Since $\bar{\nu} = \mathbf{U}^H \bar{\mathbf{h}}$, we have $\mathbf{U} = \left[\frac{\bar{\mathbf{h}}}{||\bar{\mathbf{h}}||}, \mathbf{u}_2 \cdots, \mathbf{u}_{N_t}\right]$, where $\{\mathbf{u}_i\}$ $(2 \leq i \leq N_t)$ is an arbitrary set of $(N_t - 1)$ orthonormal vectors that are orthogonal to $\bar{\mathbf{h}}$. Therefore, the optimal transmission directions to minimize outage probability are the vector channel mean and its orthog-

onal directions, which are the same as in the case of maximizing the ergodic capacity [38].

After **U** is obtained, we still need to determine the optimal power allocation, **D**, in the different transmission directions. Here we will briefly analytically study this problem and will follow up with numerical results in Section D. Let $Z_i = d_i|\nu_i|^2$ in equation (4.6). Given the optimal **U** as above, Z_1 is non-central χ^2-distributed with 2 degrees of freedom, mean $d_1(\sigma_h^2 + \xi)$ and a non-centrality parameter $d_1\xi$. The Z_i, $2 \le i \le N_t$, are each central χ^2-distributed with 2 degrees of freedom and mean $d_i\sigma_h^2$. Therefore, the outage probability can be evaluated (after some simple change of variables) as:

$$
\begin{aligned}
\epsilon(t) &= \Pr\left[\sum_{i=1}^{N_t} Z_i < t\right] \\
&= \left(\frac{\xi}{\sigma_h^2}\right)^{N_t} \frac{e^{-\frac{\xi}{\sigma_h^2}}}{\prod_{i=1}^{N_t} d_i} \iint_{\mathcal{D}} e^{-\frac{\xi}{\sigma_h^2}\sum_{i=1}^{N_t}\frac{z_i}{d_i}} I_0\left(2\sqrt{\frac{z_1}{d_1}\frac{\xi}{\sigma_h^2}}\right) dz_1 dz_2 \cdots dz_{N_t} \quad (4.10)
\end{aligned}
$$

where

$$
\mathcal{D} \triangleq \left\{(z_1, z_2, \cdots, z_{N_t}) \in \mathbb{R}^{N_t} \,\Big|\, \sum_{i=1}^{N_t} z_i < t/\xi, \;\; z_i \ge 0 \text{ for } 1 \le i \le N_t\right\}
$$

and $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind. We define the ratio $\mathrm{SNR}_{fb} \triangleq \frac{\xi}{\sigma_h^2}$ as the feedback SNR. We see that the minimum outage probability is strongly a function of SNR_{fb}, which is a measure of the accuracy of CSIT. When t is sufficiently small, $I_0\left(2\sqrt{\frac{z_1}{d_1}\frac{\xi}{\sigma_h^2}}\right)$ and $e^{-\frac{\xi}{\sigma_h^2}\sum_{i=1}^{N_t}\frac{z_i}{d_i}}$ are approximately 1 and the outage probability is mainly affected by $\frac{1}{\prod_{i=1}^{N_t} d_i}$, which is minimized by making the d_i equal. Therefore, for sufficiently small t (associated with a very small outage probability and a low target rate for given transmission power P and noise covariance σ_n^2), the optimal power allocation tends to spread power over different transmission directions. On the contrary, as t increases, $I_0\left(2\sqrt{\frac{z_1}{d_1}\frac{\xi}{\sigma_h^2}}\right)$ could increase very rapidly as z_1 increases.

Making d_1 larger, and thus the rest of the d_is smaller reduces $I_0\left(2\sqrt{\frac{z_1}{d_1}}\frac{\xi}{\sigma_h^2}\right)$. Therefore, beamforming, which concentrates power in the direction of the channel mean might be optimal for larger outage probability. $e^{-\frac{\xi}{\sigma_h^2}\sum_{i=1}^{N_t}\frac{z_i}{d_i}}$ as a function of d_i is relatively hard to evaluate due to its dependency on z_i. So, the above analysis is only approximate, but well matches the numerical results to be presented later.

C. Covariance Feedback

Given a fixed power allocation scheme, transmitting along the eigenvectors of the channel covariance matrix Σ is proven to be necessary and sufficient to achieve minimum outage probability in the covariance feedback case [41]. Note that this is the same strategy as maximizing the ergodic capacity as shown in [38]. Here, to make our presentation complete, we provide a slightly different, but much more concise proof using an inequality from [40].

Proposition 2. *Let* $\mathbf{h} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, *the EVD of* Σ *be* $\Sigma = \mathbf{U}_\Sigma \mathbf{D}_\Sigma \mathbf{U}_\Sigma^H$, *and* $tr[\Sigma] = N_t$[1]. *Choosing* $\mathbf{S} = \mathbf{U}_\Sigma \mathbf{D_S} \mathbf{U}_\Sigma^H$, *where* $\mathbf{D_S}$ *is a diagonal matrix, is necessary and sufficient for minimizing the outage probability.*

Proof. We first assume $\Sigma \neq \mathbf{I}$. Letting $\mathbf{S} = \mathbf{U}_\Sigma \Psi \mathbf{U}_\Sigma^H$, we need to show that Ψ is diagonal in order to minimize outage probability. As in the last section, minimizing the outage probability at a given rate R is equivalent to minimizing

$$\epsilon(t) = \Pr[\mathbf{h}^H \mathbf{S} \mathbf{h} < t] \tag{4.11}$$

$$= \Pr[\nu^H (\mathbf{D}_\Sigma^{1/2} \Psi \mathbf{D}_\Sigma^{1/2})\nu < t], \tag{4.12}$$

$$\text{subject to:} \quad tr(\mathbf{S}) = tr(\Psi) = 1, \tag{4.13}$$

---

[1] *If* $tr[\Sigma] = N_t\sigma^2$, *we can always normalize the channel by absorbing* σ^2 *into the transmit power* P

where $\nu \triangleq \mathbf{D}_{\Sigma}^{-1/2}\mathbf{U}_{\Sigma}^{H}\mathbf{h}$ so that ν is white with distribution of $\mathcal{N}(\mathbf{0}, \mathbf{I})$. Let $\Pi \triangleq \mathbf{D}_{\Sigma}^{1/2}\Psi\mathbf{D}_{\Sigma}^{1/2}$ with EVD $\Pi = \mathbf{U}_{\Pi}\mathbf{D}_{\Pi}\mathbf{U}_{\Pi}^{H}$. Then, the minimum outage probability problem is rewritten as

$$\min_{\Psi} \epsilon(t) \quad = \quad \min_{\Pi} \Pr\left[\nu^{H}\Pi\nu < t\right], \tag{4.14}$$

$$\text{subject to:} \quad \text{tr}\left(\mathbf{D}_{\Sigma}^{-1/2}\Pi\mathbf{D}_{\Sigma}^{-1/2}\right) = \text{tr}\left(\Psi\right) = 1. \tag{4.15}$$

Letting $\nu' \triangleq \mathbf{U}_{\Pi}^{H}\nu$, we can further transform (4.14) to:

$$\min_{\Psi} \epsilon(R) = \min_{\Pi} \Pr\left[\nu'^{H}\mathbf{D}_{\Pi}\nu' < t\right]. \tag{4.16}$$

It is shown in [40] that

$$\text{tr}\left(\mathbf{D}_{\Sigma}^{-1/2}\mathbf{D}_{\Pi}\mathbf{D}_{\Sigma}^{-1/2}\right) \le \text{tr}\left(\mathbf{D}_{\Sigma}^{-1/2}\Pi\mathbf{D}_{\Sigma}^{-1/2}\right) = 1. \tag{4.17}$$

Comparing equations (4.14) and (4.16), since ν' and ν have the same distribution, the minimum outage probability achieved by the optimal Π can always be achieved by a diagonal matrix $\mathbf{D}_{\Pi}$ with at least the same or even more stringent power constraint. Therefore, choosing Π as diagonal is sufficient to minimize outage probability. Equivalently, this means Ψ is diagonal since $\Pi \triangleq \mathbf{D}_{\Sigma}^{1/2}\Psi\mathbf{D}_{\Sigma}^{1/2}$. Necessity is easily proven using the fact that the first equality in equation (4.17) is satisfied if and only if Π is diagonal. For the trivial case of $\Sigma = \mathbf{I}$, because any set of orthonormal vectors can be used as beamforming directions, the theorem is easily proved. $\square$

Next, it remains to determine the optimal power allocation in different transmission directions, i.e., the diagonal entries of $\mathbf{D_S} \triangleq \text{diag}\{p_1, p_2, \ldots, p_{N_t}\}$. Let

$\mathbf{D}_\Sigma \triangleq \operatorname{diag}\{\lambda_1, \lambda_2, \ldots, \lambda_{N_t}\}$. The optimal power allocation problem is defined as:

$$\min_{\{p_i\}} \epsilon(t) \quad = \quad \min_{\{p_i\}} \Pr\left[\sum_{i=1}^{N_t} |\nu_i|^2 \lambda_i p_i < t\right], \tag{4.18}$$

$$\text{subject to:} \quad \sum_{i=1}^{N_t} \lambda_i = N_t \text{ and } \sum_{i=1}^{N_t} p_i = 1, \tag{4.19}$$

where $|\nu_i|^2$ is exponentially distributed. Although tedious, $\epsilon(t)$ in this case can be computed in a closed form expression. Therefore, we could solve the constrained optimization problem of equation (4.18) numerically. However, the optimization problem is not convex, which makes it difficult to find the global optimal $\{p_i\}$. Here, we coarsely identify the conditions when the optimum strategy is to spread power over different directions and when to concentrate all the power in a single direction using majority theory [46], to which we give a brief introduction next.

Majorization: Given two real positive vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ having equal summation of all entries. $\mathbf{a}$ is majorized by $\mathbf{b}$, denoted as $\mathbf{a} \prec \mathbf{b}$, if the sum of the k smallest entries of $\mathbf{a}$ is greater than or equal to the same sum for $\mathbf{b}$ for all $k = 1, 2, \cdots, n$. This is a mathematical description of the vague concept of $\mathbf{a}$ is "less spread out" than $\mathbf{b}$.

Schur-convexity: A real-valued function ϕ defined on a set of $\mathbb{R}^n$ is Schur-convex, if $\mathbf{a} \prec \mathbf{b} \implies \phi(\mathbf{a}) \leq \phi(\mathbf{b})$. ϕ is Schur-concave, if $\mathbf{a} \prec \mathbf{b} \implies \phi(\mathbf{a}) \geq \phi(\mathbf{b})$. Schur-convexity and Schur-concavity can be viewed as extensions of the increase or decrease functions defined on $\mathbb{R}$.

Now we consider two special cases for the problem in (4.18):

Case 1: $N_t = 2$. Then,

$$\epsilon(R) \quad = \quad \int_0^{\frac{t}{(1-p_1)\lambda_2}} \left[1 - \exp\left(-\frac{t - x\lambda_2(1-p_1)}{\lambda_1 p_1}\right)\right] e^{-x} dx \tag{4.20}$$

$$= \quad 1 - \frac{\lambda_2(1-p_1)}{\lambda_2(1-p_1) - \lambda_1 p_1} e^{-\frac{t}{(1-p_1)\lambda_2}} + \frac{\lambda_1 p_1}{\lambda_2(1-p_1) - \lambda_1 p_1} e^{-\frac{t}{p_1 \lambda_1}}. \tag{4.21}$$

In particular, when $\lambda_1 = \lambda_2$, we can show that $\epsilon(R)$ is a Schur-convex function when $t < t_0$, and Schur-concave function when $t \geq t_0$ where the constant $t_0 = 1.2564\ldots$ is a solution to the equation $e^{-t} - (1 + 2t)e^{-2t} = 0$. Therefore, if $R \leq \log\left(1 + \frac{t_0 P}{\sigma_{\hat{n}}^2}\right)$, we need to spread power equally on two orthogonal transmission directions to minimize outage probability; otherwise, we should concentrate on a single transmission direction to minimize outage.

Case 2: $\Sigma = \mathbf{I}$, and $N_t \geq 2$.

Consider optimizing p_1, p_2 for some fixed power allocation $\check{p}_3, \cdots, \check{p}_{N_t}$ on the remaining directions,

$$\min_{\{p_1, p_2\}} \epsilon(t', p_1, p_2) = \min_{\{p_1, p_2\}} \Pr\left[|\nu|_1^2 p_1 + |\nu|_2^2 p_2 < t'\right], \qquad (4.22)$$

where $t' \triangleq t - \sum_{i=3}^{N_t} |\nu|_i^2 \check{p}_i$. As a function of p_i for all i, $\epsilon(t, p_1, p_2, p_3, \cdots, p_{N_t})$ is defined symmetrically for all p_i. Since $\epsilon(t', p_1, p_2)$ is a Schur-concave function for vector $[p_1, p_2]$ if $t' \geq t_0$ and a Schur-convex function if $t' < t_0$, we can conclude that $\epsilon(t', p_1, p_2, p_3, \cdots, p_{N_t})$ is Schur-convex for some small enough t or Schur-concave for some large enough t according to Theorem 3.A.5 of $[46]^2$. Therefore, for a given channel received SNR $\text{SNR}_{rv} \triangleq \frac{P}{\sigma_{\hat{n}}^2}$, the optimal power allocation scheme tends to spread power when target rate R is low, and concentrate power when R is large. However, in the latter case, the outage probability could be very large (say > 0.5), and not of practical interest. So, spreading power over different transmission directions is not necessarily optimum for minimizing outage probability when $\Sigma = \mathbf{I}$. Through numerical results, we can also show that when $\Sigma \neq \mathbf{I}$, the optimal power allocation schemes still tend to spread power when outage probability is low and concentrate

---

^{2}Let $f(\mathbf{x})$ be symmetric in each element of $\mathbf{x}$, where $\mathbf{x} = [x_1, x_2, ...x_N]$. Theorem 3.A.5 states that to prove $f(\mathbf{x})$ is Schur-convex in $\mathbf{x}$, it is sufficient to show $f(\mathbf{x})$ is Schur-convex in a vector composed of any two elements of $\mathbf{x}$ by fixing the other elements.

power otherwise.

D. Numerical Results

Figures 5 and 6 plot the achievable minimum outage probability of mean-feedback for different rates and SNR_{fb}, for a fixed average received SNR, $\mathrm{SNR}_{rv} \triangleq \frac{P(\xi + \sigma_h^2)}{\sigma_n^2} = 8$ dB. We can see outage probability decreases exponentially as the target rate decreases. SNR_{fb} affects outage capacity significantly, especially at low outage probability. Note here that the capacity of an AWGN channel at SNR $= 8$ dB is about $C_{8dB} = 2.87$ bits/channel-use. To achieve 1% outage, the required SNR_{fb} for 1.15 bits/channel-use (40% of C_{8dB}) and for 1.72 bits/channel-use (60% of C_{8dB}) is about 9 dB and 12 dB, respectively. Figure 7 compares the optimal power allocation over different transmission directions for mean-feedback. When "good" channel feedback is available, the optimal solution tends to beamform to the direction indicated by the mean with all available power. On the contrary, "bad" channel feedback may require multiple beams to be transmitted. Moreover, at higher rate or larger outage probability, beamforming to the direction indicated by the channel mean with all power is optimal. Figure 8 shows the minimum outage probability achievable by covariance feedback with the optimal power and the equal power allocation schemes for $N_t = 2$. Let the EVD of Σ be $\Sigma = \mathbf{U}_\Sigma \mathbf{D}_\Sigma \mathbf{U}_\Sigma{}^H$ and λ_1 and λ_2 ($\lambda_1 + \lambda_2 = 2$) be the two diagonal elements of $\mathbf{D}_\Sigma$. Even when $\lambda_1 \ll \lambda_2$, which indicates high spatial correlation between transmit antennas, optimal power allocation cannot significantly reduce outage probability compared to the scheme in which power is uniformly allocated to both orthogonal directions.
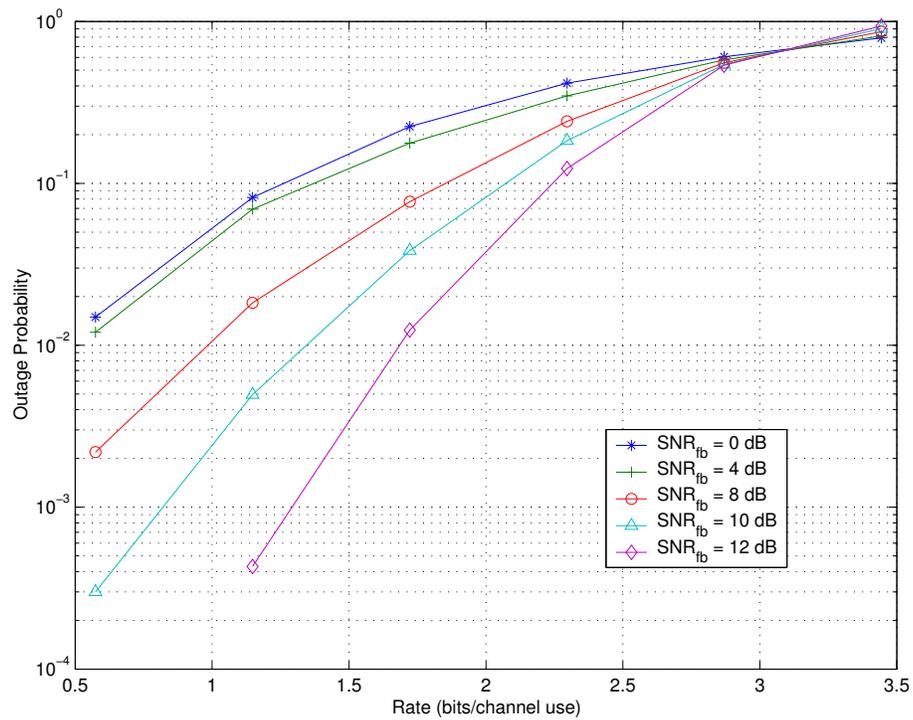
Fig. 5. Achievable minimum outage probability vs target rate of mean-feedback for $N_t = 2$. Channel received SNR, $\mathrm{SNR}_{rv} \triangleq \frac{P(\xi + \sigma_h^2)}{\sigma_n^2}$, is fixed at 8dB. $\mathrm{SNR}_{fb} \triangleq \frac{\xi}{\sigma_h^2}$.
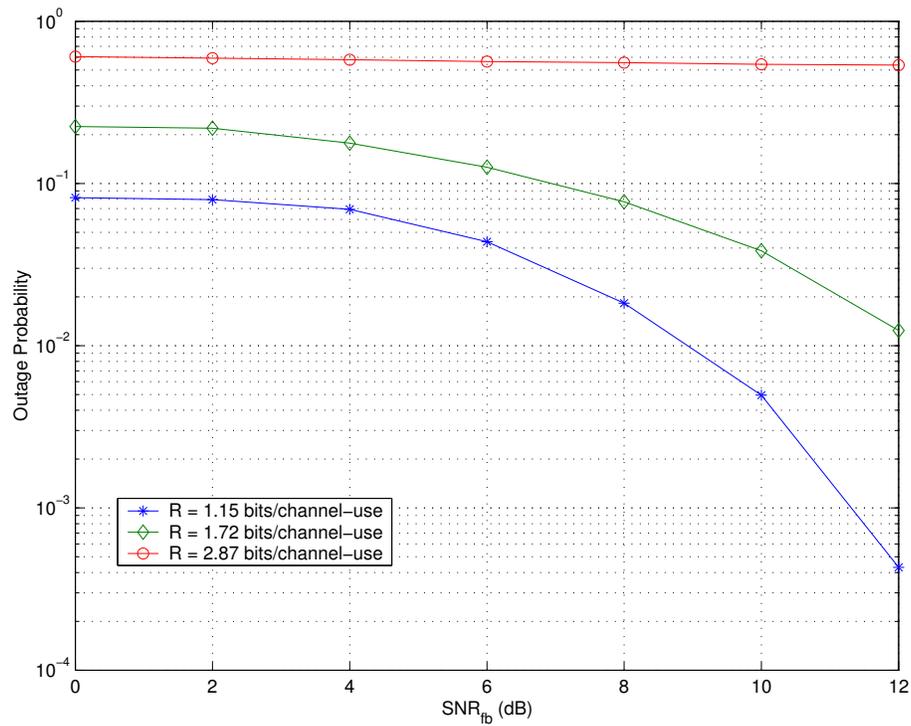
Fig. 6. Effect of SNR_{fb} on achievable minimum outage probability of mean-feedback for $N_t = 2$. Channel received SNR, $\mathrm{SNR}_{rv} \triangleq \frac{P(\xi+\sigma_h^2)}{\sigma_n^2}$, is fixed at 8dB. $\mathrm{SNR}_{fb} \triangleq \frac{\xi}{\sigma_h^2}$.
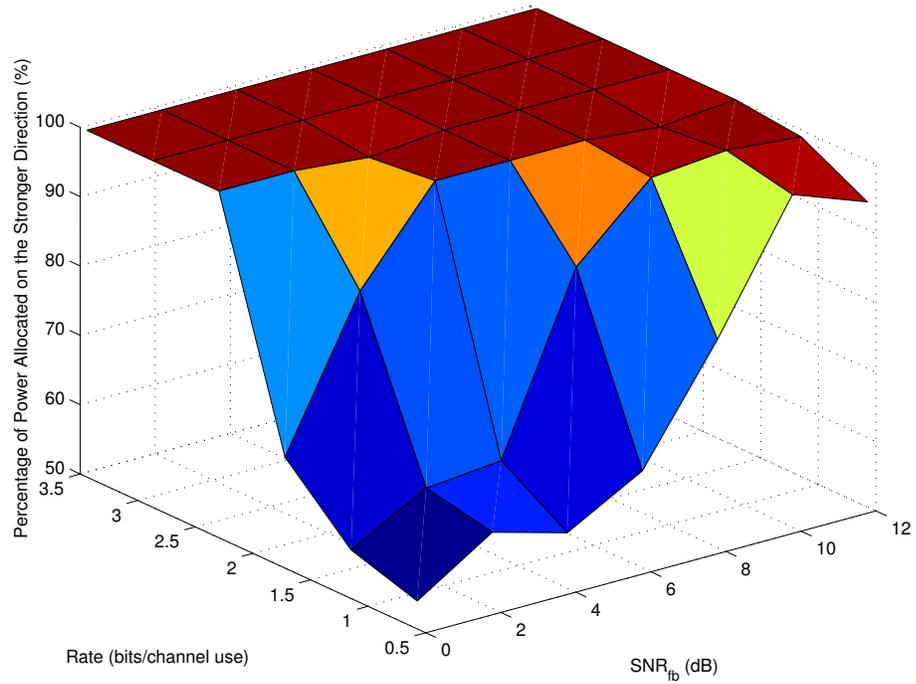
Fig. 7. Optimal power allocation over different transmission directions of mean-feedback for $N_t = 2$. Channel received SNR, $\text{SNR}_{rv} \triangleq \frac{P(\xi + \sigma_h^2)}{\sigma_n^2}$, is fixed at 8dB. $\text{SNR}_{fb} \triangleq \frac{\xi}{\sigma_h^2}$.
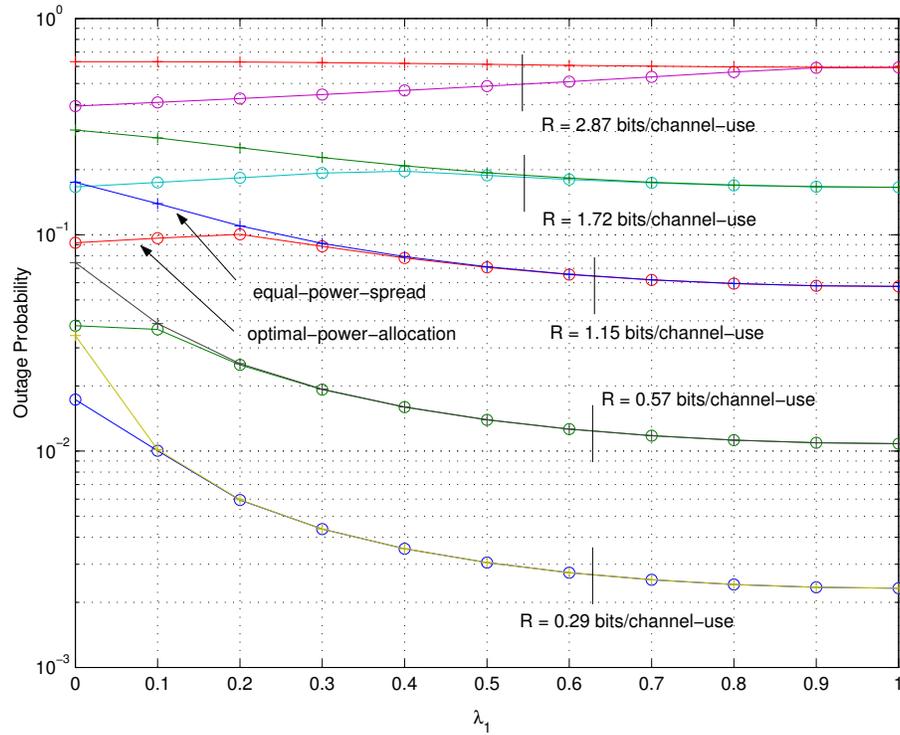
Fig. 8. Minimum outage probability achievable by optimal and equal power allocation over the directions indicated by the two eigenvectors of covariance matrix Σ. Note that the two eigenvalues, λ_1 and λ_2 satisfy $\lambda_1 + \lambda_2 = 2$. Channel received SNR, $\text{SNR}_{rv} \triangleq \frac{P}{\sigma_n^2}$, is fixed at 8dB.

E.   Conclusion

We have studied the problem of minimum outage probability transmission for a MISO fading channel in the cases of mean feedback and covariance feedback. In the case of mean feedback, the optimal transmission strategy is proven to be transmitting several independent data streams in the direction of the channel mean vector and its orthogonal directions. When SNR_{fb} is high, the optimal strategy tends to beamform to the direction indicated by the channel mean. The quality of the channel information, measured by SNR_{fb} affects the outage probability significantly. For both mean and covariance feedback, we show that the optimum power allocation scheme which minimizes outage probability is closely related to the target rate. It is more desirable to spread the power over all transmission directions than beamforming to a single direction for sufficiently small target rates.

CHAPTER V

OPTIMAL BANDWIDTH ALLOCATION FOR THE DATA AND FEEDBACK
CHANNELS IN MISO-FDD SYSTEMS

A.   Introduction

Perfect channel state information at the transmitter (CSIT) has been shown to im-
prove significantly the performance of many wireless systems. For example, at high
SNR, the ergodic capacity of a $N_t \times 1$ multi-input, single-output (MISO) system with
perfect CSIT is approximately $\log_2(N_t)$ bits larger than that of the same system with
only perfect CSI at the receiver [1]. Perfect CSIT also significantly improves the chan-
nel outage capacity for both single and multiple antenna systems [28, 36]. However,
perfect CSIT can be too optimistic in practice. In frequency division duplex (FDD)
systems, CSI is usually estimated by the receiver and then fed back to the transmitter
through a reliable link, which inevitably requires additional bandwidth and power.
If one views bandwidth and power as common resources that can be shared by the
data and feedback channels, the question is whether the increased capacity is worth
the penalty paid for it. A recent paper studied the problem of how much training is
needed to estimate CSI at the receiver in a MIMO system [47]. To our knowledge,
as far as CSIT is concerned, most literature simply ignores the feedback penalty and
considers the optimization of the data transmission and channel feedback separately.
In this chapter, we study the problem of optimal bandwidth allocation between the
data channel and the feedback channel that maximizes the average throughput of the
data channel in a MISO system. Our solution uses the beamforming scheme [37, 48]
as the performance metric of the data channel and considers two models of the partial
CSIT: the noisy CSIT that models CSI as jointly Gaussian distributed with the ac-

tual channel state, and the quantized CSIT. In the first model, we use distortion-rate theory to relate the CSIT accuracy to the feedback bandwidth. In the second model, we propose a lower bound on the capacity of the data channel based on the ensemble of a set of random quantization schemes. In both cases, we can explicitly formulate the achievable rate using a beamforming scheme in the data channel as a function of the number of feedback bits, and therefore solve for the optimal bandwidth allocated to the feedback channel.

In the rest of this chapter, Section B introduces the FDD system model assumed, including the MISO fading data channel, the error-free feedback channel and the assumption on channel prediction. Section C formulates the problem of joint optimal bandwidth allocation for the two models of partial CSIT introduced above. Section D provides numerical results and Section E concludes.

B. The Channel Model

Consider a FDD system consisting of a forward data channel and a CSI-feedback channel, each using a portion of the total bandwidth, W_{tot} Hz. We also assume that W_d and W_f are well separated, therefore the data channel and the feedback channel are uncorrelated. Assuming both channels fully use W_{tot} with ideal pulse shaping, we have

$$W_{tot} = W_d + W_f, \qquad (5.1)$$

where W_d and W_f denote the bandwidths allocated to the data channel and the feedback channel, respectively. Assuming these are passband bandwidths, this implies ISI-free baud-rates for the data and feedback links equal to the corresponding bandwidths. Thus, in the sequel, we will also refer to W_d and W_f as the baud-rates

of the data and feedback channel, respectively. We describe the model of the data channel and the model of channel prediction and CSI feedback in the following two sub-sections.

1. Data Channel

After matched-filtering and sampling at the symbol rate at the receiver, we assume an $N_t \times 1$ discrete MISO frequency flat fading model for the data channel as:

$$y(k) = \mathbf{x}(k)^H \mathbf{h}(k) + w(k), \tag{5.2}$$

where N_t is the number of transmit antennas, $w(k) \sim N(0, \sigma_n^2)$ is circularly symmetric complex white Gaussian noise and $\mathbf{h}(k)$ is the $N_t \times 1$ zero-mean circularly symmetric complex Gaussian channel vector at time k; $\mathbf{x}(k)$ is an $N_t \times 1$ complex vector denoting the transmit signal vector at time k, with power constraint $\mathrm{tr}(\Sigma_x) \leq P$, where $\Sigma_x \triangleq \mathbf{E}[\mathbf{x}(k)\mathbf{x}(k)^H]$. Note that we assume P is a constant which does not adapt to the channel $\mathbf{h}$. Although water-filling type optimal power control could achieve a higher average throughput than the fixed-power scheme, the gain in our case is insignificant at the considered SNR levels [4]. We assume there is no spatial correlation between antennas, which implies $\Sigma_h \triangleq \mathbf{E}[\mathbf{h}(k)\mathbf{h}(k)^H] = \sigma_h^2 \mathbf{I}$. Denoting the ith element of $\mathbf{h}(k)$ as $h_i(k)$ and the maximum Doppler frequency as f_d, we assume that $\{h_i(k)\}$ over the discrete time domain is stationary with its associated continuous time process bandlimited to f_d, and has the same power spectrum for all i.

Based on the feedback, the transmitter establishes some, but not perfect CSI and uses it to optimize transmission for the data channel. However, the optimal scheme which achieves the ergodic capacity with partial CSIT is generally unknown. Even though capacity can be computed numerically in some special cases, it cannot be expressed explicitly to allow joint optimization of the data channel and feedback

channel. In the sequel we assume a simplified system where a beamforming scheme (to a single spatial direction) is used for the data channel [15, 48]. Denoting the unit-energy weight vector at time k as $\mathbf{u}(k)$, the transmitted signal can be expressed as $\mathbf{x}(k) = s(k) \cdot \mathbf{u}(k)$, where $s(k)$ is a scalar signal determined by the source information bits with $\mathbf{E}[s^*(k)s(k)] = P$. Therefore, the maximum average achievable rate per channel use, C_b, associated with this beamforming scheme is

$$C_b = \mathbf{E}\left[\log_2\left(1 + \frac{P|\mathbf{u}^H(k)\mathbf{h}(k)|^2}{\sigma_n^2}\right)\right], \tag{5.3}$$

where the expectation is over the joint distribution of $\mathbf{h}(k)$ and $\mathbf{u}(k)$. Again, we note that (5.3) may not yield the capacity since the beamforming scheme in this case may not be optimal [37]. For the case when the partial CSI and the actual channel state are jointly Gaussian, [48] has shown that transmitting to multiple orthogonal directions rather than beamforming to one direction may be required to achieve capacity in some cases. However, the capacity loss due to beamforming is usually small with reasonably accurate CSIT. We will also show later that a significant portion of the capacity of the perfect CSIT can be achieved by a beamforming scheme with a few bits of feedback per transmit antenna, without the Gaussian assumption. Therefore, the optimal fraction of bandwidth allocated to the feedback channel derived based on the beamforming scheme should be close to that derived based on the capacity achieving scheme.

2.   Channel Prediction and CSI Feedback

To distinguish from the discrete time index k of the forward data channel, we use n to denote the discrete time index of the feedback channel. Note that k and n correspond to different baud-rates (bandwidths) W_d and W_f, respectively. We assume that the receiver has perfect causal CSI about $\mathbf{h}(k)$ and the feedback channel is error-free

with a capacity of K bits/channel-use. We use a quantization codebook of B bits to describe the channel state at a given time and assume that the transmitter cannot update the CSI until it receives all B bits. Obviously, for a given capacity K, we should choose $B = \tau K$ to fully utilize the channel capacity, where $\tau \geq 1$ is an integer denoting the number of channel uses. In the case of $\tau > 1$, the quantization bits are fed back by τ channel-uses.

Since feedback delay is inevitable, the feedback CSI may become outdated in a fast fading scenario, and could significantly degrade system performance if it is used directly to adapt channel transmission at time n. To improve the accuracy of the CSI, channel prediction based on past observations can be used either at the transmitter or/and at the receiver according to the requirements of different applications. Here, we consider a simple prediction scheme as follows. At time $n - \tau$, the receiver predicts $\mathbf{h}(n)$ using the past CSI up to $n - \tau$, quantizes it and feeds it back to the transmitter through τ channel-uses. Thus, at time n, the transmitter receives the quantized version of the predicted $\mathbf{h}(n)$ and uses it to select an optimal beamforming vector, which remains unchanged until an update of CSI is received at time $n + \tau$. So, the channel information is actually updated at the rate of $\frac{W_f}{\tau}$. If $\frac{W_f}{\tau} > 2f_d$, the discrete fading process associated with the sampling rate of $\frac{W_f}{\tau}$ becomes a band-limited process; thus, perfect prediction with arbitrary small mean-square error is possible if the prediction filter's energy and length are unconstrained [49]. Therefore, in the subsequent discussion, we assume $W_f = 2\gamma\tau f_d$, where γ is a scaling factor, whose value is greater than 1 and depends on practical system considerations. With this choice of W_f, we further assume $\mathbf{h}(n)$ can be perfectly predicted at time $(n - \tau)$ at the receiver. The introduction of γ makes our prediction model more flexible and practical. A larger value of γ could translate into a less stringent requirement on the prediction filter design in practical systems. Since the data rate is usually much

larger than the feedback rate, multiple data symbols could be transmitted using the same beamforming vector. A relatively large value of γ also ensures that the selected beamforming vector remains optimal for the data channel before it is updated.

C.   Optimal Bandwidth Allocation

Based on the channel model described above, we now study the problem of optimal bandwidth allocation in terms of maximizing the average throughput for the data channel with a beamforming scheme. Assuming W_{tot}, f_d, K and γ are fixed parameters of the system, we can compute the required feedback bandwidth as a function of B as:

$$W_f = \frac{2\gamma f_{nd}B}{K}W_{tot}. \tag{5.4}$$

where $f_{nd} \overset{\triangle}{=} \frac{f_d}{W_{tot}}$ is the normalized maximum Doppler frequency with respect to the total available bandwidth. Let R_d (bits/second) denote the average throughput of the data channel. Since W_{tot} can be arbitrary, we are only interested in the normalized rate $C_d \overset{\triangle}{=} \frac{R_d}{W_{tot}}$, which is the average achievable rate of the data link per unit total bandwidth and can be expressed as

$$C_d(B) = \left(1 - \frac{W_f}{W_{tot}}\right) C_b(B) = (1 - \eta B) \, C_b(B), \tag{5.5}$$

where $\eta \overset{\triangle}{=} \frac{2\gamma f_{nd}}{K}$ and $C_b(B)$ is the maximum average achievable rate in bits/channel-use using beamforming, as a function of B. Large values of B improve the accuracy of CSIT, but also reduce the available bandwidth for data transmission. Thus, there is an optimum value of B that maximizes C_d.

In the rest of this section we will explicitly derive $C_b(B)$ and solve for the optimal number of feedback bits, B_{opt}, using two different models of partial CSIT as proposed

in [37]. The optimal portion of bandwidth allocated to the feedback channel, α_{opt}, can then be computed as $\alpha_{opt} = \eta B_{opt}$.

1.   Noisy Side Information

In this model, the transmitter uses the feedback bits to establish an estimate $\hat{\mathbf{h}}(n)$ of $\mathbf{h}(n)$, which is assumed jointly Gaussian distributed with the actual channel state $\mathbf{h}(n)$. Assuming $\hat{h}_i(n)$ is i.i.d. for different i ($1 \leq i \leq N_t$), we use $\sigma_{\hat{h}}^2$ to denote the variance of each $\hat{h}_i(n)$, and $\rho \triangleq \mathbf{E}[h_i(n)\hat{h}_i(n)^*]/\sigma_{\hat{h}}\sigma_h$ to denote the correlation between $h_i(n)$ and $\hat{h}_i(n)$. Conditioned on $\hat{h}_i(n)$, the transmitter can model the ith channel as $h_i(n) \sim N(\frac{\sigma_h}{\sigma_{\hat{h}}}\rho\hat{h}_i(n), \sigma_{\epsilon}^2)$, where σ_{ϵ}^2 is the variance of the estimation error and satisfies $\sigma_{\epsilon}^2 = \sigma_h^2(1 - |\rho|^2)$.

In the following discussion, we will simplify the notation by ignoring the time index. Given $\hat{\mathbf{h}}$, the beamforming vector $\mathbf{u}$ is chosen as $\mathbf{u} = \frac{\hat{\mathbf{h}}}{||\hat{\mathbf{h}}||}$, where $|| \cdot ||$ denotes Euclidean norm. Conditioned on $\hat{\mathbf{h}}$, the composite gain $t \triangleq \mathbf{u}^H \mathbf{h}$ due to beamforming and channel fading is circularly complex Gaussian distributed as $t \sim N\left(\frac{\sigma_h}{\sigma_{\hat{h}}}\rho||\hat{\mathbf{h}}||, \sigma_{\epsilon}^2\right)$. Therefore, the average achievable rate (bits/channel use) using beamforming and assuming Gaussian input is

$$
\begin{aligned}
C_b(B) &= \mathbf{E}\left[\log_2\left(1 + |t|^2 \frac{P}{\sigma_n^2}\right)\right] &\text{(5.6)} \\
&= \mathbf{E}\left\{\mathbf{E}\left[\log_2\left(1 + \lambda\frac{P\sigma_h^2}{\sigma_n^2}\right) \mid \hat{\mathbf{h}}\right]\right\} \\
&= \int_0^\infty p(\kappa) \int_0^\infty \log_2\left(1 + \lambda\frac{P\sigma_h^2}{\sigma_n^2}\right) p(\lambda|\kappa) d\lambda d\kappa &\text{(5.7)}
\end{aligned}
$$

where $\lambda \triangleq \frac{|t|^2}{\sigma_h^2}$ denotes the normalized composite power channel gain and is a non-central χ^2-distributed random variable with 2 degrees of freedom, conditioned on $\hat{\mathbf{h}}$. The non-centrality parameter $\kappa \triangleq \frac{||\rho\hat{\mathbf{h}}||^2}{\sigma_{\hat{h}}^2}$ itself is a central χ^2-distributed random variable with $2N_t$ degrees of freedom. $p(\kappa)$ and $p(\lambda|\kappa)$ denote the probability density

function (PDF) of κ and the conditional PDF of λ given κ, respectively. Here we can see that more accurate CSIT (equivalently larger ρ) helps to change the distribution of the normalized composite channel power gain λ to one with a larger mean. For the extreme cases, when there is no CSIT, λ is central χ^2-distributed with 2 degrees of freedom and a mean equal to $\mathbf{E}[\lambda] = 1$; when there is perfect CSIT, λ is central χ^2-distributed with $2N_t$ degrees of freedom and a mean of $\mathbf{E}[\lambda] = N_t$.

Now, we need to relate $\hat{\mathbf{h}}$ to the number of feedback bits B. It is natural to define the mean-squared-error (MSE) $\frac{1}{N_t}\mathbf{E}[||\mathbf{h} - \hat{\mathbf{h}}||^2]$ as the distortion measure due to the Gaussian assumption introduced above. Since $\mathbf{h}$ is complex Gaussian distributed, we can use the distortion-rate function to bound the mean-squared-error as [5][37]

$$\sigma_\epsilon^2 = \sigma_h^2(1 - |\rho|^2) \geq D(R) = \sigma_h^2 2^{-B/N_t}. \tag{5.8}$$

Note that the lower bound is achievable only when N_t goes to infinity. However, we simply use the lower bound here to relate CSIT to B even when N_t is small. It is easy to show that ρ as a function of B in this case satisfies

$$|\rho(B)|^2 = 1 - 2^{-B/N_t}. \tag{5.9}$$

Note that we only use a beamforming scheme with a fixed power over time at the transmitter, which only requires knowledge of the vector channel direction; thus, the feedback information regarding the vector channel norm, $||\mathbf{h}||$, is not utilized by the scheme.

Combining (5.5), (5.6) and (5.9), the problem of maximizing $C_d(B)$ in (5.5) with respect to B can then be formalized as follows:

$$\max_B C_d(B) = \max_B (1 - \eta B)\int_0^\infty p(\kappa; B)\int_0^\infty \log_2(1 + \lambda\frac{P\sigma_h^2}{\sigma_n^2})p(\lambda; B|\kappa)d\lambda d\kappa \tag{5.10}$$

where we have included explicitly in the arguments of the density functions above

their dependence on B (through the normalized correlation $|\rho(B)|$.)

2. Quantized Side Information

The optimization problem in (5.10) assumes a system that achieves the distortion-rate bound on the feedback channel. We now consider a somewhat more practical case when the CSIT is represented by B quantization bits of the actual channel state $\mathbf{h}$; but in contrast to the previous model, we do not assume the CSIT and the actual channel state are jointly Gaussian. We assume a quantization codebook that consists of $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_N\}$, where $N = 2^B$ is the size of the codebook. Based on a maximum channel throughput criterion, the quantizer divides the space of channel vectors $\mathbf{h}$ into N regions $\{\Phi_1, \Phi_1, ..., \Phi_N\}$ defined as:

$$\Phi_i \triangleq \left\{ \mathbf{h} : \ \log_2 \left(1 + \frac{P||\mathbf{h}||^2}{\sigma_n^2} |\mathbf{v}_h^H \mathbf{u}_i|^2 \right) > \log_2 \left(1 + \frac{P||\mathbf{h}||^2}{\sigma_n^2} |\mathbf{v}_h^H \mathbf{u}_j|^2 \right), \ \forall j \neq i \right\}$$

(5.11)

where $\mathbf{v}_h \triangleq \frac{\mathbf{h}}{||\mathbf{h}||}$ is a unit vector in the channel direction. Since power adaptation to $||\mathbf{h}||^2$ is not allowed here, quantization of $\mathbf{v}_h$ instead of $\mathbf{h}$ is enough. If channel state $\mathbf{h}$ belongs to region Φ_i, then $\mathbf{u}_i$ is selected as the beamforming vector and communicated to the transmitter by transmitting its corresponding B bits through the feedback channel. Since the logarithm is a monotonic function, the quantization process can be simplified as choosing the output point $\mathbf{u}_i$ that maximizes $|\mathbf{v}_h^H \mathbf{u}_i|$. This will be referred to as the *maximum absolute inner-product criterion* in later discussions.

The optimal quantization scheme for a given N in general is not known. The Lloyd-Max algorithm can be used to find a locally optimal set of $\{\Phi_i\}$ and its corresponding $\{\mathbf{u}_i\}$ numerically [5, 37]. However, to perform a joint design, we need to explicitly relate the capacity of the data channel to the size of the quantization

codebook (we used the distortion-rate function in the previous case to do this). To overcome this problem, we consider here the average rate achievable by the ensemble of a set of randomly selected quantization schemes. The obtained average data rate can serve as both a lower bound to the capacity and a performance benchmark for practical quantization schemes because there exists at least one practical quantization scheme which can achieve a rate the same as or higher than the ensemble average rate. We refer to this lower bound as the *random quantization lower bound* in subsequent discussions.

We first introduce some notation and preliminary results to be used later. We define the unit hypersphere Ω in the $N_t > 1$ dimensional complex vector space $\mathbb{C}^{N_t}$ as $\Omega \triangleq \{\mathbf{v} : ||\mathbf{v}||^2 = 1, \mathbf{v} \in \mathbb{C}^{N_t}\}$. The surface area of Ω can be computed as [50, 51]:

$$A(\Omega) = \frac{2\pi^{N_t}}{(N_t - 1)!}. \tag{5.12}$$

We say $\mathbf{v}$ and $\mathbf{v}'$ are of square norm of inner product (SNIP) r if $|\mathbf{v}^H \mathbf{v}'|^2 = r$. For a given $\mathbf{v}$, we define the surface of a spherical cap as $\Psi(\mathbf{v}, r) \triangleq \{\mathbf{v}' : |\mathbf{v}^H \mathbf{v}'|^2 \geq r, \mathbf{v}' \in \Omega\}$, where $0 \leq r \leq 1$. Denoting the area of $\Psi(\mathbf{v}, r)$ as $A(r)$, which is independent of the center point $\mathbf{v}$, we have [50]:

$$A(r) = \frac{2\pi^{N_t}(1 - r)^{N_t - 1}}{(N_t - 1)!}. \tag{5.13}$$

It is easy to show that the channel direction $\mathbf{v}_h$ is uniformly distributed on Ω. The random vector quantizer can be viewed as a mapping of the unit hypersphere Ω into a set of N output unit vectors which are chosen independently according to the uniform distribution over Ω. Therefore, for a given $\mathbf{v}_h$, the probability of an output point being within $\Psi(\mathbf{v}_h, r)$, defined as $P(r)$, can be computed as:

$$P(r) = \frac{A(r)}{A(\Omega)} = (1 - r)^{N_t - 1}. \tag{5.14}$$

For the entire quantizer, the probability that no output point is within $\Psi(\mathbf{v}_h, r)$, or the probability that the maximum SNIP is less than r is $(1 - P(r))^N$. In other words, the cumulative distribution function (CDF) of $z \triangleq \max_i |\mathbf{v}_h^H \mathbf{u}_i|^2$ is

$$F(z) = \left(1 - (1 - z)^{N_t - 1}\right)^N, \ 0 \le z \le 1. \tag{5.15}$$

Defining the probability density function (PDF) $f(z)$ as the derivative of $F(z)$, we can evaluate the average achievable rate of the beamforming scheme over both the distribution of channel state and the ensemble of the uniformly distributed random $N = 2^B$ level quantizers (after integration by parts) as:

$$
\begin{aligned}
C_b(B) &= \int_0^\infty f_{||\mathbf{h}||^2}(t) \int_0^1 \log_2\left(1 + \frac{Pt}{\sigma_n^2} z\right) f(z) dz dt \\
&= \int_0^\infty f_{||\mathbf{h}||^2}(t) \log_2\left(1 + \frac{Pt}{\sigma_n^2}\right) dt - \log_2(e) \int_0^\infty f_{||\mathbf{h}||^2}(t) \int_0^1 \frac{Pt}{\sigma_n^2 + Ptz} F(z) dz dt
\end{aligned}
\tag{5.16}
$$

where $f_{||\mathbf{h}||^2}(t)$ denotes the PDF of $||\mathbf{h}||^2$, which is central χ^2-distributed with $2N_t$ degrees of freedom. Note that the first term on the RHS of equation (5.16) is the ergodic capacity with perfect CSIT. The second term, which is an increasing function of N_t and a decreasing function of N (and thus of B), can be viewed as the average capacity loss due to non-perfect CSIT. In particular, at high SNR, the capacity loss in the case of $N_t = 2$ can be easily shown to be $\frac{\log_2(e)}{N}$. Substituting $C_b(B)$ into equation (5.5), we can formulate the joint optimization problem for this case.

D. Simulation Results

We first evaluate the random quantization lower bound in (5.16) for different N_t at SNR of 10 dB in Fig. 9, where SNR is defined as $\frac{P\sigma_h^2}{\sigma_n^2}$. The achievable rate is normalized by the capacity of perfect CSIT. We show that 95% of the capacity of a
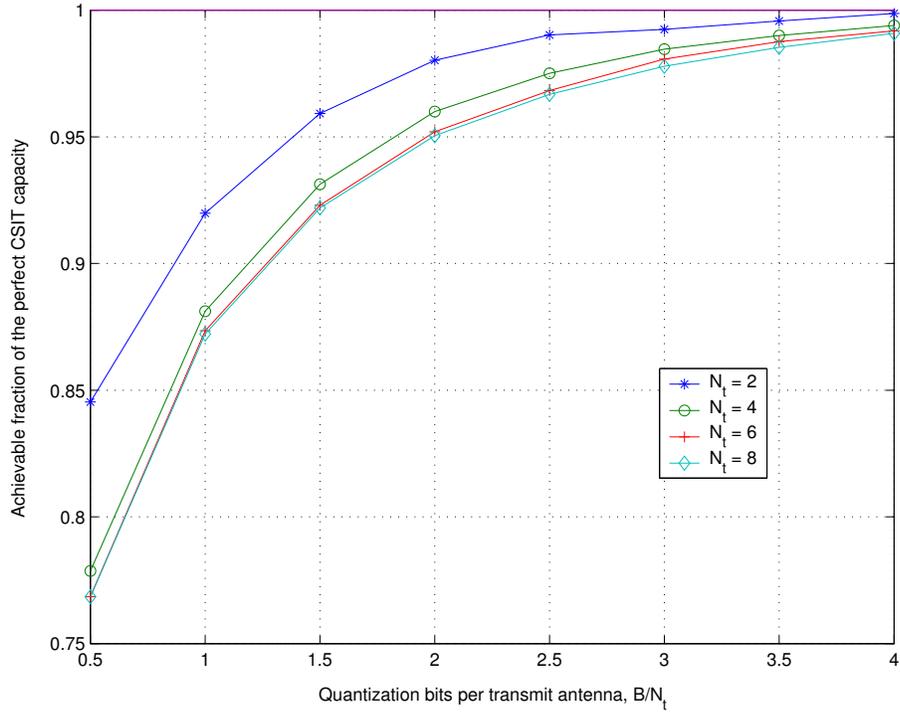
Fig. 9. The random quantization lower bound to the achievable rate for the beam-forming scheme vs $\frac{B}{N_t}$ for different N_t. Channel SNR, defined as $\frac{P\sigma_h^2}{\sigma_n^2}$, is fixed at 10 dB.

perfect CSIT MISO system of $N_t = 2, 4, 6, 8$ can be achieved by a scheme using two bits of feedback per transmit antenna at SNR of 10 dB. At a given $\frac{B}{N_t}$, the largest portion of the capacity with perfect CSIT is achieved when $N_t = 2$. The performance tends to change little for $N_t \geq 4$.

Next, we compare the derived random quantization lower bound with practical quantization schemes in the interesting case of $N_t = 4$. Since we could not find practical quantization schemes for a wide range of feedback bits, B, in the current literature, we constructed the following three classes of schemes denoted as **S1**, **S2** and **S3**, respectively, for different B.

- **S1**: Assume B is a multiple of 3. Normalize $\mathbf{h}$ by h_1, and uniformly quantize

the phases of the complex scalars h_2/h_1, h_3/h_1 and h_4/h_1 each with $B/3$ bits, respectively. The result of the quantization will be of the form $\frac{1}{2}[1, e^{j\theta_1}, e^{j\theta_2}, e^{j\theta_3}]$, where θ_1, θ_2 and θ_3 denote the independently quantized phases of h_2/h_1, h_3/h_1 and h_4/h_1, respectively.

- **S2**: Assume $B = 4, 6, 8, 10$. Two out of B bits are used to indicate the removal of the element in $\mathbf{h}$ with the least power. The remaining three elements form a new 3×1 vector $\mathbf{h}'$, which is then normalized to $\frac{\mathbf{h}'}{h_1'}$. Uniformly quantize the phase of h_2'/h_1' and h_3'/h_1' independently each with $(B-2)/2$ bits. This scheme is motivated by the idea that one of the four paths will most probably fade with power significantly lower than the rest. So this weakest path should not be used. If the first element is removed, the quantization result will be something like $\frac{1}{\sqrt{3}}[0, 1, e^{j\theta_1}, e^{j\theta_2}]$, where θ_1 and θ_2 denote the quantized phases of h_2'/h_1' and h_3'/h_1', respectively.

- **S3**: In the case of $B = 2$, choose any set of orthonormal bases to form the codebook. In the case of $B > 2$, randomly generate a set of 2^{B-2} 4×4 unitary matrices $\mathbf{U}_i$, $i = 1, 2, ..., 2^{B-2}$. The codebook consists of all the columns of the matrix $\mathbf{U} \overset{\triangle}{=} [\mathbf{U}_1 \ \mathbf{U}_2 \ ...\mathbf{U}_{2^{B-2}}]$ as the quantizer outputs. The maximum absolute inner product criterion is used to select the quantization output.

Fig. 10 plots the achievable rates of the three classes of schemes and the random quantization lower bound for $N_t = 4$ at SNR of 10 dB. Here in the cases of $\frac{B}{N_t} = 0.75$, $\frac{B}{N_t} = 1$ and $\frac{B}{N_t} = 2$ in class **S3**, the achievable rate is associated with the best of one hundred randomly generated codebooks. Except for the scheme of $\frac{B}{N_t} = 0.75$ in **S1**, all the other schemes in class **S1** and class **S2** cannot reach the performance of the random quantization lower bound. All schemes in class **S3** outperform the random quantization lower bound. This is because the maximum norm of inner
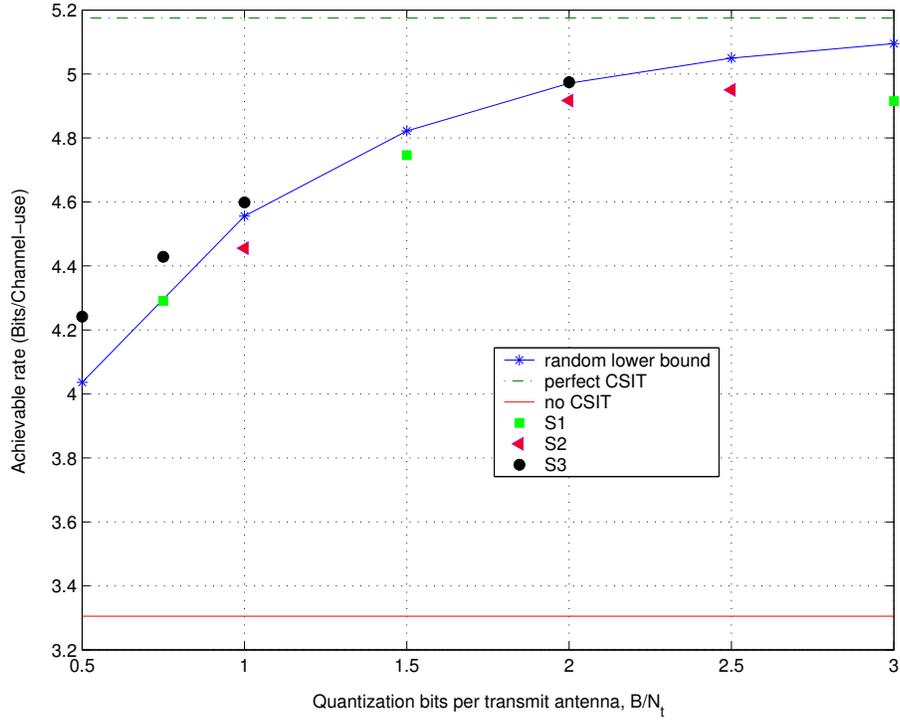
Fig. 10. Achievable rates of practical quantization schemes for $N_t = 4$ and SNR=10dB.

product quantization criterion is optimal in terms of maximizing the rate while the independent quantization of each element of $\mathbf{h}$ is sub-optimal. However, we should note that the encoding complexity of the schemes of class $\mathbf{S3}$ increases exponentially with B. In addition, as B increases, the gain of the schemes of $\mathbf{S3}$ over the random quantization lower bound decreases as shown in the figure. It can also be easily shown that the scheme of $\frac{B}{N_t} = 0.5$ in class $\mathbf{S3}$ is optimal among all quantization schemes of $B = 2$ for $N_t = 4$, and equivalent to selection diversity. Finally, we note that the achievable rates plotted in both Fig. 9 and Fig. 10 do not consider the cost of feedback.

The joint design problem and its optimal performance is demonstrated in Fig. 11, Fig. 12 and Fig. 13. Fig. 11 plots the maximum achievable data rate per unit total bandwidth, C_d^*, vs $\eta \triangleq \frac{2\gamma f_{nd}}{K}$ with the value of C_d^* normalized by the channel
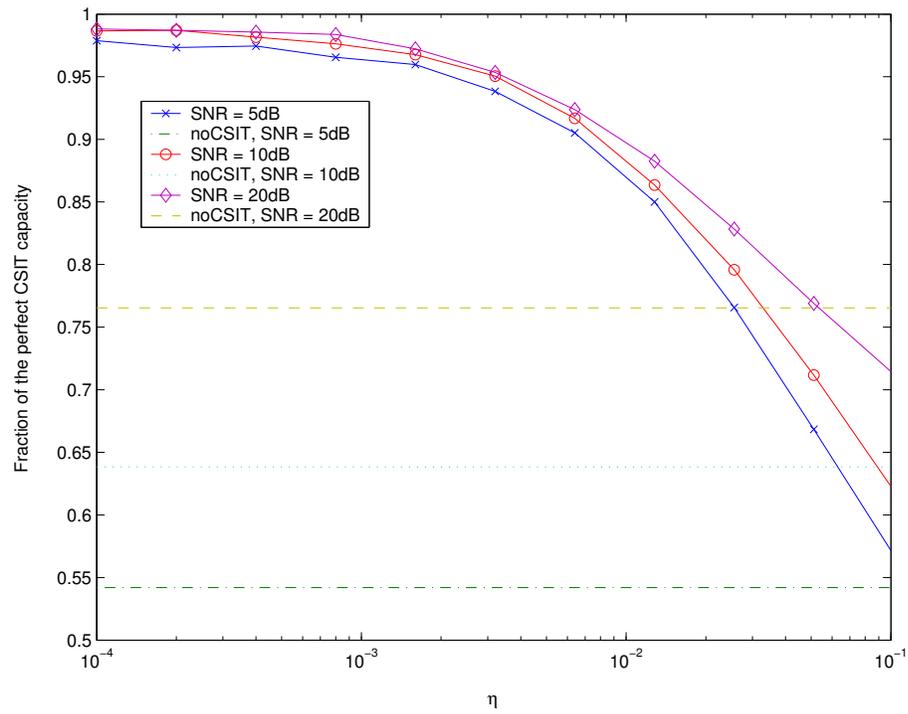
Fig. 11. Maximum achievable rate C_d^*, normalized by the capacity with perfect CSIT obtained using the random quantization lower-bound for different η. $N_t = 4$.

Fig. 12. Optimal fraction of the total bandwidth allocated to the feedback channel for different η in order to achieve C_d^*. $N_t = 4$.

Fig. 13. Comparison of achievable rates assuming two different partial CSIT models: $N_t = 32$, $\eta = 0.005$ and SNR $= 10$ dB.

capacity with perfect CSIT. Fig. 12 plots the optimal fraction of total bandwidth allocated to the feedback channel for d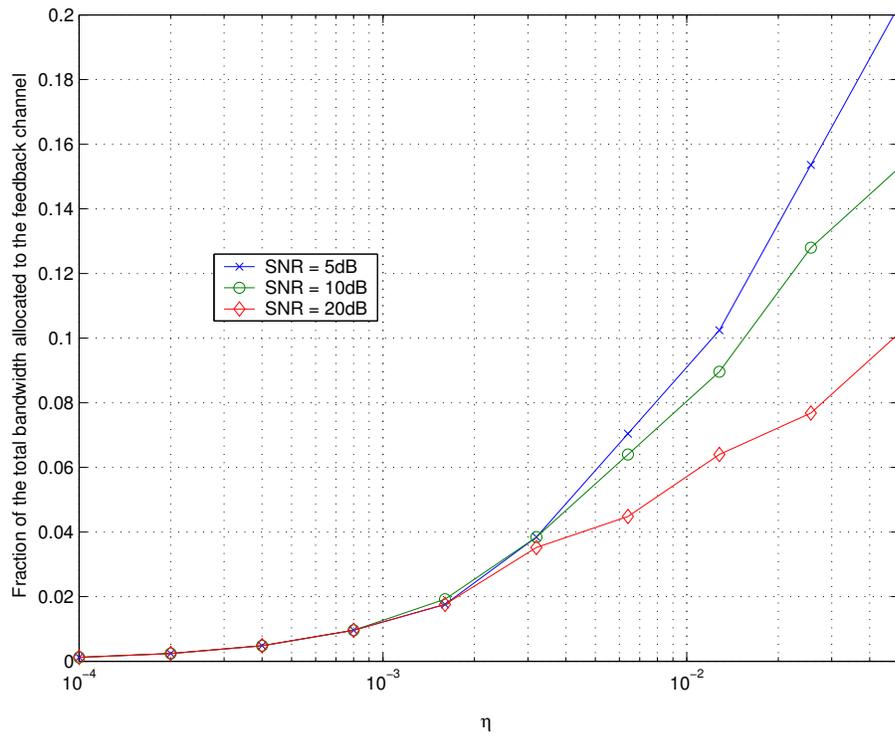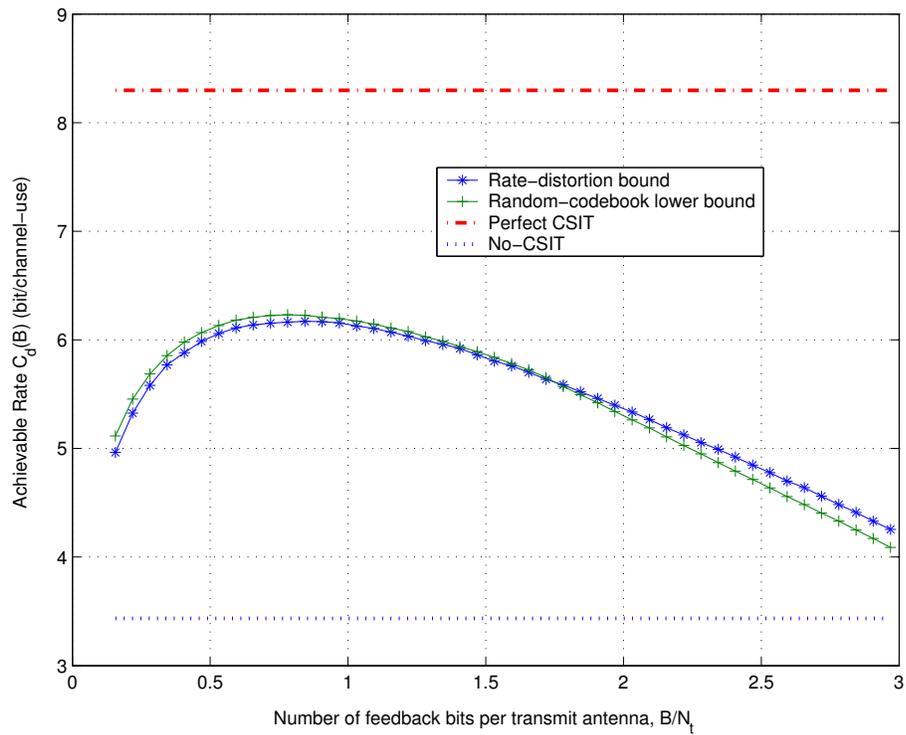ifferent η to achieve the corresponding rate in Fig. 11. In both figures, we assume $N_t = 4$. The random quantization lower bound is used in evaluating the achievable rate of the data channel. These two figures provide information on how to jointly design the FDD MISO system to maximize the average throughput. For example, if uncoded 4-PSK with maximum ratio combining of the N_t antennas is assumed in the feedback channel, a bit-error rate of about 10^{-3} can be achieved at an SNR level of 10 dB [52]. (In this case, the feedback channel is a SIMO fading channel). Therefore, we can reasonably assume $K = 2$. If we further assume $\gamma = 4$ for the prediction assumption to be valid, we have $\eta = 0.04$ if the normalized Doppler frequency $f_{nd} = 0.01$. From Fig. 11, we can achieve a rate of about 74% of the capacity of perfect CSIT at SNR of 10 dB. Fig. 12 suggests that about 14.5% of the total bandwidth should be allocated to the feedback channel to achieve this rate. As SNR increases, a larger portion of the capacity of CSIT can be achieved with a smaller portion of the bandwidth allocated to the feedback channel for the same η. A rate larger than the capacity of no-CSIT (only CSI at the receiver) can be achieved for a wide range of η, which may translate into a wide range of Doppler frequencies for a fixed γ. For example, assume $\gamma = 4$, $K = 2$ and SNR of 10 dB. According to Fig. 11, $\eta \leq 0.09$ is required to achieve a rate larger than the capacity of no-CSIT. This is equivalent to $f_{nd} = \frac{\eta K}{2\gamma} < \frac{0.09}{\gamma} = 0.0225$. If $W_{tot} = 100$ kHz, the tolerable maximum Doppler frequency can be as high as 2.25 kHz.

Fig. 13 compares the achievable rates assuming the two different partial CSIT models vs the number of feedback bits per transmit antenna, $\frac{B}{N_t}$. We assume $N_t = 32$ (the rate distortion bound is tight only when N_t is large), $\eta = 0.005$ and SNR $= 10$ dB. Both models demonstrate similar performance, although strictly speaking the two cases assume different scenarios on what is fed back to the transmitter. The noisy CSI

model shows a slightly lower achievable rate at smaller values of $\frac{B}{N_t}$ because the mean-squared distortion measure implicitly quantizes the norm $||\mathbf{h}||^2$, which is redundant in our setting since we use the beamforming scheme with fixed power allocation over time. Therefore, MSE quantization is less efficient than that based on maximum absolute inner product. However, as B increases, the portion of B used to quantize the norm apparently decreases. Eventually, the noisy CSI model indicates a larger achievable rate than that of the random quantization lower bound. This is because the achievable rate evaluated using the distortion-rate function translates into an upper-bound on the achievable rate of the beamforming scheme.

E.   Conclusion

We have studied the problem of optimal bandwidth allocation between a MISO fading data channel and an error-free feedback channel in a FDD system. We proposed a simple but flexible prediction model based on which the maximum average achievable rate of the beamforming scheme for the data channel and the associated optimal bandwidth allocated to the feedback channel are evaluated under two different assumptions of the partial CSIT. In addition, we proposed a lower bound on the average achievable rate of the beamforming scheme using quantized feedback, which can be used as a performance benchmark for practical channel quantization schemes.

CHAPTER VI

SAGE-AIDED DETECTION OF MULTIPLE TRANSMIT ANTENNA SYSTEMS

A. Introduction

In a MIMO channel, the received signal at each receive antenna is a superposition of transmitted signals from different transmit antennas. If the MIMO system has N_t transmit antennas and uses a constellation of size C, maximum-likelihood (ML) detection, which searches through all the possible transmitted signals, requires a complexity proportional to $O(C^{N_t})$, which is hard to implement when C and N_t are large. Many sub-optimal detectors were proposed in order to reduce the complexity, such as BLAST detection, zero-forcing (ZF) and MMSE detection. However, all these schemes perform fairly far from the ML detection scheme. Recently, the sphere detection algorithm which searches a vicinity of the received signal vector for the optimum solution was proposed [11][12]. The complexity of the proposed sphere detection algorithm is polynomial in N_t for a wide range of system parameters [53]. In this chapter, we propose a sub-optimal detection scheme for the MIMO system based on the Space Alternating Generalized Expectation-Maximization (SAGE) algorithm and the List-BLAST algorithm. The proposed scheme can achieve performance close to that of the ML detection scheme with a complexity of the order of $O(CN_t^2)$ - $O(CN_t^3)$, and a flexible trade-off between complexity and performance.

B.   EM and SAGE for Detecting Superimposed Signals

Consider the discrete model of a MIMO frequency-flat fading channel with N_t transmit antennas and N_r (assume $N_r \geq N_t$) receive antennas:

$$\mathbf{y} = \mathbf{Hx} + \mathbf{w}, \tag{6.1}$$

where $\mathbf{w} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I})$ is a $N_t \times 1$ circularly symmetric complex Gaussian noise vector with variance $\frac{N_0}{2}$ per dimension. $\mathbf{y}$ is the received signal vector of dimension $N_r \times 1$. $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_1, \ldots, \mathbf{h}_{N_t}]$ is a $N_r \times N_t$ MIMO channel with each element modelled as i.i.d. zero mean complex Gaussian with normalized variance. We assume the complex $N_t \times 1$ transmit signal vector $\mathbf{x} = [x_1, x_2, \ldots, x_{N_t}]^T$ satisfy the component-wise power constraint $\mathbf{E}[||x_i||^2] = \frac{E_s}{N_t}$, where E_s is the total transmit energy in each channel use.

Consider the case that X is composed of uncoded QAM or QPSK signals and $\mathbf{H}$ is assumed perfectly known on the receiver side. The maximum-likelihood (ML) detector can be expressed as

$$\hat{\mathbf{x}}_{ml} = \arg \min_{\mathbf{x} \in \Omega^{N_t}} ||\mathbf{y} - \mathbf{Hx}||^2 \tag{6.2}$$

where Ω^{N_t} denotes the set of constellation points in the complex N_t dimensional space. Since the search for the ML solution over the whole set of Ω^{N_t} is too complex to be implementable, we take a different approach using the Expectation Maximization (EM) type algorithm [1].

The EM algorithm was proposed to iteratively solve the maximum-likelihood estimation problem [54]. Feder and Weinstein [55] proposed an EM solution for the general parameter estimation problem from superimposed signals. Fessler and

---

[1]In the rest of the chapter, we refer to the classical EM algorithm simply as EM algorithm. EM-type algorithms always refer to both EM and SAGE.

Hero [56] extended the classical EM algorithm to the SAGE algorithm in which one alternates between several hidden-data spaces rather than using just one "complete" data space and updates only a subset of the elements of the parameter vector in each iteration. They also apply the SAGE algorithm to estimate superimposed signals in Gaussian noise and show that SAGE converges much faster than EM. We have recently applied both methods to channel estimation in a multiple transmit antenna OFDM system [57], which will be treated in detail in the next section. However, we are now facing a detection problem, where the parameter set is discrete.

We first consider the EM algorithm. We view the observed data $\mathbf{y}$ as the "incomplete" data and define the "complete" data $\mathbf{z}_i$ as

$$\mathbf{z}_i = \mathbf{h}_i x_i + \mathbf{w}_i, \quad 1 \leq i \leq N_t, \tag{6.3}$$

where $\sum_{i=1}^{N_t} \mathbf{w}_i = \mathbf{w}$; thus, $\sum_{i=1}^{N_t} \mathbf{z}_i = \mathbf{y}$. We assume $\{\mathbf{w}_i\}$ is i.i.d complex Gaussian distributed with zero mean vector and covariance matrix $\frac{\sigma_n^2}{N_t}\mathbf{I}$. The $N_r \times 1$ vector $\mathbf{z}_i$ can be viewed as the component of the received signal transmitted by the ith antenna through the channel $\mathbf{h}_i$ corrupted by a fraction of the actual noise.

Denote $f(\mathbf{z}; \mathbf{x})$ as the probability density function of $\mathbf{z}$ parameterized by $\mathbf{x}$. Denote $\hat{\mathbf{x}}^{(k)}$ as the estimate of $\mathbf{x}$ in the k th iteration. In the E-step, we compute $U(\mathbf{x}, \hat{\mathbf{x}}^{(k)}) \triangleq \mathbf{E}\{\log f(\mathbf{z}; \mathbf{x})\}$ as expressed below, where the expectation operation $\mathbf{E}\{\cdot\}$ is with respect to the conditional distribution of $f(\mathbf{z}|\mathbf{y}, \hat{\mathbf{x}}^{(k)})$.

$$
\begin{aligned}
U(\mathbf{x}, \hat{\mathbf{x}}^{(k)}) &= d - \mathbf{E}\{\sum_{i=1}^{N_t} ||\mathbf{z}_i - \mathbf{h}_i x_i||^2\} \tag{6.4} \\
&= e + x_i^* \mathbf{h}_i^H \bar{\mathbf{z}}_i + x_i \bar{\mathbf{z}}_i^H \mathbf{h}_i - |x_i|^2 \mathbf{h}_i^H \mathbf{h}_i, \tag{6.5}
\end{aligned}
$$

where d contains constant terms and e contains all terms independent of $\mathbf{x}$. $\bar{\mathbf{z}}_i$ denote the conditional mean of $\mathbf{z}_i$ given $\mathbf{y}$ and $\hat{\mathbf{x}}^{(k)}$. Since $\mathbf{z}_i$ and $\mathbf{y}$ are jointly Gaussian, we

have

$$\bar{\mathbf{z}}_i = \mathbf{h}_i \hat{x}_i^{(k)} + \frac{1}{N_t} \left(\mathbf{y} - \sum_{j=1}^{N_t} \mathbf{h}_j \hat{x}_j^{(k)} \right). \tag{6.6}$$

In the maximization step, we compute

$$\hat{\mathbf{x}}^{(k+1)} = \arg\max_{\mathbf{x}} U(\mathbf{x}, \hat{\mathbf{x}}^{(k)}). \tag{6.7}$$

It can be easily shown that

$$\hat{\mathbf{x}}^{(k+1)} = (\mathbf{h}_i^H \mathbf{h}_i)^{-1} \mathbf{h}_i^H \bar{\mathbf{z}}_i. \tag{6.8}$$

In the above derivation, we did not consider the fact that $\mathbf{x}$ is discrete and x_i $(1 \leq i \leq N_t)$ has to be a constellation point. To force this condition, an intuitive way is to quantize $\hat{x}_i^{(k)}$ to its nearest constellation point in each iteration.

Denote $a_{i,j} \triangleq \mathbf{h}_i^H \mathbf{h}_j / (\mathbf{h}_i^H \mathbf{h}_i)$ and $b_i \triangleq \mathbf{h}_i^H \mathbf{y} / (\mathbf{h}_i^H \mathbf{h}_i)$. Substituting equation (6.6) in (6.8) and considering the quantization process, we can summarize the EM iteration as follows:

$$\tilde{x}_i^{(k+1)} = \hat{x}_i^{(k+1)} + \frac{1}{N_t} [b_i - \sum_{j=1}^{N_t} a_{i,j} \hat{x}_j^{(k)}], \text{ for } 1 \leq i \leq N_t. \tag{6.9}$$

$$\hat{x}_i^{(k+1)} = Q(\tilde{x}_i^{(k+1)}) \tag{6.10}$$

where $Q(\cdot)$ denote the quantization process; $\tilde{x}_i^{(k+1)}$ and $\hat{x}_i^{(k+1)}$ are the unconstrained estimation and the constrained detection of x_i, respectively, in the $(k+1)^{th}$ iteration.

The convergence rate of the EM algorithm is inversely related to the Fisher information of its complete-data space [56]. In the above algorithm, the noise variance is distributed over $\mathbf{z}_i$ for all i; therefore, the Fisher information of $\mathbf{z}_i$ for $\mathbf{x}$ is relatively large for a certain i. To improve the convergence rate, the SAGE algorithm chooses

the hidden data space as $\mathbf{z}_i$ for $i = 1, 2, \cdots, N_t$ alternately in each iteration and, thus, associate all the noise variance with it. Following similar derivation shown above, the SAGE algorithm can be expressed as:

- Initialize with some $\hat{x}_i^{(0)}$ for $1 \leq i \leq N_t$.

- At the $(k+1)^{th}$ iteration ($k = 0, 1, 2, ...$):

  For $i = 1 + [k \bmod N_t]$, compute

$$\hat{x}_i^{(k+1)} = Q \left(\hat{x}_i^{(k+1)} + \left[b_i - \sum_{j=1}^{N_t} a_{i,j} \hat{x}_j^{(k)} \right] \right). \tag{6.11}$$

  For $1 \leq j \leq N_t$ and $j \neq i$,

$$\hat{x}_j^{(k+1)} = \hat{x}_j^{(k)}. \tag{6.12}$$

1. Initialization

A proper selection of the initial value of $\mathbf{x}$ is very important for the convergence of both algorithms. Note that the convergence of the EM algorithm to even a local maximum has not been proved in the case of discrete parameter spaces [58]. In our simulation, we found that the convergence property of EM is very poor. Even in the SAGE case, the algorithm usually stopped in one to three iterations according to our simulations. Therefore, we need to use multiple initial points in order to increase the probability that the iteration will converge to the ML solution. The solutions of ZF, MMSE and BLAST are good candidates as initial points. However, we usually need more in order to improve the performance. Here, we propose a scheme called List-BLAST detection to produce initial points.

Denote the QR decomposition as $\mathbf{H} = \mathbf{QR}$, where $\mathbf{Q}$ is a unitary matrix and $\mathbf{R}$ is an upper triangular matrix. We can perform a linear transformation on the received

signal as $\mathbf{y}' = \mathbf{Q}^H \mathbf{y}$; the system can be expressed as

$$\mathbf{y}' = \mathbf{R}\mathbf{x} + \mathbf{w}', \tag{6.13}$$

where $\mathbf{w}' = \mathbf{Q}^H \mathbf{w}$ has the same distribution as $\mathbf{w}$ since $\mathbf{Q}$ is unitary. In the triangulized model above, each row denotes a different transmission/detection layer with the kth layer interferenced only by layers with indexes larger than k. In BLAST, one first detects $\hat{x}_{N_t}$; assuming $\hat{x}_{N_t}$ is correct, the interference of $r_{N_t-1,N_t}\hat{x}_{N_t}$ can be subtracted from layer $N_t - 1$ and $\hat{x}_{N_t-1}$ can be detected as in a scalar channel. Similarly, layer $N_t - 2, N_t - 3, \cdots, 1$ can be detected in order. In the proposed List-BLAST scheme, we assume $\hat{x}_{N_t}$ could take as values all points in the constellation; for a given $\hat{x}_{N_t}$, we use the BLAST algorithm to detect the remaining elements of the vector $[\hat{x}_{N_t-1}, \hat{x}_{N_t-2}, \dots, \hat{x}_1]$. Therefore, we could list C candidate points, each of which is a vector in the complex N_t dimensional space. Finally, We can select the one which has the minimum Euclidean distance to $\mathbf{y}$ as the detected symbol vector. It can be easily shown that the List-BLAST algorithm for $N_t = 2$ is actually the ML detection algorithm.

We can also use the listed candidates as initial points in the SAGE algorithm, which converge to another set of C points. We then compare these C points and select the one which has the minimum Euclidean distance to $\mathbf{y}$. We refer to this detection scheme as SAGE-aided list-BLAST detection in the sequel.

It is well known that the performance of the BLAST detection can be improved by ordering the sequence of nulling and cancelling. Each different order of nulling and cancelling corresponds to a unique ranking of the columns of the channel matrix $\mathbf{H}$ in the above implementation using QR decomposition. Thus, we can also extend the list-BLAST algorithm as follows.

- List-Ranked-BLAST: In this extension, the worst layer with the lowest signal-to-noise ratio is detected first; the remaining layers are detected from the best (with the highest SNR) to worst. Note that the optimal detection sequence in traditional BLAST detection is to detect from the best layer to the worst.

- List-Shifted-BLAST: In this extension, we cyclicly shift (either right or left) the columns of $\mathbf{H}$ by one, and apply the List-BLAST algorithm as described above to each shifted $\mathbf{H}$. If shifting is performed K times, where $1 \leq K \leq N_t$, we will get $C \times K$ points to initialize the SAGE algorithm. The detected signal vector will be selected from the $C \times K$ SAGE solutions using the minimum distance criterion. A larger value of K results in better performance, as will be shown in the simulation results, but higher complexity. Therefore, the Shifted-List-BLAST algorithm provides a flexible trade-off between complexity and performance. In stead of cyclicly shifting the columns of $\mathbf{H}$, random permuting can also be used in a similar way.

2. Implementation

For complex PSK modulation, sphere decoding can be implemented directly over the N_t dimensional complex space [12]. However, QAM modulation is usually handled by decoupling the real and the imaginary components; thus, the sphere detector may need to search over a $2N_t$ dimensional real space. In contrast, the SAGE-aided List-BLAST schemes solve both QAM and PSK detection in the same fashion.

3. Complexity

Assume a block fading channel. We need to consider the computational complexity for a whole block and that for each vector symbol in the block [59]. We denote the

first kind of complexity as pre-detection complexity and the second kind of complexity as the detection complexity. For SAGE-aided detection, the pre-detection complexity requires $O(N_t^3)$ computations for QR decomposition or pseudo-inverse of the matrix channel $\mathbf{H}$ depending on whether BLAST or ZF is used to produce the initial points. If ordered BLAST is required, the computation complexity is still of $O(N_t^3)$ by using some fast algorithm [14]. Similarly, sphere detection requires computing of both the QR decomposition and pseudo-inverse $\mathbf{H}$ with a complexity of $O(N_t^3)$ [12].

Since the channel remains a constant during each transmission block which could be composed of hundreds of vector symbols, the pre-detection complexity can be very low per vector symbol and the detection complexity dominates. For computation overhead for each vector symbol detection, the List-BLAST and the List-Ranked-BLAST detection require a computation of $O(CN_t^2)$. List-Shifted-BLAST requires complexity of $O(KCN_t^2)$. Each run of SAGE starting from a single initial point requires a complexity of $O(N_t^2)$. Note that most of the time the SAGE algorithm converges in $1-3$ iterations; the number of iterations when $N_t < 8$ is not related to N_t according to our observation in the simulations. Therefore, if we perform SAGE aided detection on top of List-BLAST or List-Ranked-BLAST, the complexity is still at $O(CN_t^2)$. If SAGE aided List-shifted-BLAST is performed with $K = N_t$, the complexity is $O(CN_t^3)$.

The complexity of sphere detection is closely related to SNR and channel realizations. At low SNR, it could require an average complexity of $O((2N_t)^4)$ - $O((2N_t)^{4.5})$ [53]. Some "bad" (with spread singular values of $\mathbf{H}$) channel realizations require more computation. Some recent publications show that sphere detection requires an average complexity of $O((2N_t)^3)$ for complex Gaussian fading channels if a good initial point is selected [60][61]. The complexity of sphere detection also increases with C [61]. However, the exact relation of C to the complexity order is not clear. In

our simulations, we found that sphere detection requires much more complexity than SAGE aided List-BLAST detection schemes.

4.   Soft-output Detection

The List-BLAST type algorithm provides us a natural way to decode and generate soft-information. We assume that the information bits have been encoded with a channel code, randomly interleaved, Gray-mapped to the constellation, and then transmitted through N_t different antennas. Therefore, $N_t M$ coded bits are transmitted per channel use, where $M = \log_2 C$.

At the receiver, MAP joint demodulation and detection can be used. The a posteriori L-value of the coded bits b_k, $k = 0, 1, \ldots, N_t M - 1$, conditioned on the received vector $\mathbf{y}$, is

$$L_D(b_k|\mathbf{y}) = \ln \frac{P[b_k = +1|\mathbf{y}]}{P[b_k = -1|\mathbf{y}]}. \tag{6.14}$$

Assume $\{b_k\}$ are independent due to the random interleaver, Equation (6.14) can be further expressed as:

$$L_D(b_k|\mathbf{y}) = L_A(b_k) + \ln \frac{\sum_{\mathbf{x} \in \mathbb{X}_{k,+1}} P[\mathbf{y}|\mathbf{x}] \cdot \exp \sum_{j \in \mathbb{J}_{k,x}} L_A(b_j)}{\sum_{\mathbf{x} \in \mathbb{X}_{k,-1}} P[\mathbf{y}|\mathbf{x}] \cdot \exp \sum_{j \in \mathbb{J}_{k,x}} L_A(b_j)}. \tag{6.15}$$

where $\mathbb{X}_{k,+1}$ is the set of $2^{N_t M - 1}$ bit vectors $\mathbf{x}$ having $b_k = +1$; $\mathbb{X}_{k,-1}$ is the set of $2^{N_t M - 1}$ bit vectors $\mathbf{x}$ having $b_k = -1$. $L_A(b_j) = \ln \frac{P[b_j=1]}{P[b_j=-1]}$. $\mathbb{J}_{k,x}$ is the set of indices j with

$$\mathbb{J}_{k,x} = \{j|j = 0, 1, 2, N_t M - 1, j \neq k, b_k = 1\}. \tag{6.16}$$

The second term on the RHS of (6.15) is the extrinsic L-value, and is denoted as $L_E(b_k|\mathbf{y})$ to be used later. Since exhaustive listing of $\mathbb{X}_{k,+1}$ and $\mathbb{X}_{k,-1}$ is usually too complex, we can use the List-Shifted-BLAST algorithm together with the SAGE

algorithm to generate a set $\mathbb{L}$ consisted of possible candidates, which can be similarly divided into two sets $\mathbb{L}_{k,+1}$ and $\mathbb{L}_{k,-1}$. Using the max-log approximation, the extrinsic L-value can be approximated as [12]

$$L_E(b_k|\mathbf{y}) \approx \max_{\mathbf{x}\in\mathbb{L}_{k,+1}} \left\{ -\frac{||\mathbf{y}-\mathbf{Hx}||^2}{N_0} + \mathbf{b}_{[k]}^T \mathbf{L}_{A,[k]} \right\} - \max_{\mathbf{x}\in\mathbb{L}_{k,-1}} \left\{ -\frac{||\mathbf{y}-\mathbf{Hx}||^2}{N_0} + \mathbf{b}_{[k]}^T \mathbf{L}_{A,[k]} \right\},$$
$$(6.17)$$

where $\mathbf{b}_{[k]}$ denotes the sub-vector of $\mathbf{b}$ omitting its kth element, and $\mathbf{L}_{A,[k]}$ is the vector of all L_A values, also omitting its kth element. It is more desirable to include both the List-Shifted-BLAST solutions (the initial points in the SAGE algorithm) and the converged points after the SAGE iterations in $\mathbb{L}$ for two reasons. Firstly, the List-Shifted-BLAST algorithm ensures that $\mathbb{L}_{k,+1}$ and $\mathbb{L}_{k,-1}$ will not be a null set due to the exhaustive listing of the constellation points for each transmit antenna. Secondly, the SAGE iteration will produce some candidate points in the vicinity of the received vector [2]. These candidates are more reliable to be used in computing (6.17) using the max-log approximation. We note that the ML solution $\hat{\mathbf{x}}_{ml}$ may not necessarily be the candidate $\mathbf{x}$ which maximizes one of the two terms in the RHS of (6.17), which could be relatively far away from $\mathbf{y}$ due to the fact that turbo-coded system usually operate at very low SNR. Therefore, if one uses a sphere decoder to list the candidates as in [12], the radius of search in the sphere decoder should be much larger than that in the case of high SNR, therefore increasing the complexity of the sphere decoding algorithm.

---

[2]$\mathbf{x}$ is in the vicinity of $\mathbf{y}$ in the sense that $||\mathbf{y}-\mathbf{Hx}|| < \delta$, where δ is small.
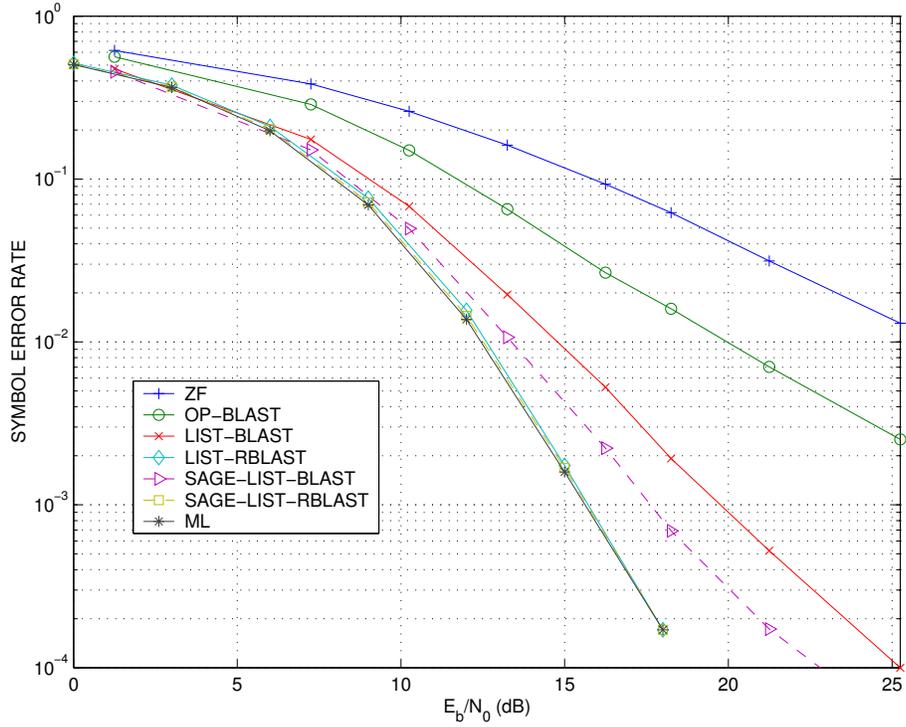
Fig. 14. Symbol Error Rate of different detectors for a 4×4 MIMO system with uncoded 8-PSK modulation.

C.   Simulation Results

In the following simulations, we define E_b as the signal energy per transmitted information bit at the receiver. Thus, we have

$$\frac{E_b}{N_0} = \frac{E_s}{N_0} + 10\log_{10}\cdot\frac{N_t}{RN_rM} \tag{6.18}$$

We conducted simulations to evaluate the performance of the proposed SAGE and the List-BLAST type of algorithms assuming an independently faded MIMO channel in each channel use. Note that for uncoded system, the average symbol error rate (SER) and the bit error rate (BER) of the independently faded MIMO channel are the same as those of the blocked faded MIMO channel. We first consider an un-coded system, in which the channel coding rate $R = 1$. Fig. 14 and Fig. 15 show

Fig. 15. Bit Error Rate of different detectors for a 4×4 MIMO system with uncoded 8-PSK modulation.
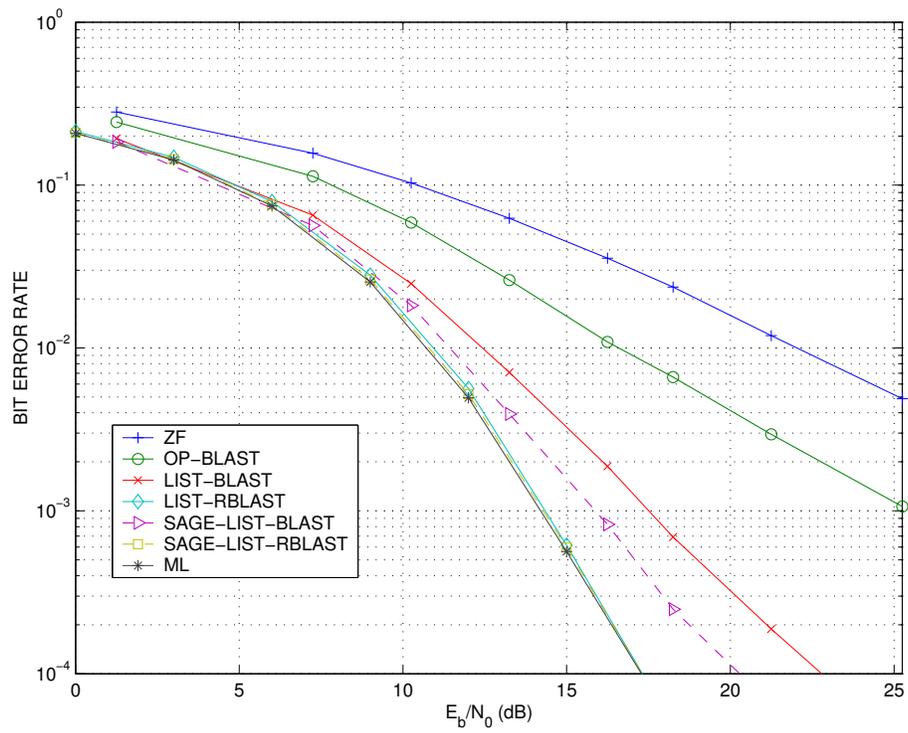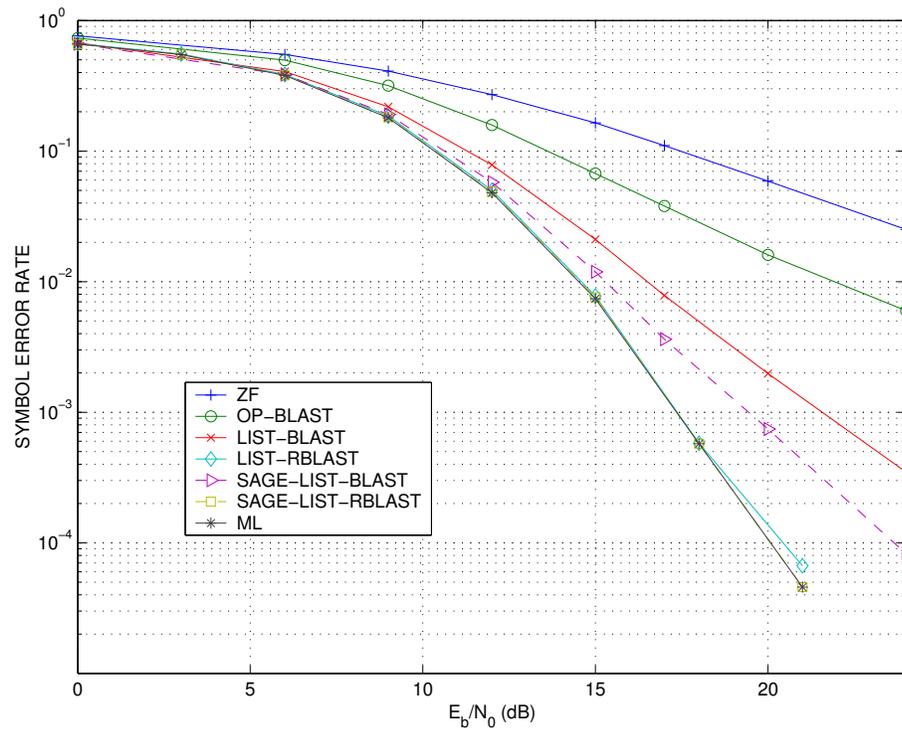
Fig. 16. Symbol Error Rate of different detectors for a 4×4 MIMO system with uncoded 16-QAM modulation.
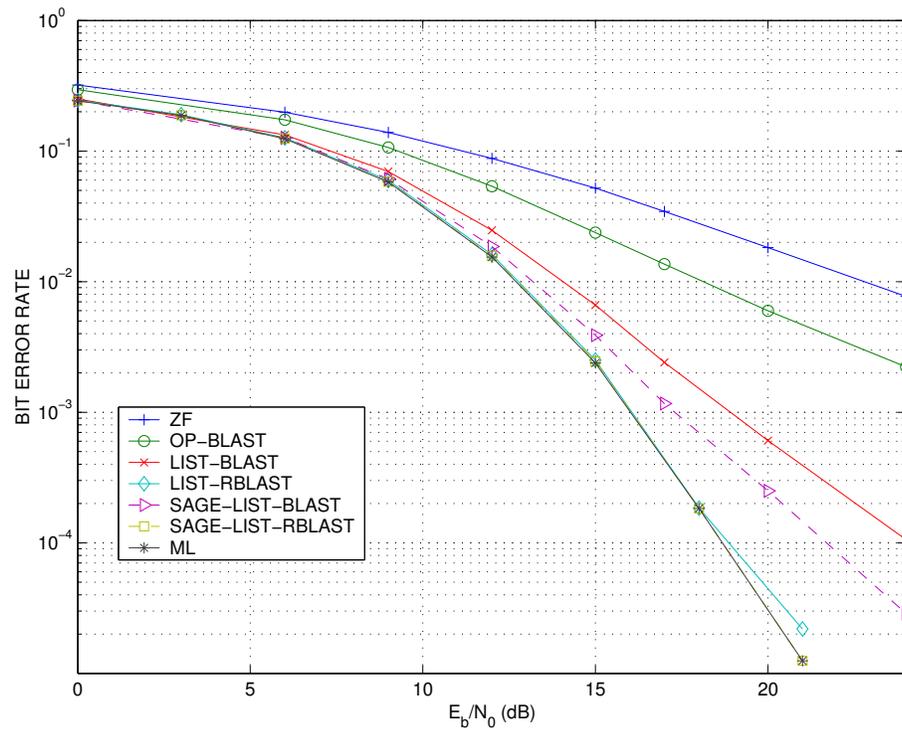
Fig. 17. Bit Error Rate of different detectors for a 4×4 MIMO system with uncoded 16-QAM modulation.

the SER and the BER of the ML detector implemented by sphere detection, the ZF detector, the zero-forcing BLAST detector with optimal detection order (the layer with the highest SNR is detected first), the List-BLAST detector, the List-Ranked-BLAST detector, the List-Shifted-BLAST detector and the SAGE aided detectors for a 4×4 MIMO system with uncoded 8PSK modulation scheme. Fig. 16 and Fig. 17 show the SER and the BER of the different detectors for the same MIMO system with uncoded 16QAM. It is not hard to show that the ML detection achieves a spatial diversity order of four in this case. There is no spatial diversity order for the ZF detector. For both the QAM and the PSK modulation, the BLAST detector with optimal detection order, denoted as "OP-BLAST" in all the figures, achieves a spatial diversity order greater than one, but is outperformed by the List-BLAST detector denoted as "LIST-BLAST" and the List-Ranked-BLAST detector denoted as "LIST-RBLAST". The List-BLAST algorithm provides a way to avoid detection error in the first detection layer (the N_t^{th} layer), which is a Rayleigh fading channel. Therefore, in the List-BLAST case, error is dominated by the $(N_t - 1)^{th}$ layer, which has spatial diversity of order two. We can achieve a further 2.5 dB gain by performing SAGE iterations as shown by the curve denoted as "SAGE-LIST-BLAST". In the List-Ranked-BLAST case, diversity order is further improved by ordering the nulling and cancelling as described in the last section. Actually, the List-Ranked-BLAST performs almost the same as the ML detection in this 4×4 MIMO case. Therefore, SAGE iterations cannot further improve its performance, and are not required in this case. Since the List-Ranked-BLAST detection scheme is only of detection complexity order of $O(CN_t^2)$, it is an excellent candidate for detection of 4×4 uncoded MIMO systems. Note that the performance of the List-Shifted-BLAST detector in the 4×4 MIMO systems is also almost the same as the ML detector, and is not plotted in these figures. But it is more computationally complex than the List-Ranked-BLAST
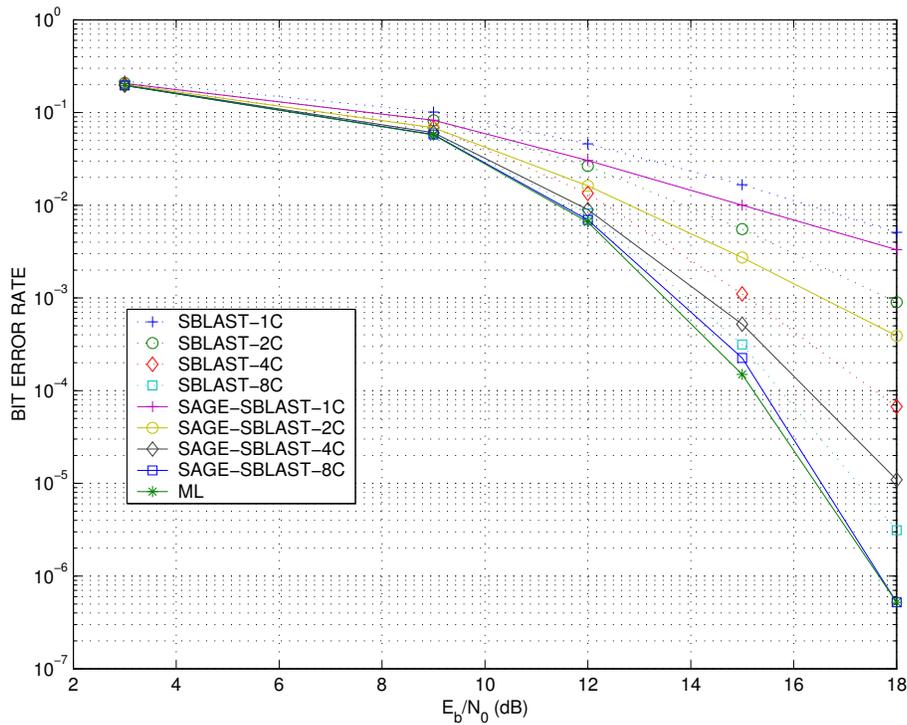
Fig. 18. Bit Error Rate of different detectors for a 8×8 MIMO system with uncoded 16-QAM modulation.

detector.

Fig. 18 shows the BER of the List-Shifted-BLAST algorithm for different values of $K = 1, 2, 4, 8$ for a 8×8 MIMO system with uncoded 16-QAM modulation. As K increases, the diversity order achievable by the List-Shifted-BLAST detector increases. When $K = 8$, the List-Shifted-BLAST detector can achieve a performance close to that of the ML detector. For the cases of $K < N_t = 8$, the SAGE-aided LIST-Shifted-BLAST detector can achieve an additional gain up to 1 dB over the List-Shifted-BLAST detector with the same value of K. In constrast to the 4×4 MIMO case, the performance of the List-Ranked-BLAST detector is not close to the ML performance in this case, and it is not shown in the figure.

Finally, we evaluated the BER performance of turbo-coded MIMO systems. The

Fig. 19. Bit Error Rate of turbo-coded 4 × 4 MIMO systems with SAGE-aided List-Shifted-BLAST decoding and the simple soft-output BLAST decoding.

rate $R = 1/2$ 16-state parallel turbo code with polynomial (23, 31) is used. Note that joint demodulation and detection is not performed here: the soft-output generated by the MIMO demodulator is passed to the turbo decoder, which has 8 iterations; the a posteriori probability of the coded bits after turbo-decoding is not passed back to the demodulator for simplicity. The curves deno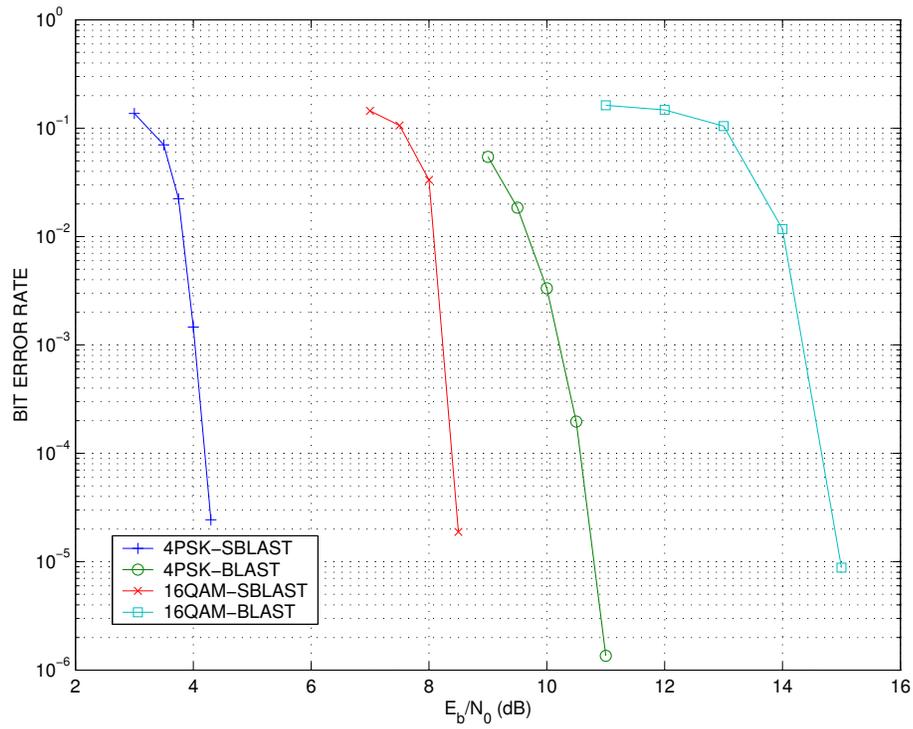ted as "4PSK-SBLAST" and "16QAM-SBLAST" in Fig. 19 show the performance of the SAGE-aided List-Shifted-BLAST detector for a turbo-coded 16-QAM and a 4-PSK 4×4 MIMO systems, respectively. Both the initial points generated using the List-Shifted-BLAST algorithm and the converged points generated using the SAGE algorithm are included in the candidate set $\mathbb{L}$ whose size is 128 and 32 in the case of 16-QAM and 4-PSK, respectively. The minimum required $\frac{E_b}{N_0}$ to achieve the capacity of the MIMO systems are 1.6 (dB) and 3.7 (dB), respectively [12]. Therefore, the performance of the SAGE-aided List-Shifted-BLAST detection scheme in these two cases is only 3 and 5 dB away from the capacity, respectively. As a comparison, we also simulated a simple soft-output BLAST detection scheme. In this scheme, traditional BLAST with optimal detection order is used. In each layer, "hard" cancellation of interference from the previously detected layers is performed; 8 iterations are used in the turbo-decoder; soft-information of the coded bits after turbo decoding is not passed back to the demodulator. The performance of this simple soft-information BLAST detection scheme for the two different modulation schemes is denoted as "4PSK-BLAST" and "16QAM-BLAST", respectively. For both cases, the simple BLAST scheme is more than 6 dB worse than the corresponding SAGE-aided List-Shifted-BLAST detection scheme.

D.   Conclusion

We have proposed a novel low complexity MIMO detector, called SAGE-aided List-BLAST detector. We use List-BLAST type algorithm to generate multiple initial points for the SAGE algorithm; after performing the SAGE iterations, the detector finally selects the minimum distance point to the received vector from those converged points generated by SAGE. We show that the List-BLAST type algorithms (including the List-Ranked-BLAST and the List-Shifted-BLAST) alone can achieve performance close to the ML detection. The SAGE algorithm can be used in combination with the List-BLAST algorithm to further improve system performance. The proposed algorithms have a complexity advantage over the sphere decoding algorithm at low SNR.

CHAPTER VII

CHANNEL ESTIMATION FOR OFDM SYSTEMS USING EM ALGORITHMS*

A. Introduction

In the previous chapters, we have been dealing with transmission and detection assuming frequency flat fading channels. In this chapter, we will study channel estimation for a frequency selective fading channel using orthogonal frequency division multiplexing (OFDM) techniques.

OFDM, which can transform a frequency-selective fading channel into many parallel flat fading sub-channels, is an efficient technique to combat multipath delay-spread in high-rate wireless systems. OFDM has already been accepted for the new wireless local area network (WLAN) standards (IEEE 802.11a), the European Telecommunications Standards Institute (ETSI)'s High Performance Local Area Network Type 2 (HIPERLAN/2) and Japan's Mobile Multimedia Access Communication (MMAC) systems [62]. With the rapid growth of the Internet, providing advanced Internet service over wide-area cellular networks is of great commercial interest [63]. Although in the WLAN standards, data rates up to 54Mbps might be achieved with conventional OFDM with single antenna, transmission at a peak rate of several Mbps is extremely challenging in a wide-area network because of significant path-loss, large delay-spread and fading [64]. To meet these challenges, OFDM schemes combined with transmitter and receiver diversity were proposed, among which space-time coded OFDM (ST-OFDM) is one of the most efficient transmitter diversity schemes [7][9][65] [66] [67].

---

*© 2003 IEEE. Reprinted, with permission, from Y. Xie and C. N. Georghiades, "Two EM-based channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106-115, Jan. 2003.

In ST-OFDM, channel state information between each transmit and receive antenna pair is required for coherent decoding. However, for each OFDM tone, since the received signals are a superposition of signals transmitted from different antennas, the simple channel estimation techniques used in single transmit antenna systems cannot be used. This chapter discusses two Expectation-Maximization (EM) type channel estimation algorithms in such scenarios. The EM-type algorithms essentially convert a multiple-input channel estimation problem into a number of single-input channel estimation problems, a much more palatable problem.

Although the discussion in the chapter is based on space-time trellis coded OFDM systems, the algorithm can be directly used for other OFDM systems with multiple transmit antennas.

The rest of the chapter is organized as follows. Section B describes transmitter diversity using space-time coding for OFDM systems and introduces the fading channel model that is used. Section C addresses least-square (LS) channel estimation for ST-OFDM systems and introduces a classical EM algorithm and a Space Alternating Generalized Expectation-Maximization (SAGE) algorithm, both of which are based on a single received OFDM symbol block. Section D compares the two proposed algorithms in terms of convergence rate and unifies them in a "message-passing" type iterative structure. Complexity of implementation and combining of the EM-type algorithms with the significant-tap-catching (STC) estimator recently proposed by Y. Li [67] are also discussed. Section E provides simulation results on the convergence of the EM-based algorithms and overall system performance. Finally, Section F concludes.
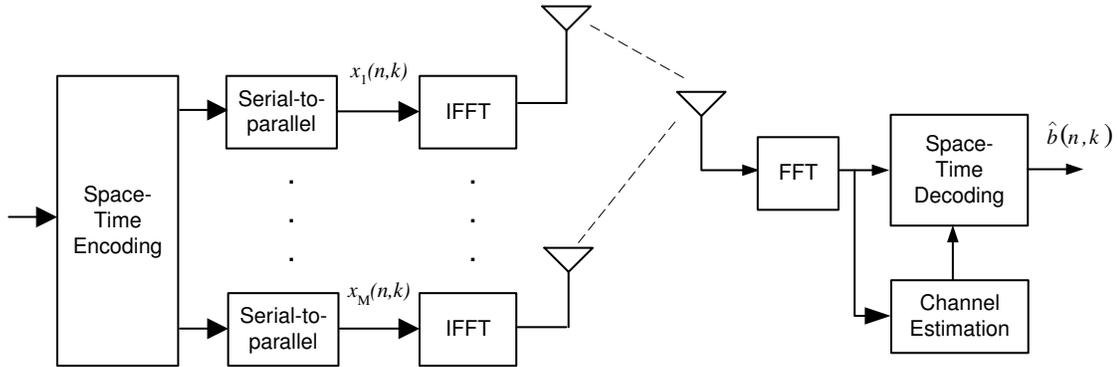
Fig. 20. ST-OFDM system.

B.   ST-OFDM Systems and Channel Model

1.   ST-OFDM Systems

An OFDM system with M transmit antennas and one receive antenna is shown in Fig. 20. At time n, a data block $\{b(n,k)\}$, $k = 0, 1, 2, \ldots, N-1$, where N is the number of sub-channels (tones), is coded into M different symbol blocks, $\{x_i(n,k)\}$, $k = 0, 1, 2, \ldots, N-1$, $i = 1, 2, \ldots, M$. Each block is transmitted through different antennas over the same bandwidth using N OFDM tones. In other words, between each transmit antenna and the receiver there is a communication link established by OFDM. Hence, the received signal after demodulation (performing a Discrete Fourier Transform (DFT)), is the superposition of M distorted transmitted signals, which can be expressed in vector form as

$$\mathbf{r}_n = \sum_{i=1}^{M} \mathbf{X}_{i,n} \mathbf{H}_{i,n} + \mathbf{W}_n. \tag{7.1}$$

Here, $\mathbf{X}_{i,n}$ is an $N \times N$ diagonal matrix with $\mathbf{X}_{i,n}[k,k] = x_i(n,k)$ representing the symbol transmitted through the ith antenna over the kth tone at time n. In the following discussion, Phase Shift Keying (PSK) modulation with unit symbol energy is assumed, such that $|x_i(n,k)| = 1$ for any i, k and n.

$\mathbf{H}_{i,n}$ is an $N \times 1$ vector with $\mathbf{H}_{i,n}[k]$ denoting the channel frequency response at the kth tone between the ith antenna and the receiver at time n. Finally, $\mathbf{W}_n$ is an $N \times 1$ zero-mean, i.i.d. Gaussian vector that models additive noise in the N sub-channels (tones) at time n. We have

$$E[\mathbf{W}_n^H \mathbf{W}_n] = \delta_n^2 \mathbf{I}_N, \tag{7.2}$$

where $\mathbf{I}_N$ is an $N \times N$ identity matrix and δ_n^2 is the variance of the additive noise at time n.

At the receiver, a Viterbi algorithm with the following metric is used to decode the space-time trellis code:

$$\left\| \mathbf{r}_n[k] - \sum_{i=1}^{M} \mathbf{X}_{i,n}[k,k]\mathbf{H}_{i,n}[k] \right\|^2, \quad \text{for } k = 0, 1, 2 \dots N - 1, \tag{7.3}$$

where $\| \cdot \|$ denotes Euclidean norm. Obviously, channel parameter estimation is essential for decoding space-time codes using the above metric.

2. The Channel Model

The impulse response of the fading channel between the ith transmit antenna and the receiver, $h_i(t)$, can be modelled as

$$h_i(t) = \sum_j \alpha_{i,j} \ \delta(t - \tau_{i,j}T_s), \tag{7.4}$$

where $\tau_{i,j}$ is the channel delay associated with the ith transmitter and the jth path, $\alpha_{i,j}$ are zero-mean, complex Gaussian random variables with a power-delay profile $\theta(\tau_{i,j})$ and T_s is the sampling interval of the OFDM system. Let T_c denote the length of the cyclic prefix which satisfies $T_c = N_c T_s$, where N_c is the number of samples of the cyclic prefix. For OFDM systems, if the cyclic prefix is sufficiently long ($0 \leq \tau_{i,j}T_s \leq T_c$

for any i and j) and perfect sample timing is assumed, the discrete channel impulse response at time n can be expressed as an $L_h \times 1$ vector $\mathbf{h}_{i,n}$ with tolerable leakage [68], where L_h satisfies $L_h \leq N_c + 1$. Therefore, the frequency response vector $\mathbf{H}_{i,n}$ can be expressed as

$$\mathbf{H}_{i,n} = \mathbf{F}\mathbf{h}_{i,n}, \tag{7.5}$$

where $\mathbf{F}$ is an $N \times L_h$ matrix with $\mathbf{F}[k,l] = \frac{1}{\sqrt{N}}e^{-i2\pi kl/N}$, $0 \leq k \leq N-1$, $0 \leq l \leq L_h-1$. Obviously $\mathbf{F}$ is constructed by the first L_h columns of the $N \times N$ square DFT matrix.

C. EM-type Channel Estimation Algorithms

1. Method of Least Squares

For OFDM with transmitter diversity, channel estimation is challenging since the received signal at each tone is a function of multiple channel distortions.

Modelling the channel impulse response at time n as deterministic but unknown, a temporal estimation of the channel impulse response vector is obtained by directly minimizing the following cost function [67]

$$\hat{\mathbf{h}}_n = \arg \min_{\mathbf{h}_n} \; C(\mathbf{h}_n) = \| \mathbf{r}_n - \mathbf{G}_n \mathbf{h}_n \|^2, \tag{7.6}$$

where

$$\mathbf{G}_n \triangleq [\mathbf{X}_{1,n}\mathbf{F} \; \mathbf{X}_{2,n}\mathbf{F} \; \cdots \; \mathbf{X}_{M,n}\mathbf{F}], \tag{7.7}$$

$$\mathbf{h}_n \triangleq [\mathbf{h}_{1,n}^H \; \mathbf{h}_{2,n}^H \; \cdots \; \mathbf{h}_{M,n}^H]^H, \tag{7.8}$$

$$\hat{\mathbf{h}}_n \triangleq [\hat{\mathbf{h}}_{1,n}^H \; \hat{\mathbf{h}}_{2,n}^H \; \cdots \; \hat{\mathbf{h}}_{M,n}^H]^H. \tag{7.9}$$

Ignoring the leakage due to non-uniform channel tap spacing and assuming (7.5) is the correct channel model, the least squares (LS) solution of (7.6) is also the

maximum-likelihood (ML) channel estimate (assuming known transmitted symbols). If $\mathbf{G}_n$ is of full column rank, then the ML solution, $\hat{\mathbf{h}}_n^{<ML>}$, can be uniquely determined by

$$\hat{\mathbf{h}}_n^{<ML>} = (\mathbf{G}_n^H \mathbf{G}_n)^{-1} \mathbf{G}_n^H \mathbf{r}_n = \mathbf{Q}_n^{-1} \mathbf{P}_n, \tag{7.10}$$

where

$$\mathbf{Q}_n \triangleq \begin{pmatrix} \mathbf{I}_{L_h} & \mathbf{F}^H \mathbf{X}_{1,n}^H \mathbf{X}_{2,n} \mathbf{F} & \cdots & \mathbf{F}^H \mathbf{X}_{1,n}^H \mathbf{X}_{M,n} \mathbf{F} \\ \mathbf{F}^H \mathbf{X}_{2,n}^H \mathbf{X}_{1,n} \mathbf{F} & \mathbf{I}_{L_h} & \cdots & \mathbf{F}^H \mathbf{X}_{2,n}^H \mathbf{X}_{M,n} \mathbf{F} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{F}^H \mathbf{X}_{M,n}^H \mathbf{X}_{1,n} \mathbf{F} & \mathbf{F}^H \mathbf{X}_{M,n}^H \mathbf{X}_{2,n} \mathbf{F} & \cdots & \mathbf{I}_{L_h} \end{pmatrix} \tag{7.11}$$

$$\mathbf{P}_n \triangleq [\mathbf{r}_n^H \mathbf{X}_{1,n} \mathbf{F} \quad \mathbf{r}_n^H \mathbf{X}_{2,n} \mathbf{F} \quad \cdots \quad \mathbf{r}_n^H \mathbf{X}_{M,n} \mathbf{F}]^H. \tag{7.12}$$

Here, $\mathbf{I}_{L_h}$ denote the $L_h \times L_h$ identity matrix. The special case of $M = 2$ of the above solution is the same as in [67].

Since $\mathbf{G}_n$ is an $N \times ML_h$ matrix, a necessary condition for the channels to be uniquely identifiable is

$$ML_h \leq N. \tag{7.13}$$

(7.13) suggests that the channels cannot be uniquely identified from one OFDM symbol if M times the number of channel delay-taps to be estimated is greater than the number of tones.

2. The EM-Based Algorithm

A drawback of directly solving (7.10) is that the calculation of the inverse of the $ML_h \times ML_h$ square matrix, $\mathbf{Q}_n$, is required. This inverse matrix requires significant

computation for large values of L_h and M. To overcome this drawback, the authors in [67] proposed to use only the L $(L < L_h)$ most significant channel taps to model the channel of length L_h, thus reducing the size of $\mathbf{Q}_n$ to $ML \times ML$, a computationally tolerable level for small L. This method is referred to as significant-tap-catching (STC) in [67]. Although in most cases this simplified method works well, it may introduce an irreducible error floor for channels with a power profile that cannot be represented adequately by the L taps used to represent the channel.

Another solution is to design the training blocks $\mathbf{X}_{i,n}$ for all $1 \leq i \leq M$ to make $\mathbf{Q}_n$ diagonal, so that matrix inversion is trivial [69]. However, this method can only provide channel estimates at the pilot blocks and therefore cannot work in a decision-directed feedback mode needed to track channel variations.

Instead of minimizing (7.6) directly, EM-type algorithms provide an iterative and more easily implementable solution. Here we apply both the EM and the SAGE algorithms to the problem at hand. Since the EM-type algorithms have been thoroughly studied and applied to a number of problems in communication over the years, we will not describe them in detail in this chapter. The reader is urged to read [54] for a general exposition to the EM algorithm and [55][56] for applications to the estimation problem related to the work herein.

In the EM algorithm, we view the observed data $\mathbf{r}_n$ as the "incomplete" data and define the "complete" data $\mathbf{Y}_{i,n}$ as

$$\mathbf{Y}_{i,n} = \mathbf{X}_{i,n}\mathbf{F}\mathbf{h}_{i,n} + \mathbf{W}_{i,n}, \quad 1 \leq i \leq M, \tag{7.14}$$

where $\sum_{i=1}^{M} \mathbf{W}_{i,n} = \mathbf{W}_n$; thus, $\sum_{i=1}^{M} \mathbf{Y}_{i,n} = \mathbf{r}_n$. $\mathbf{Y}_{i,n}$ is the component of the received signal transmitted by the ith antenna through the channel with impulse response $\mathbf{h}_{i,n}$. It is easy to show that the EM algorithm for the above particular choice of complete data takes the following form:

- E-Step: For $i = 1, 2, \ldots, M$, compute

$$\hat{\mathbf{Z}}_{i,n}^{(k)} = \mathbf{X}_{i,n}\mathbf{F}\hat{\mathbf{h}}_{i,n}^{(k)}, \tag{7.15}$$

$$\hat{\mathbf{Y}}_{i,n}^{(k)} = \hat{\mathbf{Z}}_{i,n}^{(k)} + \beta_i \left[\mathbf{r}_n - \sum_{j=1}^{M} \hat{\mathbf{Z}}_{j,n}^{(k)} \right]. \tag{7.16}$$

- M-Step: For $i = 1, 2, \ldots, M$, compute

$$\hat{\mathbf{h}}_{i,n}^{(k+1)} = \arg\min_{\mathbf{h}_{i,n}} \ \left\{ \| \hat{\mathbf{Y}}_{i,n}^{(k)} - \mathbf{X}_{i,n}\mathbf{F}\mathbf{h}_{i,n} \|^2 \right\}. \tag{7.17}$$

The superscript (k) denotes the kth iteration and the β_i are chosen such that $\sum_{i=1}^{M} \beta_i = 1$. Solving (7.17), we obtain:

$$\hat{\mathbf{h}}_{i,n}^{(k+1)} = \mathbf{F}^H\mathbf{X}_{i,n}^{-1}\hat{\mathbf{Y}}_{i,n}^{(k)}. \tag{7.18}$$

Although not necessary for implementation, equations (7.15), (7.16) and (7.18) can be combined to yield the following recursion:

$$\hat{\mathbf{h}}_{i,n}^{(k+1)} = \hat{\mathbf{h}}_{i,n}^{(k)} + \beta_i \left[\mathbf{F}^H\mathbf{X}_{i,n}^{-1}\mathbf{r}_n - \sum_{j=1}^{M} \mathbf{F}^H\mathbf{X}_{i,n}^{-1}\mathbf{X}_{j,n}\mathbf{F}\hat{\mathbf{h}}_{j,n}^{(k)} \right]. \tag{7.19}$$

Note that $\mathbf{X}_{i,n}$ is a diagonal matrix and, thus, calculation of its inverse is trivial. Also note that (7.18) is just the well-known LS channel estimation scheme for the conventional (i.e., single transmit antenna) OFDM system [68]. The motive of the EM algorithm is clear: At the E-step, it estimates the corresponding component in the received signal for each of the OFDM links. At the M-step, as in the conventional OFDM scheme, it divides the corresponding component by the reference symbols (either known from training, or previously decoded symbols) in the frequency domain and then performs an IFFT to obtain an updated estimate of the channel impulse response.

The convergence rate of the EM algorithm is inversely related to the Fisher

information of its complete-data space [56]. In the above algorithm, the noise variance is distributed over $\mathbf{Y}_{i,n}$ for all i; therefore, the Fisher information of $\mathbf{Y}_{i,n}$ for $\mathbf{h}_{i,n}$ is relatively large for a certain i. To improve the convergence rate, we can choose the hidden data space as $\mathbf{Y}_{i,n}$ for $i = 1, 2, \cdots, M$ alternately in each iteration and, thus, associate all the noise variance with it. The SAGE algorithm for this specific problem is then:

- Initialization: For $1 \le i \le M$,

$$\hat{\mathbf{Z}}_{i,n}^{(0)} = \mathbf{X}_{i,n}\mathbf{F}\hat{\mathbf{h}}_{i,n}^{(0)}. \tag{7.20}$$

- At the kth iteration $(k = 0, 1, 2, ...)$:

  For $i = 1 + [k \bmod M]$, compute

$$\hat{\mathbf{Y}}_{i,n}^{(k)} = \hat{\mathbf{Z}}_{i,n}^{(k)} + \left[\mathbf{r}_n - \sum_{j=1}^{M}\hat{\mathbf{Z}}_{j,n}^{(k)}\right], \tag{7.21}$$

$$\hat{\mathbf{h}}_{i,n}^{(k+1)} = \mathbf{F}^{H}\mathbf{X}_{i,n}^{-1}\hat{\mathbf{Y}}_{i,n}^{(k)}, \tag{7.22}$$

$$\hat{\mathbf{Z}}_{i,n}^{(k+1)} = \mathbf{X}_{i,n}\mathbf{F}\hat{\mathbf{h}}_{i,n}^{(k+1)}. \tag{7.23}$$

  For $1 \le j \le M$ and $j \ne i$,

$$\hat{\mathbf{Z}}_{j,n}^{(k+1)} = \hat{\mathbf{Z}}_{j,n}^{(k)}. \tag{7.24}$$

A proper selection of the initial value of $\hat{\mathbf{h}}_{i,n}$ is very important for the convergence speed of both algorithms. Intuitively, assuming all the signals transmitted from other than the ith antenna to be zero (though not true in practice), we can obtain an initial estimate of the channel for the EM-type iteration as follows:

$$\hat{\mathbf{h}}_{i,n}^{(0)} = \mathbf{F}^{H}\mathbf{X}_{i,n}^{-1}\mathbf{r}_n, \quad i = 1, 2, \cdots, M. \tag{7.25}$$

The introduced EM-type algorithms only provide a channel estimate at time n.

With this temporal estimate, a complete channel estimation scheme can be developed according to the time-selectivity of the fading. For high-speed wireless data packet applications, if a data packet is short compared to the channel coherence time, channel fading can be assumed to be the same for the whole packet. For each transmitter, one or two pilot symbol blocks can be sent at the beginning of each packet; the temporal estimates (or the average of the two temporal estimates in case of two pilot symbol blocks) can then be used for the whole packet [62]. In this case, each OFDM link can also be estimated alternately by transmitting at each time a training symbol from a specified antenna while transmitting no signals from others. The estimation is greatly simplified at the cost of an M-fold increase in training time. If channel parameters cannot be assumed constant over the whole packet, but almost the same for several continuous data symbols, the alternate estimation method cannot be used. In this case, the EM-type algorithms provide flexibility for both estimation and tracking. The decoded bits can be encoded again and used to estimate the current channel parameters, which will be used in the decoding of the next block. Actually, except for the first training block, we can always use the last channel estimate as the initial value for the current estimation, which will significantly reduce the number of iterations (could be less than 3 iterations as shown in the simulation) until convergence. In this case, the decision-directed EM-type algorithms can be treated as adaptive channel tracking algorithms. Of course, STC can also be used in the decision directed mode, but not as efficiently as the EM-type algorithms since it operates independently for each block and cannot use the previous channel estimates. For applications in which large delay is tolerable, a more accurate estimation of the channels at each time could be obtained by passing the temporal estimates through a Wiener filter for each communication link [62].

D. Remarks

1. Convergence

It is already proven in [56] that the SAGE algorithm converges faster than the classical EM algorithm in estimation of superimposed signals in Gaussian noise. The difference here is that the parameter associated with each superimposed signal is a vector instead of a variable as in [56]. Here we only give a single analysis on the best convergence rate of the two algorithms.

If we can write the equation for iterations as follows

$$\hat{\mathbf{h}}_n^{(k+1)} - \hat{\mathbf{h}}_n^{<ML>} = \mathbf{A}_n(\hat{\mathbf{h}}_n^{(k)} - \hat{\mathbf{h}}_n^{<ML>}), \tag{7.26}$$

for some matrix $\mathbf{A}_n$, the convergence factor can be defined as the matrix spectral radius $\rho(\mathbf{A}_n)$, the largest magnitude of eigenvalue of $\mathbf{A}_n$. Obviously, a smaller convergence factor indicates faster convergence rate.

According to Appendix C, we have for EM

$$\mathbf{A}_n = \mathbf{I}_{ML_h} - \frac{1}{M}\mathbf{Q}_n. \tag{7.27}$$

where $\mathbf{I}_{ML_h}$ denotes the $ML_h \times ML_h$ identity matrix.

In the EM algorithm, channel impulse responses associated with each OFDM link are simultaneously updated during each iteration, while in the SAGE algorithm, only one OFDM link is updated in each iteration. Since all the OFDM links are updated every M iterations with the total complexity similar to that of one iteration in the EM algorithm, for fair comparison, we will count M iterations of the SAGE algorithm as one iteration in later discussions.

For SAGE (see Appendix C):

$$\mathbf{A}_n = \mathbf{I}_{ML_h} - LT(\mathbf{Q}_n)^{-1}\mathbf{Q}_n, \tag{7.28}$$

where $LT(\mathbf{Q}_n)$ is the lower triangular part of $\mathbf{Q}_n$ (including the diagonal entries). According to [56], both the EM and SAGE algorithms converge and we have

$$\rho_{sage}(\mathbf{A}_n) \le \rho_{em}(\mathbf{A}_n). \tag{7.29}$$

$\mathbf{Q}_n$ is Hermitian and non-negative definite, its eigenvalues λ_i $(i = 1, 2, \dots, ML_h)$ are all real and non-negative and satisfy $\sum_{i=1}^{ML_h} \lambda_i = tr(\mathbf{Q}_n) = ML_h$, where $tr(\cdot)$ is the trace operation (assuming PSK is used). Therefore, we have

$$\rho_{em}(\mathbf{A}_n) \ge 1 - 1/M, \tag{7.30}$$

where equality holds iff $\mathbf{Q}_n = \mathbf{I}_{ML_h}$, corresponding to the case when the training block is designed as in [69]. So the best convergence factor achievable is $1 - 1/M$ for the EM algorithm. Note that in this best case, the initial guess of $\hat{\mathbf{h}}_{i,n}^{(0)}$ as in (7.25) happens to match the maximum likelihood solution, no further iterations are actually required. The best convergence factor of SAGE can be zero when $\mathbf{Q}_n = \mathbf{I}_{ML_h}$, which means the SAGE algorithm converges in just one iteration regardless of the initial value.

The convergence factor is a random variable because it is a function of the reference signal $\mathbf{X}_{i,n}$, which are random data blocks in the decision-feedback mode. The distribution of the convergence factor will be affected by the size of $\mathbf{G}_n$, i.e. M, L_h and N [70]. Especially for the EM algorithm, according to the conjecture in [71](page 166), the smallest eigenvalue of $\mathbf{Q}_n$ converges to $(1 - \sqrt{ML_h/N})^2$ if $N \to \infty$, and

ML_h/N is a constant. Therefore, the asymptotic convergence factor of EM, $\tilde{\rho}_{em}$, is

$$\tilde{\rho}_{em}(\mathbf{A}_n) = 1 - \frac{1}{M}(1 - \sqrt{ML_h/N})^2. \qquad (7.31)$$

Since explicit formulas for the distribution function of the spectral radius of a random matrix with finite size is not known except for some special matrices [70], we will compute the distribution of the convergence factor for EM-type algorithms through simulation in the next section.

2. A "Message Passing" Interpretation

Similar to the iterative "Message Passing" decoding structure in low density parity check codes (LDPC) [72], we can view the EM-type estimation algorithms as a special kind of "Message Passing" procedure between variable nodes associated with channel parameters to be estimated and a check node associated with the observed OFDM block, as shown in Fig. 21.

At the kth iteration, each variable node V_i passes the message $S_{V_i \to C} = \hat{\mathbf{Z}}_{i,n}^{(k)}$, which is an estimate of $\hat{\mathbf{Y}}_{i,n}^{(k)}$, to the check node C. It then combines all the incoming messages, computes and passes back an updated message, $S_{C \to V_i} = \hat{\mathbf{h}}_{i,n}^{(k+1)}$ to V_i. Note that there is a hidden constraint for each V_i, i.e., $\mathbf{h}_{i,n}$ only has L_h taps.

Based on the constraint that $\sum_{i=1}^{M} \hat{\mathbf{Y}}_{i,n}^{(k)} = \mathbf{r}_n$, the EM and SAGE algorithms use different message updating schemes at the check node.

For the EM algorithm, to compensate for a non-zero difference $\mathbf{r}_n - \sum_{i=1}^{M} \hat{\mathbf{Z}}_{i,n}^{(k)}$ at the check node, $\hat{\mathbf{Y}}_{i,n}^{(k)}$ is balanced by adding a fraction (associated with β_i) of the difference to the original value of $\hat{\mathbf{Z}}_{i,n}^{(k)}$ so that the constraint is satisfied. An outgoing message $S_{C \to V_i}$ for $(i = 1, 2, \cdots, M)$ is then computed based on $\hat{\mathbf{Y}}_{i,n}^{(k)}$.

In the SAGE algorithm, the constraint is forced to be satisfied at the check node by setting $\hat{\mathbf{Y}}_{i,n}^{(k)}$ to be $\mathbf{r}_n - \sum_{j=1(j \neq i)}^{M} \hat{\mathbf{Z}}_{j,n}^{(k)}$ with all other $\hat{\mathbf{Y}}_{j,n}^{(k)} = \hat{\mathbf{Z}}_{j,n}^{(k)}$ $(1 \leq j \leq M, j \neq i)$
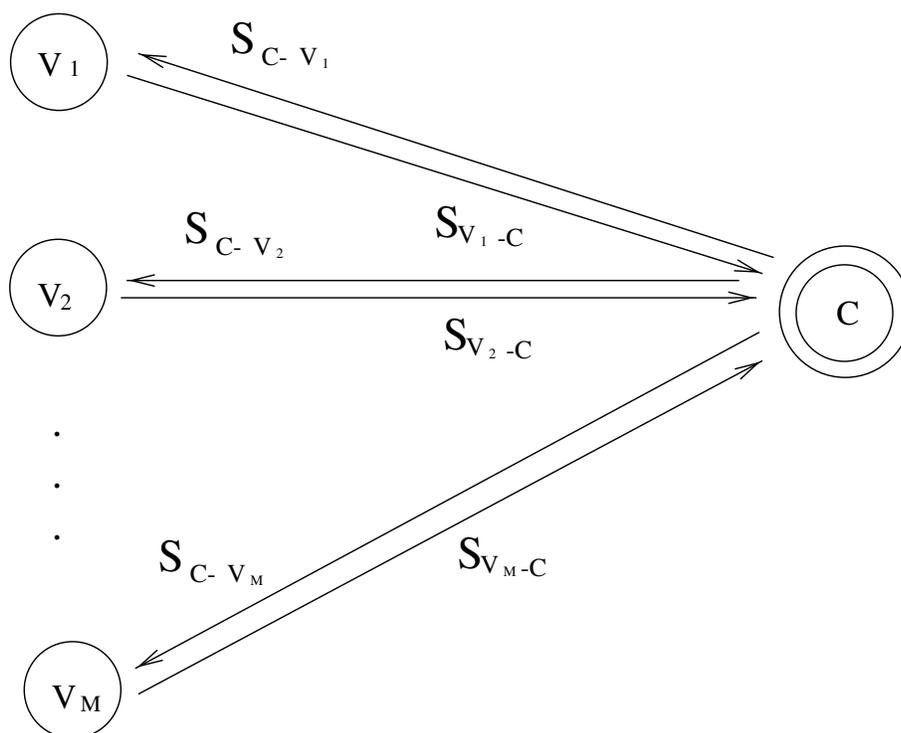
Fig. 21. A Message Passing explanation of the EM-type algorithms.

unchanged. Otherwise, if all $\hat{\mathbf{Y}}_{j,n}^{(k)}$ are updated as $\hat{\mathbf{Y}}_{i,n}^{(k)}$ simultaneously, the constraint will not be satisfied. We can also easily show that (see Appendix C) the corresponding convergence matrix $\mathbf{A}_n = \mathbf{I}_{ML_h} - \mathbf{Q}_n$, whose spectral radius is not guaranteed to be less than 1. Therefore, parallel updating in SAGE will not guarantee convergence.

3.   Implementation Complexity for STC-EM and STC-SAGE

Compared with the STC method with direct matrix inversion, EM-type algorithms have a number of implementation advantages. Note that for all M OFDM links, the estimation structure is exactly the same. No complex computation union is involved, except for the FFT/IFFT operations. Since an FFT is a necessary unit of the demodulator and, hence, already available, the estimator can use the same unit (an IFFT can be implemented by FFT as $\text{IFFT}(x) = \text{conj}(\text{FFT}(\text{conj}(x)))$) to perform the FFT and IFFT for the EM-based algorithm.

In the STC algorithm, $\mathbf{Q}_n$ and $\mathbf{P}_n$ can be efficiently computed by $(M + M(M - 1)/2)$ N-point-FFTs/IFFTs as shown in [67]. The inverse of $\mathbf{Q}_n$ is of computational complexity $O((ML)^3)$. Equation (7.10) can be also regarded as the product of the pseudo-inverse of $\mathbf{G}_n$ and the vector $\mathbf{r}_n$, so it can be computed iteratively using Greville's method without directly solving the pseudo-inverse (pp. 223 of [73]). However, the computational complexity is not reduced as shown in Appendix D. In the EM-type algorithms, the total number of FFTs/IFFTs required is $2MN_{it}$ and some extra multiplications (N_{it} is the number of iterations). Note that FFT/IFFTs (radix-2 type) requires $0.5N\log_2(N)$ of multiplications [74]. Therefore, the SAGE algorithm is preferred than the EM algorithm, and both algorithms are much more efficient in the decision feedback channel tracking mode, where N_{it} is small. (For the channel training block, $\mathbf{X}_{i,n}$ can be designed so that $\mathbf{Q}_n$ is diagonal.) As will be shown in our simulations, the number of iterations for convergence can be as low as less than

3 for the SAGE algorithm, and 10 for the EM in the case of two-transmit antennas and 17 channel taps. An example comparing the complexity of the EM-type algorithms with STC is given in Appendix D. It is shown that the EM-type algorithms, and especially the SAGE, have much lower computational complexity than the STC when M and L are large. Actually, the EM-type algorithms can also be used to solve the matrix inversion encountered in STC, when direct computation of the matrix inversion is too complex. In this case, from the "Message Passing" point of view, the hidden constraints at the variable nodes are that non-selected taps in STC are zero. In other words, the equations in EM and SAGE remain unchanged except that the DFT matrix $\mathbf{F}$ becomes an $N \times L$ matrix with $\mathbf{F}[k, l] = \frac{1}{\sqrt{N}} e^{-i2\pi kl/N}$, $0 \leq k \leq N-1$, $l \in \Omega_L$, where Ω_L denotes the indices of the selected significant taps. We denote the corresponding algorithms as STC-EM or STC-SAGE.

E. Simulation Results

Simulations were conducted to test the convergence of the EM-type algorithms and the ST-OFDM system performance using such estimators. The simulation parameters were set as follows:

- Similar to [67], the entire channel bandwidth was 800 KHz and was divided into 128 sub-carriers (or tones). The symbol duration was 160-μs. An additional 40-μs guard interval was used to provide protection from intersymbol interference (ISI) due to channel delay-spread. The Doppler shift was chosen to be 40Hz.

- Two transmit antennas and one receive antenna were employed. The 2-space-time codes of 4-PSK, 2b/s/Hz with 16 states (their trellis structure is shown in Fig. 5 of [7]) were used in the simulation.

- The simulations were carried out for two different channel delay and power

profiles, as shown in Fig. 22. The maximum channel delay of 20-μs was assumed to be known in all the simulations.

Fig. 23 shows the mean-square-error (MSE) performance of the EM-type algorithms as a function of the number of iterations. It also includes comparisons with the MSE of the ML estimator and a 9-tap STC estimator. All the results were evaluated based on 5000 OFDM simulated blocks sent from each transmit antenna. The hilly terrain (HT) channel profile shown in Fig. 22(b) was used in the simulations. The initial value of $\hat{\mathbf{h}}_{i,n}^{(0)}$ was chosen as in (7.25). β_1 and β_2 were each chosen to be 0.5. The signal-to-noise ratio (SNR) is define as E_b/N_0, where E_b denotes engergy per information bit. It is shown that the EM algorithm converges to the ML estimate within 10-20 iterations on average for SNR $\leq$ 25dB, while the SAGE algorithm converges to the ML estimate within 4-6 iterations. The 9-tap STC estimator has a larger MSE than the ML estimator at SNR = 25dB due to the ignored channel taps. However, for SNR = 15dB, the 9-tap-STC has the least MSE. This is because the OFDM system can only resolve from the HT profile a discrete channel with a small number of significant taps with the other taps having very small values. At relatively low SNRs, estimating these taps with small values as done in ML and EM-type estimators may introduce more error than just assuming them to be zero as in the STC.

Fig. 24 shows the MSE performance of the EM-type algorithm with the initial value of $\hat{\mathbf{h}}_{i,n}^{(0)}$ chosen as the channel estimate of the previous OFDM block. With the better initial value, the EM estimator converges within 2-10 iterations and the SAGE estimator converges within 1-3 iterations depending on the SNR. An interesting phenomenon is that the EM estimate first reaches a lower MSE and then converges back to that of the ML estimate. This may seem odd at first sight, but one must remember that ML estimates, which are based on one OFDM block, do not necessarily
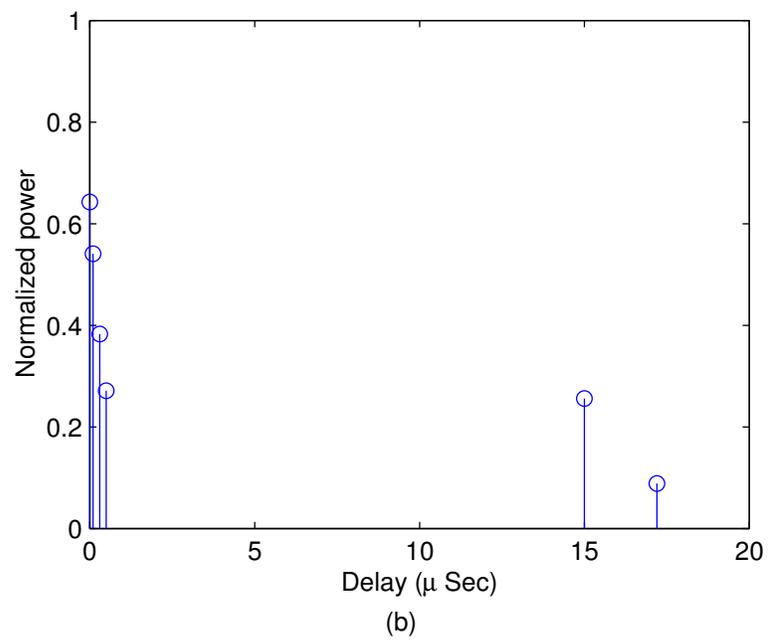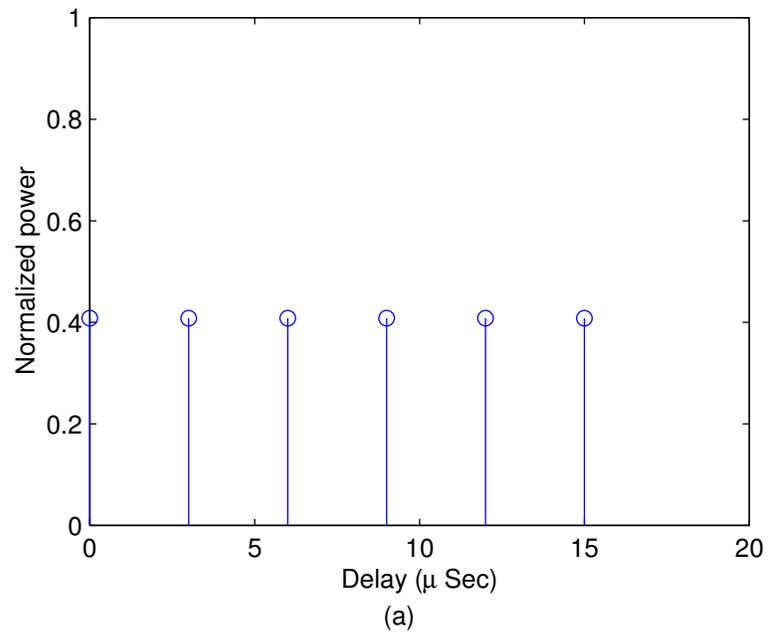
Fig. 22. (a) The uniform power and delay profile; (b) the hilly terrain profile.

Fig. 23. The convergence of MSE with respect to number of iterations of the EM-type estimators compared with the MSE of a 9-tap-STC estimator and the ML estimator. An initial channel estimate is obtained using (7.25). The HT profile as shown in (b) of Fig. 22 is assumed.
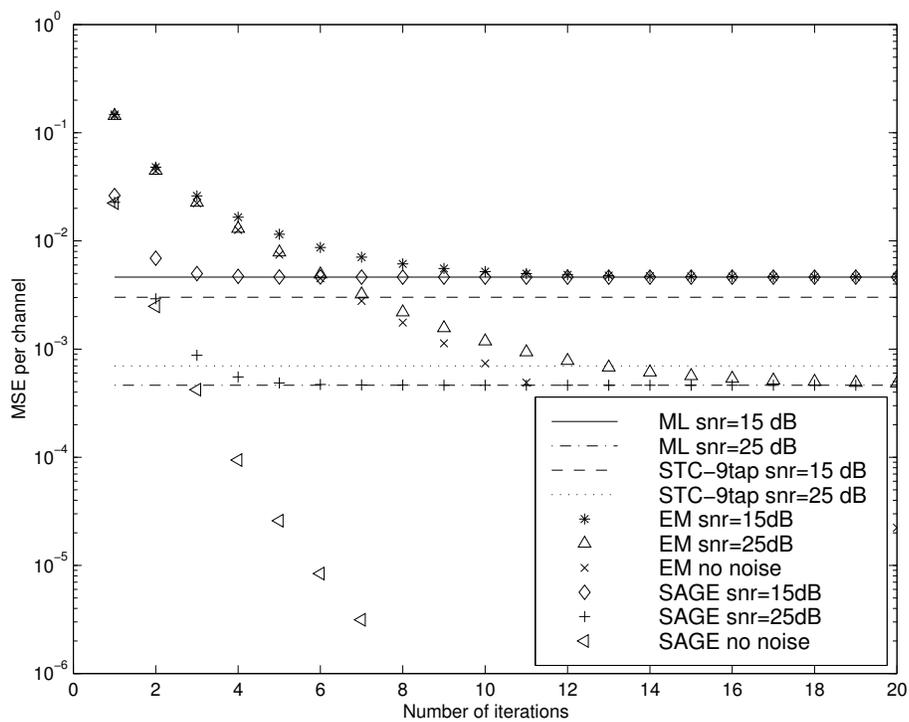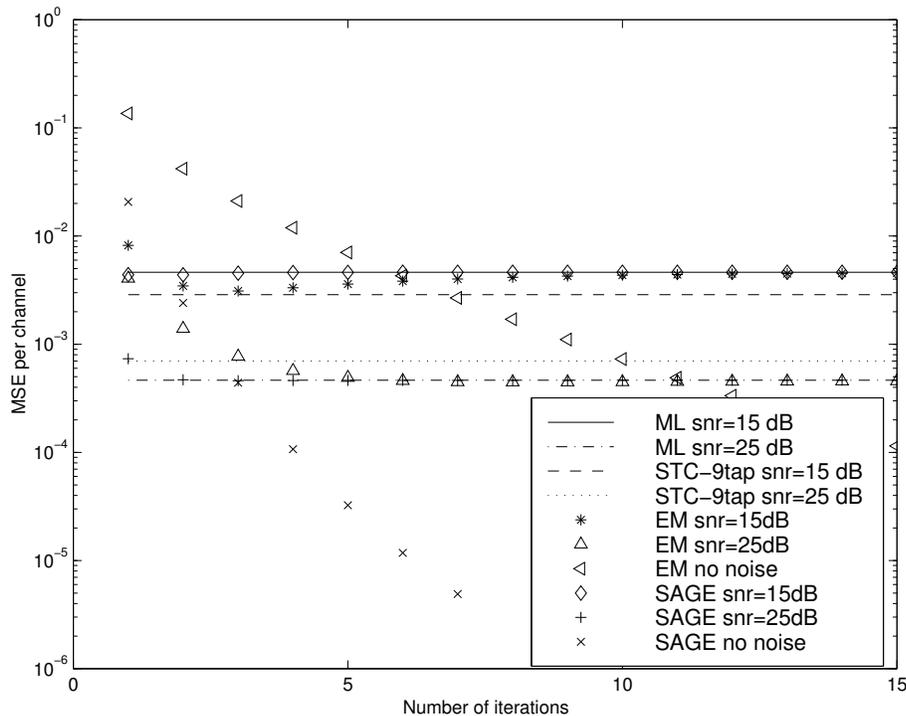
Fig. 24. The convergence of MSE with respect to number of iterations of the EM-type estimators compared with the MSE of a 9-tap-STC estimator and the ML estimator. The EM algorithm is initialized using the last channel estimate. The HT profile as shown in (b) of Fig. 22 is assumed.

minimize mean-square error.

Fig. 25 compares the MSE performance of different algorithms as a function of the number of iterations when the delay profile shown in Fig. 22(a) (taken from [75]) was used in the simulations. The initial value of $\hat{\mathbf{h}}_{i,n}^{(0)}$ was chosen as in (7.25). It is shown that the EM-type algorithms converge to the ML estimate with the same rate as in the case of HT profile. The 9-tap STC estimator has a significantly larger MSE than the ML estimator at both SNR = 15dB and SNR = 25dB.

Note that although the delays are uniformly spaced in this case, the interval is not an integer multiple of T_s. Therefore, leakage still exists.The 9-tap-STC has a significant MSE for both SNR = 15dB and SNR = 25 dB. This is because the ignored
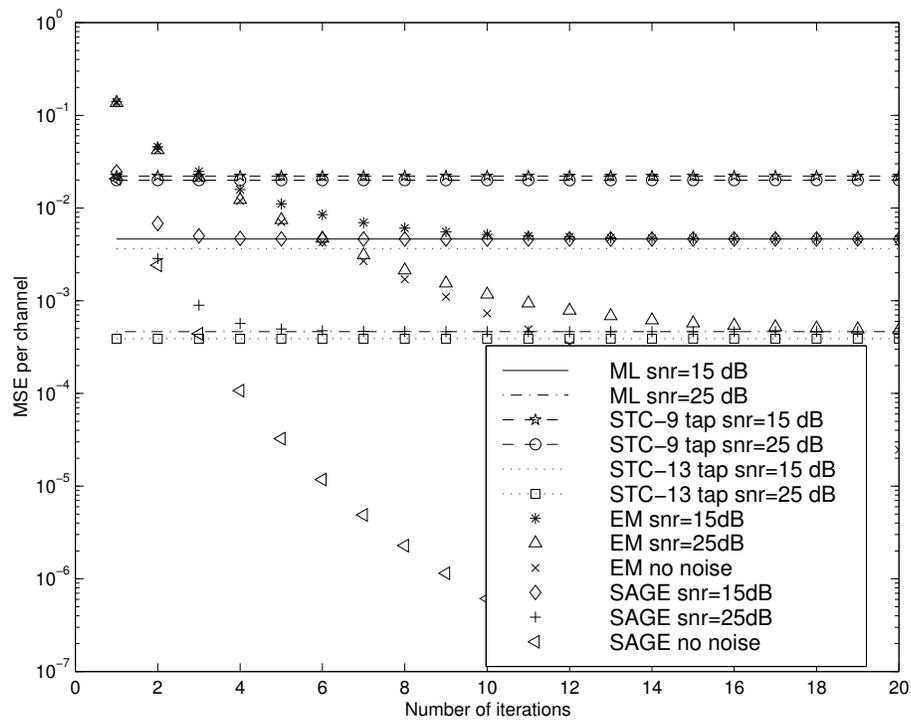
Fig. 25. The convergence of MSE with respect to number of iterations of the EM-type estimators compared with the MSE of a 9-tap-STC, 13-tap-STC estimator and the ML estimator. An initial channel estimate is obtained using (7.25). The uniform profile as shown in (a) of Fig. 22 is assumed.

channel taps for this channel power-delay profile are actually not all "insignificant". To reduce leakage in STC, we have to increase the number of significant taps to be 13, whose complexity is nearly tripled compared to that of the 9-tap-STC if direct matrix inversion is used. Note that the 13-tap-STC has a smaller MSE than that of the ML estimator even at SNR = 25dB. This suggests that almost all the power of this uniform channel profile is included in 13 taps.

Fig. 26 and Fig. 27 show the effect of channel length L_h and the number of transmit antennas M on the cumulative distribution function (CDF) of the convergence factor of EM-type algorithms. For the four antenna case, we used uncoded 4PSK due to the lack of proper space-time codes. The information bits were purely random. For all cases, the SAGE algorithm has a smaller convergence factor than that of the EM algorithm and is much less likely to encounter OFDM data blocks that cause very slow convergence (corresponding to convergence factor $\rho \approx 1$). The convergence rate of both algorithms decreases when M or L_h increases. Note that in Fig. 27, when L_h increases to 32, $\mathbf{Q}_n$ becomes a 128×128 matrix (note that $N = 128$); therefore, it is very likely that it is singular, causing the convergence factor to be near 1. Using (7.31), we can compute the asymptotic convergence factor of the EM algorithm. For example, we have $\tilde{\rho}_{em} = 7/8$ for $L_h = 16$ and $M = 2$, and $\tilde{\rho}_{em} = 0.9786$ for $L_h = 16$ and $M = 4$. In both cases, the asymptotic convergence factors roughly match the corresponding CDFs.

The bit-error-rate (BER) and word-error-rate (WER) of the ST-OFDM systems employing the 9-tap-STC estimator and the SAGE estimator, respectively, were compared assuming the HT channel profile. The results are shown in Fig. 28. In this and the following simulation, for both systems, a frame of 25 OFDM symbol blocks were transmitted from each antenna with the first OFDM block (known by the receiver) used for training. For the rest of the blocks, the decoded bits were encoded
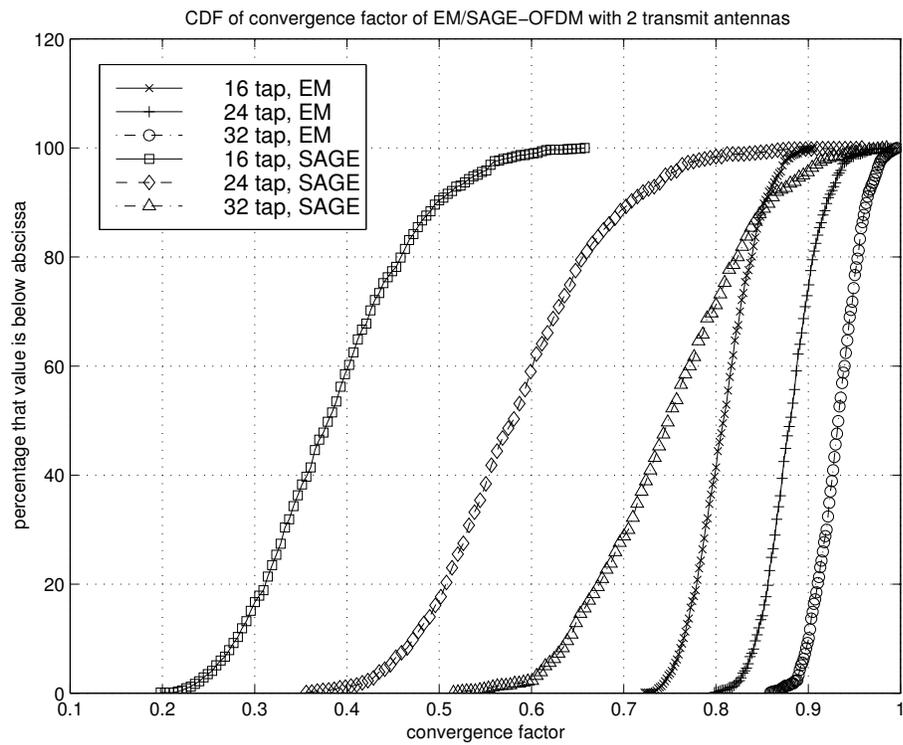
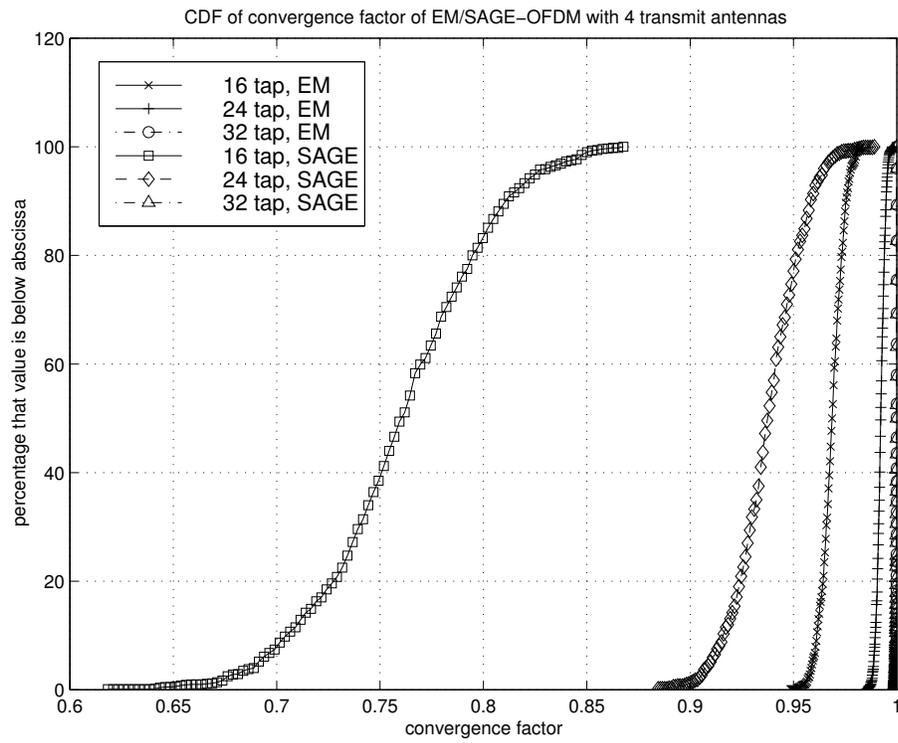Fig. 26. CDF of convergence factor of EM-type estimators with two transmit antennas used.

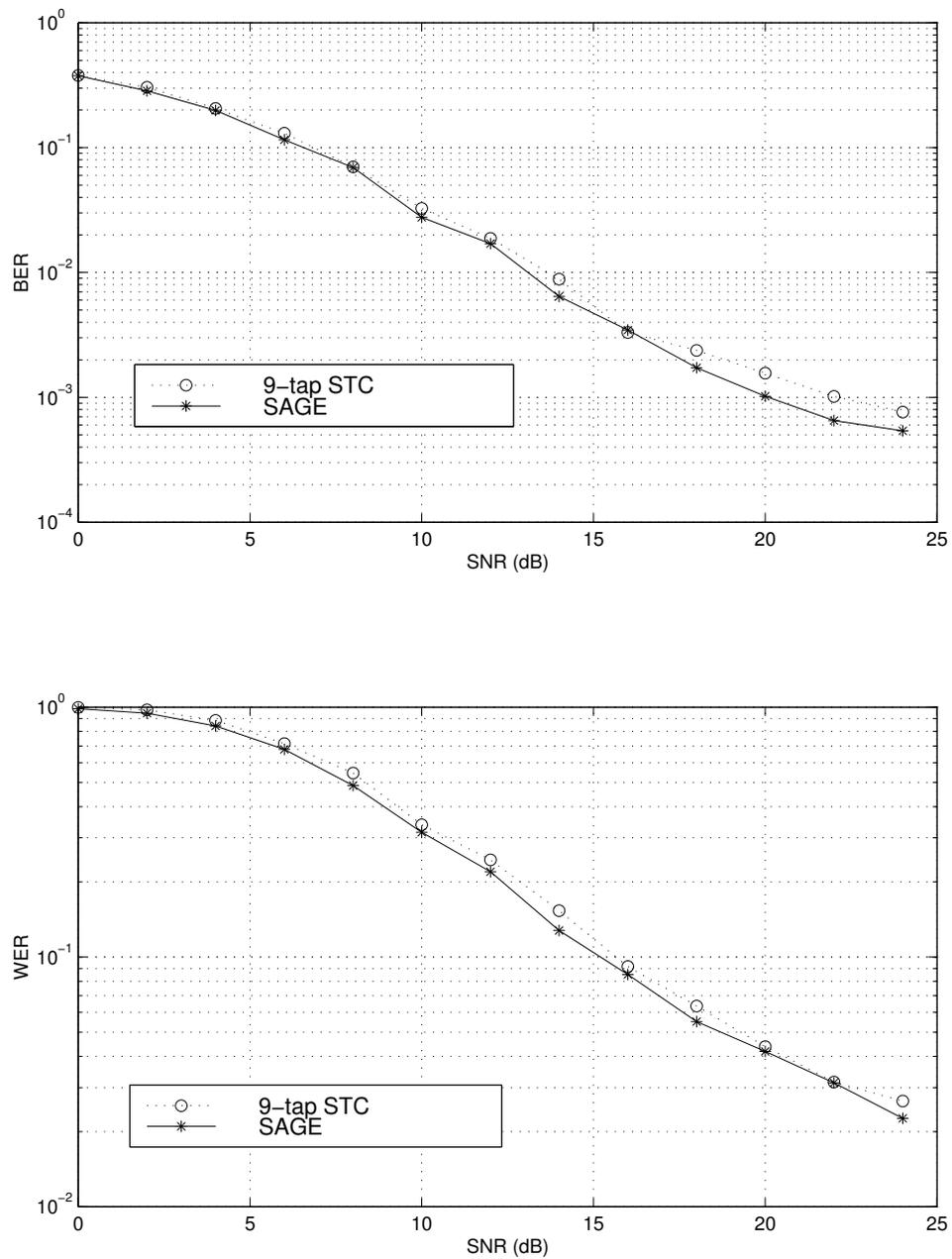Fig. 27. CDF of convergence factor of EM-type estimators with four transmit antennas used.

Fig. 28. BER and WER of ST-OFDM system for channels with the HT profile as shown in (b) of Fig. 22.

again and used as a reference for channel estimation. For the first training block, the SAGE algorithm used an initial estimate as shown in (7.25) and 10 iterations. For the remaining blocks, the previous estimate was used as initial value and 3 iterations were used.

Note that both systems show similar performance. It is not difficult to see that the discrete channel for the HT profile actually has very few significant taps. Therefore, the 9-tap STC is expected to perform well. The SAGE estimate in this case is actually as good as the ML estimate. It starts to outperform STC only at high SNRs. This is because at high SNR the channel estimation error caused by ignoring the insignificant taps becomes apparent.

The same simulations were performed for the uniform channel profile as shown in Fig. 22(a). The results are shown in Fig. 29. The system performance employing a 9-tap-STC has a large error floor due to the inaccurate channel estimation. The 13-tap-STC (direct matrix inversion) achieves almost the same performance as the SAGE estimator. The same performance can also be achieved by a 13-tap STC-SAGE with only 2 iterations used in the decision-directed channel tracking mode. The SAGE estimator assumes a complete channel model and therefore is robust to different kinds of power-delay profiles.

F.   Conclusion

Two efficient EM-based channel estimation algorithms for space-time coded OFDM systems are introduced and compared with each other in terms of convergence rate. We show that the convergence rates for both algorithms are unrelated with the channel delay profile, and the convergence rate for both algorithms decreases when the length of the channel or the number of transmit antennas increases. The EM-type es-
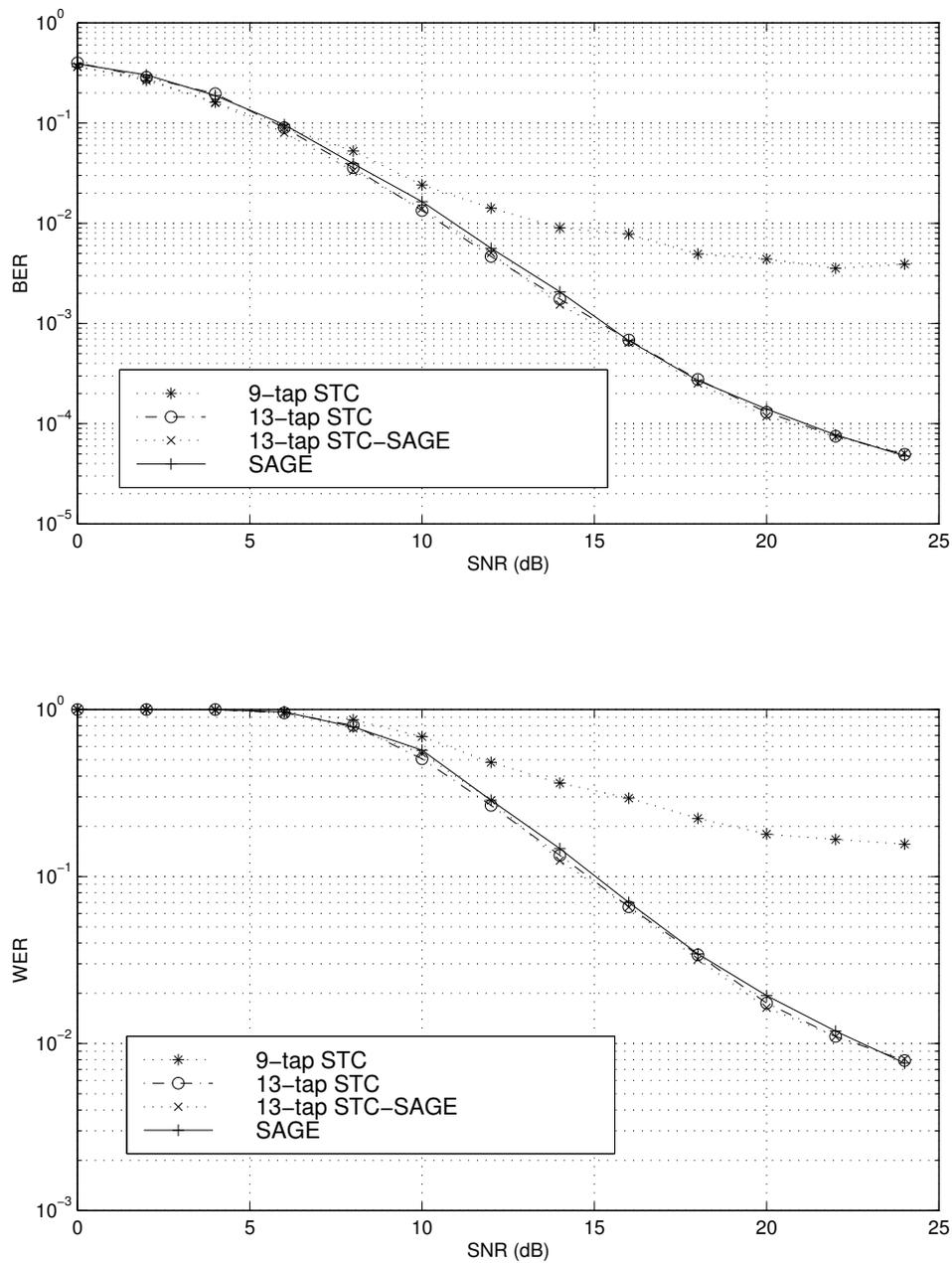
Fig. 29. BER and WER of ST-OFDM system for channels with the uniform profile as shown in (a) of Fig. 22.

timators can also be combined with the STC estimator when direct matrix inversion is computationally prohibitive in STC. Therefore, the resulting ST-OFDM system can perform well in various multipath channel profiles.

CHAPTER VIII

CONCLUSION

This work studies various topics in adaptive transmission, capacity analysis, signal detection and channel estimation for some important multiple antenna systems. The major contribution of this work is summarized as follows:

- We have evaluated the ergodic sum-rate capacity of the flat fading MIMO-BC with perfect CSIT, through both exact numerical computation and derived upper and lower bounds. Given fixed number of transmit and receive antennas in the MIMO-BC system, we show that as the number of users $K \to \infty$, the upper bound of the ergodic sum-rate capacity increases with the order of $O(\log(\ln(K)))$ asymptotically independent of the number of receive antennas. Sub-optimal transmission schemes, which use ranked known interference cancellation and zero-forcing beamforming to explore multi-user diversity, are shown to be able to achieve close to capacity performance.

- We have shown that the minimum outage probability transmission schemes for a flat MISO fading channel in the cases of mean feedback and covariance feedback are the same as the optimal schemes which maximize the ergodic capacity in terms of spatial directions. The optimum power allocation scheme over the optimal spatial directions which minimizes outage probability is closely related to the target rate. For both mean and covariance feedback, we show that it is more desirable to spread the power over all transmission directions than beamforming to a single direction for sufficiently small target rates.

- We have studied the optimal bandwidth allocation between the data channel and the feedback channel in a FDD MISO flat-fading system. Based on a sim-

ple but flexible prediction model, the maximum average achievable rate of the beamforming scheme for the data channel and the associated optimal bandwidth allocated to the feedback channel are evaluated under two different assumptions of the partial CSIT. Additionally, we proposed a lower bound on the average achievable rate of the beamforming scheme using quantized feedback bits, which can be used as a performance benchmark for practical channel quantization schemes.

- We have proposed a novel low complexity MIMO detector, which uses both the SAGE detection algorithm and the List-BLAST detection algorithm. We show that the List-BLAST algorithms (including the List-Ranked-BLAST and the List-Shifted-BLAST) alone can achieve performance close to the ML detection. The SAGE algorithm can be used in combination with the List-BLAST algorithm to further improve the system performance

- We have proposed two efficient EM-based channel estimation algorithms for OFDM systems with transmit diversity to be used in the decision-directed tracking mode. We show that the convergence rate for both algorithms is unrelated with the channel delay profile, and decreases when the length of the channel or the number of transmit antennas increases.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Guassian channels," *European Trans. Telecommun. (ETT)*, vol. 10, no. 6, pp. 585-596, Nov. 1999.

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

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

[4] E. Biglieri, J. Proakis and S. Shamai (Shitz), "Fading channels: information-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

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

[6] W. Yu, W. Rhee and J. M. Cioffi, "Optimal power control in multiple access fading channels with multiple antennas," in *Proc. IEEE Int. Conf. Commun.*, Helsinki, Finland, June 2001, pp. 575-579.

[7] 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.

[8] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456-1467, July 1998.

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

[10] B. Hassibi and B. Hochwald, "High rate codes that are linear in space and time," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1804-1824, July 2002.

[11] M. O. Damen, H. El-Gamal and G. Caire, "On maximum likelihood detection and the search for the closest lattice point," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2389-2402, Oct. 2003.

[12] B. M. Hochwald and S. T. Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, no. 3, pp 389-399, March 2003.

[13] G. J. Foschini,"Layered space-time architechture for wireless communication in a fading environment when using multielement antennas," *Bell Labs. Tech. J.*, vol. 1, no. 2, pp. 41-59, Autumn 1996.

[14] J. Benesty, Y. Huang and J. Chen, "A fast recursive algorithm for optimal sequential signal detection in a BLAST system," *IEEE Trans. Signal Processing*, vol. 51, no. 7, pp. 1722-1730, July 2003.

[15] K. Rohani, M. Harrison and K. Kuchi, "A comparison of base station transmit diversity methods for third generation cellular standards," in *Proc. IEEE Vehicular Technology Conf.*, Houston, TX, 1999, pp. 351-355.

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

[17] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 433-441, May 1983.

[18] R. Zamir, S. Shamai and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1250-1276, June 2002.

[19] R. Knopp and P. A. Humblet, "Information capacity and power control in single-cell multiuser communications," in *Proc. IEEE Int. Conf. Commun.*, Seattle, WA, June 1995, pp. 331-335.

[20] W. Yu and J. M. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inform. Theory*, submitted, Nov. 2001. Available: www.comm.toronto.edu/weiyu/publications.html. Accessed: Dec. 2002.

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

[22] P. Viswanath and D. Tse, "Sum capacity of the multiple antenna Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inform. Theory*, vol. 49, no. 8, pp. 1912-1921, Aug. 2003.

[23] Y. Xie and C. N. Georghiades, "Some results on the sum-rate capacity of MIMO fading broadcast channels," in *Proc. IEEE International Symposium on Advances in Wireless Communications*, Victoria, British Columnbia, Canada, Sept. 2002, pp. 4-5.

[24] M. Brandt-Pearce and A. Dharap, "Transmitter-based multiuser interference rejection for the down-link of a wireless CDMA system in a multipath environ-ment," *IEEE J. Selected Areas in Commun.*, vol. 18, no. 3, pp. 407-417, Mar. 2000.

[25] B. Hochwald and S. Vishwanath, "Space-Time multiple access: linear growth in sum rate," in *Proc. Allerton Conf. on Communications, Control and Computing*, Monticello, IL, Oct. 2002, pp. 387-396.

[26] H. C. Huang, S. Venkatesan and H. Viswanathan, "Downlink capacity evaluation of cellular networks with known interference cancellation," *IEEE J. Selected Areas in Commun.*, vol. 21, no. 5, pp. 802-811, June 2003.

[27] W. Yu, W. Rhee, S. Boyd and J. Cioffi, "Iterative water-filling for Gaussian vector multiple access channels," *IEEE Trans. Inform. Theory*, vol. 50, no. 1, pp. 145-152, Jan. 2004.

[28] G. Caire, G. Taricco and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468-1489, July 1999.

[29] N. Jindal, S. Jafar, S. Vishwanath and A. Goldsmith, "Sum power waterfilling for Gaussian broadcast channels," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2002, pp. 1518-1522.

[30] L. Vandenberghe, S. Boyd and S. P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 19, no. 2, pp. 499-533, 1998.

[31] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U. K.: Cambridge University Press, 1985.

[32] H. A. David, *Order Statistics*. New York: Wiley-Interscience, 1970.

[33] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1277-1294, June 2004.

[34] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis* (Second Edition). New York: John Wiley & Sons, 1984.

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

[36] E. Biglieri, G. Caire and G. Taricco,"Limiting performance of block-Fading channels with multiple antennas," *IEEE Trans. Inform. Theory*, vol. 47, no. 4, pp. 1273 - 1289, May 2001.

[37] A. Narula, M. J. Lopez, M. D. Trott and G. W. Wornell, "Efficient use of side information in multiple antenna data transmission over fading channels", *IEEE Trans. Inform. Theory*, vol. 16, no. 8, pp. 1423-1435, Oct. 1998

[38] E. Vistotsky, U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632-2639, Sept. 2001.

[39] G. G. Raleigh and V. K. Jones, "Adaptive antenna transmission for frequency duplex digital wireless communications," in *Proc. IEEE Int. Conf. Commun.*, vol. 2, June 1997, pp. 641-646.

[40] S. A. Jafar, S. Vishwanath and A. J. Goldsmith, "Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback," in *Proc. IEEE Int. Conf. Commun.*, vol.7, June 2001, pp. 2266 – 2270.

[41] S. H. Simon and A. L. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Selected Areas in Commun.*, vol. 21, no. 3, pp. 406-417, Apr. 2003.

[42] S. A. Jafar and A. J. Goldsmith, "On optimality of beamforming for multiple antenna systems with imperfect feedback," in *Proc. IEEE Int. Symp. Inform. Theory*, Washington DC, June 2001, p. 321.

[43] E. Jorswieck and H. Boche, "On transmit diversity with imperfect channel state information," in *Proc. IEEE ICASSP*, Orlando, FL, May 2002, pp. 2181-2184.

[44] A. L. Moustakas and S. H. Simon, "Optimizing multi-transmitter-single-receiver (MISO) communication systems with general Gaussian channels: nontrivial covariance and nonzero mean," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2770-2780, Oct. 2003.

[45] Y. Xie, C. N. Georghiades and A. Arapostathis, "Outage probability for MISO fading channels with imperfect feedback," in *Proc. IEEE Globecom*, San Francisco, CA, Dec. 2003, pp. 1674-1678.

[46] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press, 1979.

[47] B. Hassibi, B. H. Hochwald, "How much training is needed in multiple antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951-963, Apr. 2003.

[48] E. Vistotsky, U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632-2639, Sept. 2001.

[49] P. P. Vaidyannathan, "On predicting a band-limited signal based on past sample values," *Proc. of IEEE*, vol. 75, pp. 1125-1127, Aug. 1987.

[50] K. K. Mukkavilli, A. Sabharwal, E. Erkip and B. Aazhang, "On beamforming with finite rate feedback in multiple antenna systems," *IEEE Trans. Inform.*

Theory, vol.49, no. 10, pp. 2562-2579, Oct. 2002.

[51] M. G. Kendall, *A Course in the Geometry of n Dimensions.* Griffin's Statistical Monographs & Courses, New York: Hafner Pub. Co, 1961.

[52] J. G. Proakis, *Digital Communications, Third Edition.* New York: McGraw-Hill, 1995.

[53] M. O. Damen, K. Abed-Meraim and M. S. Lemdani, "Further results on the sphere decoder," in *Proc. IEEE Int. Symp. Inform. Theory*, Washington DC, June, 2001, p. 333.

[54] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, pp. 1-17, 1977.

[55] M. Feder, E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 36, no. 4, pp. 477-489, Apr. 1988.

[56] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2664-2677, Oct. 1994.

[57] Y. Xie and C. N. Georghiades, "Two EM-based channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106-115, Jan. 2003.

[58] P. Spasojevic, "Sequence and channel estimation for channels with memory," Ph.D. dissertation, Dept. Elec. Eng., Texas A&M University, College Station, TX, Dec 1999. Available: ee-wcl.tamu.edu/formerstudents.html.

[59] H. Artes, D. Seethaler and F. Hlawatsch, "Efficient detection algorithms for MIMO channels: a geometrical approach to approximate ML detection," *IEEE Trans. Signal Processing*, vol 51, no. 11, pp. 2808-2820, Nov 2003.

[60] B. Hassibi and H. Vikalo, "On the expected complexity of sphere decoding" in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2001, pp. 1051–1055,

[61] B. Hassibi and H. Vikalo, "On the sphere decoding algorithm: Part II, generalizations, second-order statistics, and applications to communications," *IEEE Trans. Signal Processing*, submitted, 2004. Available: www.systems.caltech.edu/EE/Faculty/babak/pubs/journal.html. Accessed: May 2004.

[62] R. V. Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Boston: Artech House Publishers, 2000.

[63] L. J. Cimini, J.C. Chuang, and N.R. Sollenberger, "Advanced cellular internet service (ACIS)," *IEEE Commun. Magazine*, vol. 36, pp. 150-159, Oct. 1998.

[64] L. J. Cimini, B. Daneshrad, and N.R. Sollenberger, "Clustered OFDM with transmitter diversity and coding," in *Proc. 1996 IEEE Globecom*, London, U.K., Nov 1996, pp. 703-707.

[65] D. Agrawal, V. Tarokh, A. Naguib, N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. IEEE Vehicular Technology Conference*, Ottawa, Ontario, Canada, May 1998, pp. 2232-2236.

[66] Y. Li, J. C. Chuang, N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high-rate data wireless networks," *IEEE J. Selected Areas in Commun.*, vol. 17, no. 7, pp. 1233-1243, July 1999.

[67] Y. Li, N. Seshadri, L. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channel," *IEEE J. Selected Areas in Commun.*, vol. 17, no. 3, pp. 461-471, Mar. 1999.

[68] J. J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, "On channel estimation in OFDM systems," in *Proc. IEEE Vehicular Technology Conference*, Chicago, IL, July 1995, pp. 815-819.

[69] Y. Li, "Optimum training sequences for OFDM systems with multiple transmit antennas," in *Proc. IEEE Globecom*, San Francisco, CA, Nov. 2000, pp. 1478-1482.

[70] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Dept. Math., Massachusetts Institute of Technology, Cambridge, MA, 1989.

[71] J. E. Cohen, H. Kesten and C. M. Newman, *Random Matrices and Their Applications.* Contemporary Mathematics, American Mathematical Society, Providence, RI, vol. 50, 1984.

[72] T. J. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 599-618, Feb. 2001.

[73] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses Theory and Applications*, New York: John Wiley & Sons, Inc., 1974.

[74] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing: Principles, Algorithms, and Applications.* Upper Saddle River, NJ: Prentice Hall, Inc., 1996.

[75] R. Steele, *Mobile Radio Communications.* London, U. K.: Pentech Press 1992.

APPENDIX A

FORMALIZATION OF THE OPTIMIZATION PROBLEM

The authors in [30] proposed an efficient solution of the following constrained determinant maximization problem:

$$\min_{X} C^T X + \log \left| \mathbf{G}(X)^{-1} \right| \tag{A.1}$$

subject to:

$$\mathbf{G}(X) \succ \mathbf{0} \tag{A.2}$$

$$\mathbf{F}(X) \succeq 0, \tag{A.3}$$

where $\mathbf{G}(X) = \mathbf{G}_0 + \sum_{i=1}^{m} x_i \mathbf{G}_i$ and $\mathbf{F}(X) = \mathbf{F}_0 + \sum_{i=1}^{m} x_i \mathbf{F}_i$; C is some real constant vector having the same dimension as X. $X = [x_1, x_2, \ldots, x_m]^T$ is required to be a real vector. To transform the optimization problem of equations (3.4)-(3.6) to the above format, we can drop the expectation, and consider the equivalent problem of $\max_{\mathbf{S}_k} \log |\Psi|$ for some given $\mathbf{H}$. Let $X = [RV(\mathbf{S}_1), RV(\mathbf{S}_2), \cdots, RV(\mathbf{S}_K)]^T$, where $RV(\mathbf{A})$ denotes the $N^2 \times 1$ real vector formed by stacking the N diagonal entries (which are real), the real and the imaginary part of the $N(N-1)/2$ lower-triangular off-diagonal entries of an $N \times N$ Hermitian matrix $\mathbf{A}$. For any vector $z \in \mathbb{C}^n$ and matrix $\mathbf{A} \in \mathbb{C}^{n \times m}$, define $\hat{z} = [Re(z) Im(z)]$ and $\hat{\mathbf{A}} = \begin{bmatrix} Re(\mathbf{A}) & -Im(\mathbf{A}) \\ Im(\mathbf{A}) & Re(\mathbf{A}) \end{bmatrix}$. According to Lemma 1 of [1], we have $\log |\Psi| = 0.5 \log |\hat{\Psi}|$. Comparing with the standard objective function of equation (A.1), we have $\mathbf{G}(X) = \hat{\Psi}$ and $C = \mathbf{0}$. Note that any matrix $\mathbf{A}$ can be written as $\mathbf{A} = \sum_{n,m} \mathbf{E}_{n,m} \mathbf{A}[n, m]$, where n and m are row and

column indexes, respectively; $\mathbf{E}_{n,m}$ is the natural basis of the matrix, or the matrix with only one non-zero entry $\mathbf{E}_{n,m}[n,m] = 1$. Since each entry of the matrix $\hat{\Psi}$ is the linear combination of the elements in X, $\mathbf{G}_i$ can be determined by identifying the positions of $X[i]$ in $\hat{\Psi}$ with its associated linear coefficients, and then constructed using the corresponding linear combination of $\mathbf{E}_{n,m}$.

The constraints of both equation (3.5) and (3.6) can also be transformed into the desired format of $\mathbf{F}(X)$. To simplify illustration, we consider the case where $\mathbf{S}_k$ is real, $K = 2$ and $N_r = 2$. We have $\mathbf{F} = \mathbf{F}'$, where

$$\mathbf{F}' = \begin{pmatrix} \mathbf{S}_1[1,1] & \mathbf{S}_1[1,2] & 0 & 0 & 0 \\ \mathbf{S}_1[2,1] & \mathbf{S}_1[2,2] & 0 & 0 & 0 \\ 0 & 0 & \mathbf{S}_2[1,1] & \mathbf{S}_2[1,2] & 0 \\ 0 & 0 & \mathbf{S}_2[2,1] & \mathbf{S}_2[2,2] & 0 \\ 0 & 0 & 0 & 0 & t \end{pmatrix}, \tag{A.4}$$

where $t = P - \mathbf{S}_1[1,1] - \mathbf{S}_1[2,2] - \mathbf{S}_2[1,1] - \mathbf{S}_2[2,2]$. For other values of K and N_r, $\mathbf{F}'$ can be similarly constructed. We can easily prove that the constraints as shown in (3.5) and (3.6) are equivalent to $\mathbf{F} \succeq \mathbf{0}$.

In order to deal with the case where $\mathbf{S}_k$ is complex, we use Corollary 2 of [1], which states that $\mathbf{F} \succeq \mathbf{0}$ and $\hat{\mathbf{F}} \succeq \mathbf{0}$ are equivalent. Therefore, in this case, the constraints of (3.5) and (3.6) are equivalent to $\mathbf{F} = \hat{\mathbf{F}}' \succeq \mathbf{0}$. $\mathbf{F}_i$ associated with x_i can be similarly determined as determining $\mathbf{G}_i$. Since $\mathbf{F} \succeq \mathbf{0}$ guarantees $\mathbf{G}(X) \succ \mathbf{0}$ in our problem, $\mathbf{G}(X) \succ \mathbf{0}$ does not need to be implemented explicitly as in the standard format.

APPENDIX B

PROOF OF LEMMA 1 IN CHAPTER IV

We restate Lemma 1 as follows:

Given $X \sim \mathcal{N}(\bar{x}, 1)$, $Y \sim \mathcal{N}(\bar{y}, 1)$, X, Y independent, and $\gamma > 1$, then

$$\min_{\bar{x}^2 + \bar{y}^2 = m^2} \Pr(X^2 + \gamma^{-2} Y^2 < q^2)$$

is attained at $\bar{x} = m$, $\bar{y} = 0$.

Proof. Parameterize $\bar{x}$ and $\bar{y}$, as $\bar{x} = m\cos(\varphi)$, $\bar{y} = m\sin(\varphi)$, with $\varphi \in \left[0, \frac{\pi}{2}\right]$, and let

$$\mathcal{D} \triangleq \left\{ (x, y) \in \mathbb{R}^2 \mid x^2 + \gamma^{-2} y^2 < q^2 \right\}.$$

Thus,

$$\mathcal{I}(\varphi) \triangleq \Pr(X^2 + \gamma^{-2} Y^2 < q^2) = \frac{1}{2\pi} \iint_{\mathcal{D}} e^{-\frac{(x-\bar{x})^2}{2}} e^{-\frac{(y-\bar{y})^2}{2}} \, dx \, dy$$
$$= \frac{1}{2\pi} \iint_{\mathcal{D}} e^{-\frac{(x - m\cos(\varphi))^2}{2}} e^{-\frac{(y - m\sin(\varphi))^2}{2}} \, dx \, dy. \tag{B.1}$$

We convert to polar coordinates $x = r\cos(\theta)$, $y = r\sin(\theta)$, $\theta \in [0, 2\pi)$. Then, $\mathcal{D}$ takes the form

$$\mathcal{D} = \left\{ (r, \theta) \in \mathbb{R}^2 \mid r^2 \cos^2(\theta) + \frac{r^2}{\gamma^2} \sin^2(\theta) < q^2 \right\}.$$

For $r \in [0, \gamma q]$, let

$$\Theta_r \triangleq \left\{ \theta \in [0, 2\pi) \mid r^2 \cos^2(\theta) + \frac{r^2}{\gamma^2} \sin^2(\theta) < q^2 \right\}.$$

The integral in (B.1) transforms to

$$\mathcal{I}(\varphi) = \frac{1}{2\pi} \int_0^{\gamma q} \left[\int_{\Theta_r} e^{-\frac{(r\cos(\theta) - m\cos(\varphi))^2}{2}} e^{-\frac{(r\sin(\theta) - m\sin(\varphi))^2}{2}} \, d\theta \right] r \, dr. \tag{B.2}$$

Note that

$$
\Theta_r = \begin{cases} [0, 2\pi), & \text{if } 0 \le r \le q, \\[2mm] \left[\alpha_r, \frac{\pi}{2}\right] \cup \left[\frac{\pi}{2}, \pi - \alpha_r\right] \cup \left[\pi + \alpha_r, \frac{3\pi}{2}\right] \cup \left[\frac{3\pi}{2}, 2\pi - \alpha_r\right], & \text{if } q < r \le \gamma q, \end{cases}
\tag{B.3}
$$

where

$$
\alpha_r = \cos^{-1}\left(\sqrt{\tfrac{\gamma^2 q^2 - r^2}{r^2(\gamma^2 - 1)}}\right), \quad q < r \le \gamma q, \quad \alpha_r \in \left[0, \tfrac{\pi}{2}\right].
$$

Decomposing the inner integral in (B.2) along the partition of Θ_r in (B.3) and using the symmetry properties of the functions, we obtain

$$
\int_{\Theta_r} e^{-\frac{(r\cos(\theta) - m\cos(\varphi))^2}{2}} e^{-\frac{(r\sin(\theta) - m\sin(\varphi))^2}{2}} \, d\theta = 2 \int_{\alpha_r}^{\frac{\pi}{2}} e^{-\frac{m^2 + r^2}{2}} M_\varphi(r, \theta) \, d\theta, \quad r \in (q, \gamma q],
$$

where

$$
M_\varphi(r, \theta) \triangleq 2\cosh\big(mr\cos(\varphi)\cos(\theta)\big)\cosh\big(mr\sin(\varphi)\sin(\theta)\big)
$$
$$
= \cosh\big(mr\cos(\theta - \varphi)\big) + \cosh\big(mr\cos(\theta + \varphi)\big).
\tag{B.4}
$$

Therefore, by (B.2),

$$
\mathcal{I}(\varphi) = \frac{1}{2\pi} \int_0^q \left[\int_0^{2\pi} e^{-\frac{m^2 + r^2}{2}} e^{mr\cos(\theta - \varphi)} \, d\theta\right] r \, dr + \frac{1}{\pi} \int_q^{\gamma q} \left[\int_{\alpha_r}^{\frac{\pi}{2}} e^{-\frac{m^2 + r^2}{2}} M_\varphi(r, \theta) \, d\theta\right] r \, dr.
\tag{B.5}
$$

The first integral in (B.5) is independent of φ and is equal to

$$
\frac{1}{2\pi} \int_0^q \left[\int_0^{2\pi} e^{mr\cos(\theta)} \, d\theta\right] e^{-\frac{m^2 + r^2}{2}} r \, dr.
$$

Hence, in order to establish that $\mathcal{I}(\varphi) \ge \mathcal{I}(0)$ for all $\varphi \in \left[0, \frac{\pi}{2}\right]$, it is sufficient to show that

$$
\int_{\alpha_r}^{\frac{\pi}{2}} M_\varphi(r, \theta) \, d\theta \ge \int_{\alpha_r}^{\frac{\pi}{2}} M_0(r, \theta) \, d\theta, \qquad \forall \varphi \in \left[0, \tfrac{\pi}{2}\right], \quad \forall r \in (q, \gamma q].
$$

We have,

$$\int_{\alpha_r}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta + \varphi)\big)\, d\theta = \int_{\alpha_r+\varphi}^{\frac{\pi}{2}+\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta$$

$$= \int_{\alpha_r+\varphi}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta)\big)\, d\theta + \int_{\frac{\pi}{2}}^{\frac{\pi}{2}+\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta$$

$$= \int_{\alpha_r+\varphi}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta)\big)\, d\theta + \int_{\frac{\pi}{2}-\varphi}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta)\big)\, d\theta,$$

$$\text{(B.6)}$$

and

$$\int_{\alpha_r}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta - \varphi)\big)\, d\theta = \int_{\alpha_r-\varphi}^{\frac{\pi}{2}-\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta$$

$$= \int_{\alpha_r-\varphi}^{\alpha_r+\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta + \int_{\alpha_r+\varphi}^{\frac{\pi}{2}-\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta.$$

$$\text{(B.7)}$$

By (B.4), (B.6) and (B.7),

$$\int_{\alpha_r}^{\frac{\pi}{2}} M_\varphi(r, \theta)\, d\theta = \int_{\alpha_r-\varphi}^{\alpha_r+\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta + 2\int_{\alpha_r+\varphi}^{\frac{\pi}{2}} \cosh\big(mr\cos(\theta)\big)\, d\theta. \quad \text{(B.8)}$$

Therefore, by (B.8),

$$\int_{\alpha_r}^{\frac{\pi}{2}} M_\varphi(r, \theta)\, d\theta - \int_{\alpha_r}^{\frac{\pi}{2}} M_0(r, \theta)\, d\theta = \int_{\alpha_r-\varphi}^{\alpha_r} \cosh\big(mr\cos(\theta)\big)\, d\theta - \int_{\alpha_r}^{\alpha_r+\varphi} \cosh\big(mr\cos(\theta)\big)\, d\theta$$

$$= \int_0^\varphi \big[\cosh\big(mr\cos(\alpha_r - \delta)\big) - \cosh\big(mr\cos(\alpha_r + \delta)\big)\big]\, d\delta$$

$$= 2\int_0^\varphi \sinh\big(mr\cos(\alpha_r)\cos(\delta)\big) \sinh\big(mr\sin(\alpha_r)\sin(\delta)\big)\, d\delta$$

$$\geq 0.$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

APPENDIX C

DERIVATION OF THE CONVERGENCE MATRIX FOR THE EM AND SAGE

ALGORITHM IN CHAPTER VII

We first derive the convergence matrix for the EM algorithm as shown in equation (7.27). Assume $\beta_i = \beta = 1/M$ for $i = 1, 2, \ldots, M$. (7.19) can be re-written as

$$\hat{\mathbf{h}}_n^{(k+1)} = \hat{\mathbf{h}}_n^{(k)} + \beta \mathbf{T}[(\mathbf{1}_M \bigotimes \mathbf{r}_n) - (\mathbf{1}_M \bigotimes \mathbf{G}_n)\hat{\mathbf{h}}_n^{(k)}], \qquad (C.1)$$

where $\mathbf{1}_M$ is a $M \times 1$ vector with all entries equal to 1 and $\bigotimes$ denotes the Kroenecker product; $\mathbf{T}$ is defined as follows:

$$\mathbf{T} \stackrel{\triangle}{=} \mathrm{diag}(\mathbf{F}^H \mathbf{X}_{1,n}^H \ \ \mathbf{F}^H \mathbf{X}_{1,n}^H \cdots \mathbf{F}^H \mathbf{X}_{M,n}^H), \qquad (C.2)$$

where $\mathrm{diag}(\cdot)$ denotes block diagonal matrix appropriately formed.

Substituting (C.1) into (7.26), we have

$$[\mathbf{A}_n + \beta \mathbf{T}(\mathbf{1}_M \bigotimes \mathbf{G}_n) - \mathbf{I}_{ML_h}]\hat{\mathbf{h}}_n^{(k)} = \beta \mathbf{T}(\mathbf{1}_M \bigotimes \mathbf{r}_n) + (\mathbf{A}_n - \mathbf{I}_{ML_h})\hat{\mathbf{h}}_n^{<ML>}. \quad (C.3)$$

(C.3) is satisfied for any value of $\hat{\mathbf{h}}_n^{(k)}$, so that we have

$$\mathbf{A}_n = \mathbf{I}_{ML_h} - \beta \mathbf{T}(\mathbf{1}_M \bigotimes \mathbf{G}_n) = \mathbf{I}_{ML_h} - \frac{1}{M}\mathbf{Q}_n. \qquad (C.4)$$

Substituting (C.4) into the RHS of (C.3), we can verify that the RHS also equals zero.

Next, we consider the SAGE Algorithm. For the kth iteration and the corresponding i, where $i = 1 + [k \bmod M]$, assume we have

$$\hat{\mathbf{h}}_n^{(k+1)} - \hat{\mathbf{h}}_n^{<ML>} = \mathbf{A}_n^{(i)}(\hat{\mathbf{h}}_n^{(k)} - \hat{\mathbf{h}}_n^{<ML>}). \tag{C.5}$$

Consider the case $k = 0$ and $i = 1$. Combining (7.20) - (7.24), we have

$$\hat{\mathbf{h}}_n^{(1)} = \mathbf{R}_1 + \mathbf{S}_1\hat{\mathbf{h}}_n^{(0)} \tag{C.6}$$

where

$$\mathbf{R}_1 \triangleq [\mathbf{r}_n^H \mathbf{X}_{1,n}\mathbf{F} \quad \mathbf{0} \quad \cdots \quad \mathbf{0}]^H \tag{C.7}$$

$$\mathbf{S}_1 \triangleq \begin{pmatrix} \mathbf{0} & -\mathbf{F}^H\mathbf{X}_{1,n}^H\mathbf{X}_{2,n}\mathbf{F} & -\mathbf{F}^H\mathbf{X}_{1,n}^H\mathbf{X}_{3,n}\mathbf{F} & \cdots & -\mathbf{F}^H\mathbf{X}_{1,n}^H\mathbf{X}_{M,n}\mathbf{F} \\ \mathbf{0} & \mathbf{I}_{L_h} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I}_{L_h} & \mathbf{0} \\ \mathbf{0} & & \cdots & \mathbf{0} & \mathbf{I}_{L_h} \end{pmatrix}. \tag{C.8}$$

Substituting (C.6) into (C.5), we have

$$(\mathbf{S}_1 - \mathbf{A}_n^{(1)})\hat{\mathbf{h}}_n^{(0)} = -\mathbf{R}_1 + (\mathbf{I}_{ML_h} - \mathbf{A}_n^{(1)})\hat{\mathbf{h}}_n^{<ML>}. \tag{C.9}$$

(C.9) is true for any value of $\hat{\mathbf{h}}_n^{(0)}$, so that we have

$$\mathbf{A}_n^{(1)} = \mathbf{S}_1 = \mathbf{I}_{ML_h} - (\mathbf{e}_1 \cdot \mathbf{e}_1^H)\mathbf{Q}_n. \tag{C.10}$$

Substitute (C.10) into RHS of (C.9), we can verify that the RHS also equals to zero. Similarly, we can show

$$\mathbf{A}_n^{(i)} = \mathbf{I}_{ML_h} - (\mathbf{e}_i \cdot \mathbf{e}_i^H)\mathbf{Q}_n, \tag{C.11}$$

where $\mathbf{e}_i = [\mathbf{0}\cdots\mathbf{I}_{L_h}\cdots\mathbf{0}]^H$, is a $ML_h \times L_h$ block matrix, with the ith $L_h \times L_h$ block equal to $\mathbf{I}_{L_h}$. $\mathbf{A}_n$ for the SAGE algorithm can be shown as (according to [56]),

$$\mathbf{A}_n = \prod_{i=1}^{M} \mathbf{A}_n^{(i)} = \mathbf{I}_{ML_h} - LT(\mathbf{Q}_n)^{-1}\mathbf{Q}_n. \tag{C.12}$$

If in the SAGE algorithm the message update at the check node is performed simultaneously for each variable node, we can show

$$\hat{\mathbf{h}}_n^{(k+1)} = \mathbf{T}(\mathbf{1}_M \bigotimes \mathbf{r}_n) + (\mathbf{I}_{ML_h} - \mathbf{Q}_n)\hat{\mathbf{h}}_n^{(k)}. \tag{C.13}$$

Following the same procedure as in the derivation of $\mathbf{A}_n$ for the EM algorithm, we can easily show that

$$\mathbf{A}_n = \mathbf{I}_{ML_h} - \mathbf{Q}_n. \tag{C.14}$$

APPENDIX D

COMPLEXITY COMPARISON OF EM-TYPE ALGORITHMS AND STC

Consider solving (7.10) using Greville's method (pp. 223 of [73]). Here, we briefly introduce the algorithm for quick reference. Suppose we need to compute $\mathbf{A}^\dagger \mathbf{y}$, where $\mathbf{A}$ is a matrix of size $N \times ML$, $\mathbf{y}$ is a $N \times 1$ vector and $\dagger$ denotes pseudo-inverse. Let $\mathbf{A}_k$ denote the matrix formed by the first k columns of matrix $\mathbf{A}$, where $k = 2, \ldots, ML$. Partition $\mathbf{A}_k$ as $\mathbf{A}_k = [\mathbf{A}_{k-1} \ \mathbf{a}_k]$ and denote $\tilde{\mathbf{A}} = [\mathbf{A} \ \mathbf{y}]$; Greville's iterative algorithm is then [73]

$$
\mathbf{A}_k^\dagger \tilde{\mathbf{A}} = \begin{pmatrix} \mathbf{A}_{k-1}^\dagger \tilde{\mathbf{A}} - \mathbf{d}_k \mathbf{b}_k^H \tilde{\mathbf{A}} \\ \mathbf{b}_k^H \tilde{\mathbf{A}} \end{pmatrix},
\tag{D.1}
$$

where $\mathbf{d}_k$ is the kth column of $\mathbf{A}_{k-1}^\dagger \tilde{\mathbf{A}}$. Let $\mathbf{c}_k = \mathbf{a}_k - \mathbf{A}_{k-1} \mathbf{d}_k$. If $\mathbf{c}_k = \mathbf{0}$, then

$$
\mathbf{b}_k^H \tilde{\mathbf{A}} = (1 + \mathbf{d}_k^H \mathbf{d}_k)^{-1} \mathbf{d}_k^H \mathbf{A}_{k-1}^\dagger \tilde{\mathbf{A}}.
\tag{D.2}
$$

If $\mathbf{c}_k \neq \mathbf{0}$,

$$
\mathbf{b}_k^H \tilde{\mathbf{A}} = (\mathbf{c}_k^H \mathbf{c}_k)^{-1} \mathbf{c}_k^H \tilde{\mathbf{A}}.
\tag{D.3}
$$

Ignore the number of multiplications that are of order $O(ML)$ or $O(N)$. At the kth iteration, assuming $\mathbf{c}_k \neq 0$, computing $\mathbf{c}_k$, $\mathbf{b}_k^H \tilde{\mathbf{A}}$ and $\mathbf{d}_k(\mathbf{b}_k^H \tilde{\mathbf{A}})$ takes $(k-1)N$, $N(ML+1)$ and $(k-1)(ML+1)$ multiplications, respectively. The total number of multiplications, V, can be computed easily as:

$$
V = \sum_{k=2}^{ML} [(k-1)N + N(ML+1) + (k-1)(ML+1)] \approx 1.5(ML)^2 N + 0.5(ML)^3.
\tag{D.4}
$$

If $\mathbf{c}_k = 0$, we have $V = 2(ML)^2 N + 0.5(ML)^3$. Therefore, Greville's algorithm does not reduce the complexity order. Here is an example of rough comparison of STC and EM-type algorithms with parameter settings similar to those used in the simulation. For $M = 2$ and $N = 128$, 13-tap-STC and 7-tap-STC require 138580 and 39004 multiplications respectively according to (D.3). In the worst case, a channel profile may require all 17-taps to be used for an accurate channel estimation; then, the required number of multiplications is 241604. If one directly computes $\mathbf{Q}_n$, $\mathbf{P}_n$ and $\mathbf{Q}_n^{-1}$, the required approximate number of multiplications for a 7-tap-STC, 13 tap-STC, and 17-tap STC are 4088, 18920 and 40648, respectively. Here we simply assume the coefficient associated with $O((ML)^3)$ is 1, thus the number of multiplications is computed as $(M + 0.5 * M(M - 1)) * 0.5 * Nlog_2 N + (ML)^3$. Note that the actual coefficient might be 2 or 2.5 as suggested by setting $N = ML$ in the equations used to compute V. In contrast, EM-type algorithms require approximately $MN_{\mathrm{iter}}N\log_2(N)$ multiplications, which is 5376, 10752 and 17920 for N_{iter} equal to 3,6 and 10, respectively. Since $N_{\mathrm{iter}} = 3$ is enough for SAGE and 10 is enough for EM in this case, SAGE is much less complex than a 13-tap-STC, while the EM algorithm is of comparable complexity as a 13-tap-STC. These comparisons might become even more favorable for SAGE as the number of transmit antennas increases.

VITA

I, Yongzhe Xie, was born in Shaoyang, Hunan Province, China in 1974, and raised in Wuhan, Hubei Province. My parents, both of whom are professors in Huazhong University of Science and Technology (HUST) in Wuhan, have been constantly encouraging me to strive for the best in my schooling. I studied four years at Shanghai Jiaotong University (SJTU) for my B.S. degree majoring in biomedical engineering and minoring in international finance. After graduation from SJTU in 1996, I studied for my master's degree in electronics and information engineering at HUST for a year before I transferred to The National University of Singapore, where I obtained my M. Eng. in electrical engineering. For me, the greatest achievement in Singapore was perhaps not the master's degree, but meeting Lan, now my wife. I came to College Station, Texas, in the Fall of 1999 to study for the Ph.D in electrical engineering with a research focus on wireless communications.

I can be contacted through my dissertation advisor, Prof. Costas N. Georghiades, whose mailing address is: Department of Electrical Engineering, Texas A&M University, College Station, TX 77843-3128.

The typist for this dissertation was Yongzhe Xie.