ON SPACE-TIME TRELLIS CODES OVER RAPID FADING CHANNELS WITH CHANNEL ESTIMATION

LI YAN

(M.Eng, Chinese Academy of Sciences)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2006

To my family

Acknowledgement

I would like to express my sincere gratitude to my supervisor Professor Pooi Yuen Kam for his valuable guidance and constant encouragement throughout the entire duration of my Ph.D course. It is he who introduced me into the exciting research world of wireless communications. His enthusiasm, critical thinking, and prudential attitude will affect me forever.

I specially thank Prof. Meixia Tao for her stimulating discussions and useful comments on parts of the work I have done. I am also grateful to Prof. Tjhung Tjeng Thiang, Prof. Chun Sum Ng, Prof. Nallanathan Arumugam and Prof. Yan Xin for their clear teaching on wireless communications, which help me broad the knowledge in this area.

I am grateful to my former and current colleagues in the Communications Laboratory at the Department of Electrical and Computer Engineering for their friendship, help and cheerfulness. Particular thanks go to Thianping Soh, Cheng Shan, Huai Tan, Zhan Yu, Rong Li, Jun He, Yonglan Zhu and Wei Cao.

I greatly appreciate my husband Zongsen Hu, who has always been with me and has given me a lot of support. He and our coming baby are the source of my happy life. They motivate me to chase my dream.

Finally, I would like to acknowledge my parents, who always encourage and support me to achieve my goals.

Contents

Summary

Space-time trellis codes (STTC) provide a promising technique to offer high data rates and reliable transmissions in wireless communications. Although most researches on STTC assume that perfect channel state information (CSI) is available at the receiver, this assumption is difficult and maybe impossible to realize in practice due to the time-varying characteristic of wireless channels. In this thesis, we examine the receiver structure and performance of linear STTC over rapid, nonselective, Rayleigh fading channels with channel estimation. Based on the performance analysis results obtained, code design and transmission schemes of STTC are investigated.

The time-varying MIMO channels are estimated by a pilot-symbol-assisted-modulation (PSAM) scheme. To achieve channel estimation of satisfactory accuracy with reasonable complexity, a systematic procedure is proposed to determine the optimal values of the design parameters used in PSAM, namely, the pilot spacing and the Wiener filter length. Based on the channel estimates obtained, the maximum likelihood (ML) receiver structure with imperfect channel estimation is derived for both independent, identically distributed (i.i.d.) and independent, non-identically distributed (i.n.i.d.) fading channels. Our results show that for the i.n.i.d. case, the channel estimation accuracy plays an important role in determining the weight on the signals received at each receive antenna. New results for the pair-wise error probability and the bit error probability are derived for the ML receiver obtained. The explicit results show clearly that the effects of channel estimation on the performance of STTC depend on the variances of the channel

estimates and those of the estimation errors. Using the performance analysis results obtained, we can optimally distribute the given average energy per symbol between the data symbols and the pilot symbols. By using the optimal pilot power allocation, performance can be improved without additional cost of power and bandwidth.

Based on the performance results obtained, a new code design criterion is proposed. This criterion gives a guide to STTC design with imperfect CSI over rapid fading channels. The key feature of our proposed criterion is the incorporation of the statistical information of the channel estimates. Therefore, the codes designed using this criterion are more robust to channel estimation errors for both i.i.d. and i.n.i.d. channels. For the i.n.i.d. case, due to the inherent unequal distributions among channels, it is more important to use our new design criterion by exploiting the statistical information of the channel estimates. To reduce the complexity of code search, an iterative code search algorithm is proposed. New STTC are designed which can work better than existing codes even when there exist channel estimation errors.

Finally, we study the closed-loop system, where it is assumed that only imperfect channel estimates are known to the receiver, and either complete or partial knowledge of this imperfect CSI is conveyed to the transmitter as the side information. A new lower bound on the capacity with imperfect CSI at both the transmitter and receiver is derived. Several optimal transmit power allocation schemes based on the side information at the transmitter are proposed.

List of Tables

List of Figures

List of Abbreviations

2G	second generation
3G	third generation
4G	fourth generation
AWGN	additive white Gaussian noise
BEP	bit error probability
BLAST	Bell-lab layered space-time architecture
CSI	channel state information
CDMA	code division multiple access
FDD	frequency division duplexing
GSM	Global System for Mobile system
ICSI	imperfect channel state information
i.i.d.	independent identically distributed
i.n.i.d.	independent non-identically distributed
IIR	infinite impulse response
LST	layered space-time
MAP	maximum a posteriori probability
MIMO	multiple-input multiple-output
ML	maximum likelihood
MPSK	M-ary phase shift keying
MRC	maximal ratio combing
MSE	mean square error
MMSE	minimum mean square error
PAG	power allocation gain
PCSI	perfect channel state information

PDF	probability density function
PEP	pair-wise error probability
PSAM	pilot-symbol-assisted-modulation
PSD	power spectrum density
SC	selection combining
SISO	single-input single-output
SNR	signal-to-noise ratio
STBC	space-time block codes
STTC	space-time trellis codes
TDM	time division multiplexing
TDMA	time division multiple access
TPA	transmit power allocation
TDD	time division duplexing
ZF	zero-forcing

Notations

Throughout this thesis, scalars are denoted by lowercase letters (a), vectors by boldface lowercase letters ($\mathbf{a}$), and matrices by boldface uppercase letters ($\mathbf{A}$).

- $\sqrt{-1} = j$.
- $P(X)$ is the probability of the event X.
- $P(X|Y)$ is the conditional probability of the event X given that the event Y has occurred.
- $p_X(x)$ is the probability density function of the random variable X.
- $\mathrm{E}[X]$ is the expectation of the random variable X.
- a^* is the conjugate of the complex scalar a.
- $\mathbf{A}^T$ is the transpose of $\mathbf{A}$.
- $\mathbf{A}^H$ is the complex conjugate transpose of $\mathbf{A}$.
- $\mathbf{0}_{T\times N}$ is the $T \times N$ zero matrix.
- $\mathbf{I}_M$ is the $M \times M$ identity matrix.
- $\|\mathbf{A}\|^2$ is the squared Euclidean norm of the $M \times N$ matrix $\mathbf{A}$ with the (m, n)th entry $[A]_{m,n} = a_{m,n}$.
- $\mathrm{rank}(\mathbf{A})$ is the rank of the matrix $\mathbf{A}$.
- $|\mathbf{A}|$ is the determinant of the matrix $\mathbf{A}$.
- $\mathrm{tr}(\mathbf{A})$ is the trace of the matrix A.
- $\mathrm{diag}[a_1, a_2, \cdots, a_M]$ is an $M \times M$ diagonal matrix with diagonal elements $a_1, a_2, \cdots, a_M$.

Chapter 1

Introduction

The history of wireless communication can be traced back to the 1890s. In 1897, Guglielmo Marconi first demonstrated the ability to communicate remotely with radio. Since then an exciting era of wireless communications has been unveiled. With the convenience of mobile communications and ease of deployment without wire, wireless communication has enjoyed rapid growth since the 1990s', and it now pervades our daily life.

1.1 Evolution of Wireless Communication

Due to limitations in analogue techniques of first generation (1G) wireless systems, second generation (2G) systems have employed digital modulation and signal processing techniques in transmission. Most of today's cellular networks are based on 2G techniques. Envisioning providing multimedia communications, third generation (3G) wireless systems are under construction, whose data rates are up to 2 megabits per second. However, the explosive growth of the Internet creates increasing demand for broadband wireless access. The data rates are set to exceed 100 megabits per second, which cannot be achieved by the current systems. Therefore, the aim of the next generation systems, the fourth generation (4G), is to provide

high transmission rate and highly reliable wireless communication. Reliable transmission with high peak data rates is expected to be 100 megabits per second to 1 gigabits per second, or higher for this 4G and systems beyond 4G. Therefore, a single wireless network that integrates both computing and communication systems can be used to provide ubiquitous services. This goal poses a tremendous challenge to design systems that are both power and bandwidth efficient, and manageable in complexity.

In wireless communication, the fundamental difficulty is the fading caused by multipath propagation which severely impacts system performance. However, the effects of fading can be substantially mitigated by using diversity techniques. Three main forms of diversity are exploited for fading channels: temporal, spectral and spatial diversity. Recently, it was found that the *space* domain can be exploited to significantly increase channel capacity, i.e., using multiple-input multiple-output (MIMO) systems, without increasing spectral and power consumption. MIMO systems are those that have multiple antenna elements at both the transmitter and receiver. In fact, antenna diversity at the receiver has long been widely used in wireless communication to combat the effects of fading. Although antenna diversity at the receiver has been studied for more than 50 years, research on transmit diversity is much more recent. Pioneering works by Winters [4], Foschini [5], and Telatar [6] show remarkable spectral efficiencies for wireless systems with multiple antennas. Under the rich scattering environments with independent transmission paths, the capacity of a MIMO system with N_T transmit and N_R receive antennas is linearly proportional to $\min(N_R, N_T)$. Thus, the capacity is increased by a factor of $\min(N_R, N_T)$ compared to a system with just one transmit and one receive antenna. The advantages of multiple antennas is due to two effects. One is diversity gain since it reduces the chances that several antennas are in a deep fade simultaneously. The other is the beamforming gain obtained by combining the signals from different antennas to achieve a higher signal-to-noise ratio (SNR). Since multiple antennas introduce a new dimension of space on top of the conventional time dimension at the transmitter, this triggers tremendous research interests on multi-dimensional

coding procedures for MIMO systems, which are generally referred to as space-time coding schemes. More detailed literature reviews on space-time coding schemes will be given in the next section.

1.2 Space-time Coding Schemes

Tarokh et al. [1] first introduced the concept of space-time coding by designing codes over both time and space dimensions. Their original work gave the well known rank-determinant and product distance code design criteria of space-time codes for quasi-static fading and rapid fading channels, respectively. For the quasi-static fading case, the fading coefficients remain constant over an entire transmission frame, while, for the rapid fading case, the coefficients vary independently from symbol to symbol. Following Tarokh's work, much research efforts have been made to develop powerful space-time codes based on different design criteria or improved search algorithms [2], [7], [8], [9], [10], [11], [12] and [13]. The family of space-time codes includes space-time trellis codes (STTC) [2], [7], [8], [9], [10], and space-time block codes (STBC) [11], [12] and [13]. The beauty of STBC is its simplicity, which can achieve the maximum diversity with a simple decoding algorithm. However, no coding gain can be provided by STBC, and non-full rate STBC reduce bandwidth efficiency. In this thesis, we will concentrate on effective space-time trellis codes (STTC), which is a joint design of coding, modulation and diversity.

So far, many papers in the literature on the design criterion of STTC considered quasi-static fading channels. To design codes with optimal performance, we first need certain performance measures. One of the most important performance measures is the error probability. Tarokh et al. proposed the well known STTC scheme in [1] by minimizing the worst pair-wise error probability (PEP). Based on their derived PEP upperbound, their code design criterion relies on the minimum determinant of codeword difference matrices. This criterion is mainly for high SNR.

Alternatively, the Euclidean distance criterion was presented by Yuan and Vucetic in [7], which indicates that when the diversity gain is reasonably large, the trace of the codeword distance matrix, or, equivalently, the minimum square Euclidean distance, will dominate the code performance. It was also found in Tao [8] that the Euclidean distance criterion should be used for moderate and low SNR. Based on these popular design criteria, several powerful STTC are obtained using computer search techniques in [2], [7], [8], and [10]. To simplify code search complexity, some systematic code design algorithms were proposed. Using delay diversity, [14] converted the two-dimensional design problems to the traditional one-dimensional problem, and greatly reduced the code search complexity. Also, a systematic search algorithm was provided in [15] to design codes with full diversity gain. Diversity gain can characterize the error probability performance at high SNR. Using diversity gain as a performance measure is more convenient, but the price is that coding gain cannot be guaranteed. Instead of using error probability as the performance measure, a novel scheme was proposed by [5] aiming to achieve the outage capacity with reasonable complexity, and this is the so-called layered space-time (LST) architecture that can attain a tight lower bound on the MIMO channel capacity. There are a number of LST architectures, depending on whether error control coding is used or not, and on the way the modulated symbols are assigned to transmit antennas. An uncoded LST structure, known as vertical Bell Laboratories layered space-time (VBLAST) scheme is first proposed in [16]. In this scheme, the input information sequence is demultiplexed into N_T sub-streams and each of them is modulated and transmitted from a transmit antenna. The receiver in [16] is based on a combination of interference suppression and cancelation. The interferences are suppressed by a zero-forcing (ZF) approach. Following this, more researches [17], [18], [19] exploited the combination of layered space-time coding and signal processing. By using a spatial interleaver, a better performance can be achieved. With the spatial interleaver, the modulated codeword of each layer is distributed among the N_T antennas, which introduces space diversity. Note that these LST systems require a quasi-static channel as the iterative cancelation process requires

a precise knowledge on the channel coefficients.

Compared with the case of the quasi-static channels, the works on the design criteria for the rapid fading channels do not follow so much the approach of [1]. There remains much room to improve on the design criteria for the case of rapid fading. Recently, an improved code design criterion by minimizing the node error probability was presented in [9] for the rapid fading case with perfect channel state information (CSI). With the development of more performance analysis results for STTC, it has been shown that the distance spectrum need to be considered to fully characterize STTC performance [20], [21]. Although new, improved STTC for the rapid fading channels are few, the performance analysis for this case has attracted lots of research interests. Some exact PEP results are provided in [22], [23]. These exact PEP expressions are not explicit, and rely either on numerical integration or residue computation. Several tighter PEP bounds than those in [1] are also provided in [24]. In addition to the PEP analysis, [23], [25] and [26] examined the BEP bounds. Note that all these papers assume perfect CSI at the receiver, and the results are not explicit. Thus, they provide little insights into how to improve the design criterion for STTC over rapid fading channels.

The space-time coding schemes mentioned above all are open-loop systems. For open-loop systems, there is no CSI available at the transmitter. However, if CSI is available, it should be utilized to improve performance. Therefore, closed-loop MIMO systems have recently attracted great research interests. In closed-loop systems, CSI at the receiver can be conveyed to the transmitter by using feedback. We call information that is known to the transmitter as side information. By incorporating side information, closed-loop systems have been shown to achieve improved performance [27], [28], [29]. With the available side information concerning the channels, the transmitter can employ strategies such as adaptive coding, and modulation schemes [30], [31], and transmit antenna selection [32]. Side information at the transmitter can also be exploited to take advantage of sophisticated signal processing techniques [33], [34]. It is well known that when perfect CSI is assumed

at the transmitter, beamforming can be used to maximize the received SNR. However, due to the limited bandwidth of the feedback channel, or the feedback errors and delays, perfect CSI at the transmitter is practically impossible. Recently, much research has been done on partial or imperfect side information scenarios [28], [35], [36]. It is shown in [28] that in the extreme of perfect feedback, the optimal strategy entails transmission in a single direction specified by the feedback, i.e., the beamforming strategy. Conversely, with no channel feedback, the optimum strategy is to transmit equal power in orthogonal independent directions, i.e., the diversity scheme. Between these two extremes, some appropriate transmitter strategies are provided when the side information at the transmitter is imperfect. In [28], both quantized and noisy side information are considered, and the optimal transmission strategy depends on the rank of its input correlation matrix given side information. In [35], two feedback schemes are proposed, namely, mean and variance feedback. For both schemes, the beamforming strategy appears to be a viable transmission strategy when meaningful channel feedback is present. However, [35] only considered the case of one receive antenna. More results are obtained for multiple antenna systems with space-time coding in [36]. All these papers [28], [35], [36] assume that perfect CSI is available to the receiver.

Throughout the development of space-time codes, most researches have focused on the idealistic assumption that perfect CSI is available to the receiver. However, in practical systems, perfect CSI may not be available due to channel estimation errors. This is especially true for the rapid fading case, where perfect CSI is generally unavailable. To overcome this problem, either noncoherent detection methods, where no CSI is needed at the receiver, or channel estimation techniques can be used. Noncoherent differential modulation schemes were developed in [37] and [38]. However, it is known that there is a performance loss with noncoherent detection. Furthermore, signal constellation design for differential modulation schemes is difficult. To achieve satisfactory performance with noncoherent differential modulation schemes, it is required that channels are constant for a sufficient long time duration. Therefore, in this thesis, we consider instead the use of channel estimation at

the receiver. In the limit of perfect channel estimation, performance can approach that of ideal coherent detection, which is optimal. Although channel estimation techniques are well understood for single-input single-output (SISO) systems [39], [40], channel estimation schemes for MIMO systems are different from those of SISO systems. To estimate MIMO channels is not a trivial problem because of the additional spatial dimension. In addition to the difficulties of MIMO channel estimation, the performance analysis and code design of STTC over MIMO systems with channel estimation errors are even more challenging. There have been a few works on STTC error performance analysis with imperfect CSI [41], [42], and [43]. However, all these works on STTC with imperfect CSI considered only the *quasi-static fading* case, where the fading coefficients are assumed to remain constant over an entire frame. For the quasi-static fading channels, Tarokh et al. [41] presented a PEP upperbound with channel estimation. Their result is a function of the correlation coefficients between channel fading coefficients and their estimates. Also, the result is only approximately correct at high SNR [42]. Therefore, the result is implicit and does not reveal explicitly the effects of channel estimation errors on code performance or code design. Garg et al. [43, eq. (39)] provided an analytical PEP expression for STTC with imperfect CSI. Their method requires the computation of residues of the characteristic function of a random variable, and the computation has to resort to some numerical softwares like MATLAB. The implicit expression obtained fails to give insights into the performance loss caused by channel estimation errors. This PEP result also makes its applications to code design cumbersome. To our knowledge, STTC performance analysis with imperfect CSI over *rapid* fading channels has not been considered so far. The performance analysis and code design of STTC over rapid fading channels are nearly untouched research areas in the literature, and these will be the key research topics in this thesis. Rapid fading channels are frequently encountered in many practical communication systems. The rapid fading may arise from complete interleaving/de-interleaving to achieve better performance. Over rapid fading, an additional form of diversity, namely, time diversity, can be exploited, and full di-

versity is equal to the product of the number of receive antenna and the minimum Hamming distance between codevectors.

1.3 Research Objectives and Main Contributions

In this thesis, we will examine linear STTC over rapid Rayleigh fading with imperfect CSI for both open-loop and closed-loop systems. The effects of channel estimation errors on the receiver structure, performance and channel code design of STTC systems are investigated. Throughout this thesis, we consider point-to-point communications with the common M-ary phase shift keying (MPSK) modulation scheme. The channels are modeled by frequency non-selective, rapid, Rayleigh fading processes. In most applications, rapid fading channels are desirable because the time diversity achieved can combat channel fading effectively. The usual way to produce the rapid fading scenario is by using interleaving/deinterleaving techniques. For illustration purpose, throughout this thesis, the rapid fading scenario is produced by perfect multiplexing/de-multiplexing of the time division multiplexing (TDM) system. This TDM technology can not only be implemented easily on the existing wireless networks, but also reduce the memory size and the transmission delay for each user. The idea behind the TDM system is that each user can experience independent channel fading over time by perfect interleaving/de-interleaving through the multiplexing/de-multiplexing with a sufficiently large number of users.

This rapid fading channel model is important from both practical and theoretical viewpoints. The widely deployed wireless network in Europe, Asia, etc. is the Global System for Mobile (GSM) system, which is based on time division multiple access (TDMA) techniques. Thus, it provides a convenient, implementation platform to boost the data rate by applying MIMO techniques to TDMA systems. In TDMA systems, the data from all users are multiplexed into frames. In each frame, each user is assigned one time slot to transmit data. Then, he must wait for a frame length to transmit again. Therefore, the data from each user are interleaved by one

frame length. When the number of users is sufficiently large, the data from each user are completely interleaved, and the fading coefficients experienced by each user are independent from one symbol to the next. By accommodating a sufficiently large number of users in the practical TDMA system, it consequently produces the rapid fading channel model. This brings out the time diversity advantage, which has been widely exploited in SISO systems. With space and time diversity, the increased capacity promised by MIMO systems [6] can be achieved. The other extensively examined fading channel model, namely, the quasi-static fading channel model can be understood in a similar way in TDMA systems. Whether the rapid fading channel model or the quasi-static model should be used first depends on the relationship between the channel coherence time T_c and the symbol duration time T_s. The coherence time is a statistical measure of the time duration over which the channel impulse response is essentially invariant. If the coherence time of the channel is much greater than the symbol period of the transmitted signal, i.e., $T_c \gg T_s$, then, the channel changes at a rate much lower than that of the transmitted signal, and the channel fading coefficients can be assumed to be constant over several symbol intervals. Otherwise, if the coherence time is approximately equal to the symbol duration, i.e., $T_c \approx T_s$, the channel fading coefficients are only symbol-wise constant. For the quasi-static fading model, the channel fading coefficients remain constant over the transmission of a block of data. Therefore, this model is only suitable for the very low mobility applications, where the coherence time of channels is sufficiently larger than the symbol duration, and the channels are assumed to be block-wise constant. On the other hand, the symbol duration is also dependent on the channel coherence bandwidth. For flat fading channels, the reciprocal of the symbol duration needs to be much less than the channel coherence bandwidth in order to avoid intersymbol interference. Thus, the channel coherence time, the channel coherence bandwidth and the symbol duration all play a role in determining the channel fading model to be used. Compared with the quasi-static fading model, the rapid fading model is more realistic since it only assumes the fading coefficients to be symbol-wise constant, which is realizable for most practi-

cal systems. Besides these practical considerations, the rapid fading channel model also provides attractive properties for theoretic analysis. For this case, explicit performance results can be obtained, which can clearly show the effects of the channel estimation errors on MIMO systems. Therefore, we concentrate on the rapid fading case in this thesis.

The time-varying MIMO channels are estimated by a pilot-symbol-assisted-modulation (PSAM) scheme. For open-loop systems, i.e., there is no side information at the transmitter, both independent, identically distributed (i.i.d.) and independent, non-identically, distributed (i.n.i.d.) fading channels are considered. For the i.i.d. case, the maximum likelihood (ML) receiver structure with imperfect channel estimation is derived. Then, performance analysis for this receiver is analyzed. The explicit results show clearly the effects of channel estimation errors on the performance of STTC. With the performance results, a new code design criterion is proposed. This criterion gives a guide to STTC design with imperfect CSI over rapid fading channels. The key feature of our proposed criterion is the incorporation of the statistical information of the channel estimates. Therefore, the codes designed using this criterion are more robust to channel estimation errors. New STTC are designed which can work better than existing codes even when there exist channel estimation errors. This is very important for practical systems where channel estimation errors are common. The codes designed with perfect CSI assumption may not be optimal in actual channel estimation conditions. After the study of the i.i.d. case, we extend the work to the i.n.i.d. channel conditions by relaxing the constraint of identical statistical distribution on each link. This is motivated by the fact that the requirement of i.i.d. fading channels may have limitations in some applications where different paths have non-identical statistics, such as for wideband code division multiple access (CDMA) and indoor ultrawide bandwidth communications [44]. The channel multipath intensity profile of IMT-2000 channel models [45] and JTC channel models [46] is variable, i.e., the mean square fading gain of each diversity branch is different. The i.n.i.d. fading channel model has been examined in [47, 48], because it is realistic and general. For the

i.n.i.d. case, the ML receiver is different from that obtained for i.i.d. channels. The i.i.d. ML receiver is only a special case of the i.n.i.d. receiver. A method different from that used in the i.i.d. case is employed to analyze the performance of STTC, where both the exact PEP and several PEP bounds are obtained. Similarly, we also examine the effect of non-identical distributions of the MIMO fading channels on the code design. The effects of both the non-identical distribution of channels and the imperfect channel estimation are all represented in the same way by the different variances of the channel estimates among the transmit antennas. With i.n.i.d. fading channels, it is even more important to exploit the statistical information of the channel estimates since the imbalance among the different transmit-receive antenna pair is greater.

After examining the open-loop systems, we next consider the closed-loop systems by assuming that side information of the imperfect CSI is available at the transmitter. We study the optimal power allocation schemes for the closed-loop STTC system with imperfect CSI. The criterion is either the channel capacity or the error probability. It is worth mentioning that the channel capacity with imperfect CSI cannot simply be obtained by replacing the channel fading matrix by its estimate. The capacity with channel estimation is also affected by the variances of the channel estimation errors. For bandwidth-limited feedback systems, the CSI available to the transmitter is partial. Based on the side information at the transmitter, different power allocation schemes are presented. In addition to studying the power allocation for the transmit power, the power allocation for pilot symbols is also examined to achieve the optimum error performance with the constraint of a fixed total transmission power. The details of the contribution and the main results obtained are listed in the following paragraphs.

First, we propose a simple PSAM scheme to estimate MIMO channels. The two design parameters of this scheme are the pilot spacing and Wiener filter length, respectively. A smaller pilot spacing and a larger Wiener filter length offer better estimation accuracy with other factors fixed. However, high transmission

rate/bandwidth efficiency dictates a large pilot spacing, while a reduced receiver complexity dictates a smaller Wiener filter length. Therefore, suitable compromise values for the pilot spacing and Wiener filter length should be chosen. We propose a systematic approach to determine suitable values of these two design parameters. It will be shown that suitable values should be chosen in a tradeoff between estimation accuracy, transmission rate/pilot overhead, and receiver complexity. With the PSAM scheme and the minimum mean square error (MMSE) estimator, the statistical information of the channel estimates is derived. By incorporating the pilot channel measurements obtained from the PSAM scheme, the ML receiver with imperfect CSI is derived, assuming that no CSI is available to the transmitter. The fading channels are assumed to be i.i.d.. The ML receiver has a simple form for MPSK modulation. A new, explicit PEP upperbound on the ML receiver is derived. Our PEP result clearly presents the effects of channel estimation errors on the error performance. The channel estimation errors introduce additional noise at the receiver as shown by *an effective noise* term. In other words, the channel estimation errors increase the noise power at each receive antenna. Furthermore, the error performance results show that the Euclidean distances between code symbols are *weighted* by the different variances of the estimated channel coefficients associated with the different transmit antennas. Thus, the structure and performance of codes are expected to be affected by the quality of the channel estimates. Based on the PEP upperbound obtained, a tight upperbound on the bit error probability (BEP) is presented using the dominant error events approach. The maximum length of the dominant error events considered increases with the number of states of the codes. It will be shown that the performance loss caused by channel estimation errors increases with the channel fade rate. For high fade rates, the transmit diversity gain of a large number of transmit antennas can be significantly offset by the increased estimation error variance.

The performance results show that STTC performance with imperfect CSI is affected by the variances of the channel estimates. This motivate us to provide an effective design criterion for STTC over rapid fading channels with imperfect CSI.

This criterion exploits the statistical information of the channel estimates in the code design, and can reduce the performance loss caused by the channel estimation errors. The statistical information of the channel estimates depends on the channel parameters, such as the variances and the fade rates of the fading processes. Here, we first concentrate on the i.i.d. fading case. For this case, the key parameter which affects the statistical information of the channel estimates is the channel fade rate. With different channel fade rates, the most suitable values for the design parameters of the estimator should be modified, and the corresponding statistical information of the channel estimates is different. Based on the knowledge of the fade rate at the transmitter, two situations for code design are considered. For fading channels with invariant fade rates, the channel fade rate can be treated as a priori-known knowledge at the transmitter. When the channel fade rate is known to the transmitter, codes whose design parameters are matched to this channel fade rate perform best. It will be shown that adaptive code design based on knowledge of the channel fade rate should be used to achieve the optimal performance. On the other hand, when the channel fade rate is unknown to the transmitter, robust codes are proposed based on the distribution of the channel fade rates. The codes are robust in a statistical sense. This means that the robust codes have the best *average* performance over all the possible channel fade rates. The average performance can be obtained by averaging the conditional performance at each fade rate over the probability distribution of channel fade rates. To achieve the optimum average performance, the robust code can be designed based on *the average variance vector* of the channel estimates. In general, robust codes are designed based on the average variance vector. However, this can be simplified, if the channel has a *dominant* fade rate, which occurs with much higher probability than the sum of the probabilities of occurrence of other fade rates. With the proposed design criteria of STTC with imperfect CSI, new codes will be designed using an search algorithm with reduced implementational complexity. As is known, one challenging problem of space-time code design is the high search complexity over a very large set of possible generator matrices, which is caused by the cross-dimensional design of space-time

codes over both space and time. Thus, we provide a new, iterative code search algorithm to greatly reduce the high code search complexity. The iterative code search algorithm with reduced complexity provides a systematic way to design codes with good performance. In the presence of channel estimation errors, the codes designed based on our new criterion assuming imperfect CSI have improved bit error probability performance compared to existing codes that were designed under the perfect CSI assumption. It is shown that the effect of channel estimation on code design increases with the channel fade rate and the number of transmit antennas. Simulation results also verify the advantages of our proposed new codes under actual channel estimation conditions.

Extending from the case of the i.i.d. fading channels, we next relax the identical distribution constraint and consider the i.n.i.d. fading channels. Both unequal variances and unequal fade rates on the different links are assumed. For the i.n.i.d. channels, the corresponding ML receiver is derived. Due to the i.n.i.d. fading channels, the receiver requires in its signal detection function, both the channel estimates and the second order statistical information of the estimates, which can be obtained from the channel estimator. The variances of the channel estimation errors increase the total noise power at each receive antenna, and lead to the effective noise being different from one receive antenna to another. Therefore, the ML receiver for the i.n.i.d. case with imperfect CSI cannot be obtained from the perfect CSI ML receiver by replacing the known channel matrix with the imperfectly estimated channel matrix. In fact, channel estimation accuracy plays an important role in determining the weight on the signals received at each receive antenna. The exact PEP result for the ML receiver is obtained using the moment generating function. Based on the exact PEP expression, several PEP bounds, which are explicit and simple to compute, are derived. The union bounds on the BEP of STTC are also derived by using both the transfer function approach and the method of dominant error events. All these results are extensions of those obtained for the i.i.d. case. Following the performance analysis, code design for i.n.i.d. channels with imperfect CSI is studied. First, we examine the effects of the differences among the

channel fade rates and the differences among the variances on code design. Both the effects of the different channel fade rates and variances can be all reflected in the same way by the statistical information of the channel estimates. Similar to the i.i.d. case, the effect of channel estimation on code design is measured by the maximum variance difference of the channel estimates. Our results show that with satisfactory channel estimation accuracy, the effects of the different variances is more important, compared with the differences among the channel fade rates. Employing the node error event as the cost function, a practical code design criterion is presented, and new STTC are obtained using the iterative search algorithm proposed early in the i.i.d. case. Due to the inherent non-identical distributions among the fading channels, it is more important to use our new design criterion by exploiting the statistical information of the channel estimates. Under non-identical channel conditions, our proposed codes perform better than the existing codes in the literature which are designed on the assumption of i.i.d. channels, and perfect CSI at the receiver. When the variance differences among channel fading processes increase, the performance gains achieved by our proposed STTC are greater. It is also shown that optimal codes matched to the channel and estimator parameters should be used when these parameters are known at the transmitter.

Finally, we turn our attention to closed-loop systems, where it is assumed that only imperfect channel estimates are known to the receiver, and either complete or partial knowledge of this imperfect CSI is conveyed to the transmitter as the side information. With the partial information at the transmitter, the optimal power allocation schemes are investigated. From information theory, the known water-filling scheme can be used to allocate the optimal power for each transmit antenna when perfect CSI is available at the transmitter. For the case of imperfect CSI at both the transmitter and receiver, the optimal power allocation scheme has not been examined. Thus, we derive MIMO capacity bounds with both imperfect CSI at both the transmitter and receiver. Unlike the commonly used capacity expression with imperfect CSI at the transmitter, where only the channel fading matrix is replaced by its estimate, it is worth noting that the power of the noise at

the receiver is also increased by the estimation errors. Since the variances of the estimation errors depends on the transmit power, the maximization of the capacity is now different from that of the perfect CSI case. The optimization problem now changes to a fractional programming problem, and the conventional water-filling scheme cannot apply here. Based on the capacity lower bound, we propose a new power allocation scheme using a general nonlinear programming method. Besides using the capacity lower bound as the objective function, we also employ the upper bound on the PEP as the objective function to design the optimal transmit weighting matrix. Therefore, two additional power allocation schemes are proposed based on the knowledge of estimated CSI at the transmitter. In the first case, we assume that the knowledge of the estimated channel fading matrix is available at the transmitter. Thus, the conditional PEP given the pilot channel measurements is used as the objective function. With imperfect CSI at the receiver, both the estimated channel fading matrix, and a matrix, which depends on the variances of the channel estimation errors and the average SNR, should be used *jointly* to minimize the error performance. When the channel estimates are sufficiently reliable, the transmitter should transmit signals along the direction of the eigenvector corresponding to the largest eigenvalue of a matrix, where the estimated channel fading matrix is modified by its corresponding estimation accuracy. This power allocation scheme has lower complexity than that based on the capacity lowerbound. However, it also requires the knowledge of the estimated channel matrix be known at the transmitter. To be more practical, the second case considers the bandwidth-limited feedback systems, where only partial CSI is available at the transmitter. Then, the above two schemes are not suitable, since the estimated channel fading matrix is unavailable at the transmitter due to the limited feedback bandwidth. For this case, we assume that only the variances of the channel estimates are known at the transmitter. Then, the average PEP is used here as the metric. Compared with the optimal weighting vector obtained for the first case, the correlation matrix of the estimated channel fading matrix is replaced by the matrix consisting of the variances of the channel estimates. It is intuitively clear that the performance gain

achieved by the optimal weighting vector of the second case is worse than that of the first case, due to the loss of some information. However, it will be shown by our numerical results that there is a tradeoff between the performance gain and the complexity of feedback.

1.4 Organization of the Thesis

The organization of the thesis is given as follows.

In Chapter 2, the MIMO communication systems and some basic ideas of channel estimation are introduced.

In Chapter 3, performance analysis of STTC over i.i.d. rapid, Rayleigh, fading channels with channel estimation is presented. The ML receiver is derived by incorporating the pilot channel measurements. Tight PEP and BEP bounds are obtained. In Chapter 4, we discuss the code design of STTC with channel estimation. The encoder structure is given. An improved code design criterion is proposed, and an iterative code search algorithm is introduced. When the channel fade rate is known at the transmitter, the imperfect CSI codes for that fade rate can be used. On the other hand, when the chanel fade rate is time-variant, the robust codes are designed based on the probability model of the channel fade rates.

After the study on the performance analysis and code design of STTC for i.i.d. fading channels, the whole work is extended to the i.n.i.d. fading channels in Chapter 5. More general results are obtained. In Chapter 6, we focus on the power allocation scheme for closed-loop STTC systems with imperfect CSI at the both transmitter and receiver. Given the side information at the transmitter, several power allocation schemes are proposed.

Finally, we provide concluding remarks and suggestions for future research in Chapter 7.

Chapter 2

MIMO Communication Systems with Channel Estimation

In this chapter, we first introduce the MIMO communication systems. The capacity of MIMO channels with perfect CSI known at the receiver is the maximum achievable data rate for an arbitrarily low probability of error, if the signal may be encoded by an arbitrarily long space-time code. Then, the MIMO channel model used in this thesis is described. The simulator to generate the fading channels is presented. Finally, mean-square channel estimation is presented.

2.1 MIMO Communication Systems

We consider a point-to-point MIMO system with N_T transmit and N_R receive antennas. The system block diagram is shown in Fig. 2.1. The input data stream is encoded by the space-time encoder. The encoded data is split into N_T streams. Each stream is pulse-shaped and modulated. At each time slot t, a $N_T \times 1$ signal vector $\mathbf{s}(t) = [s_1(t) \ \cdots \ s_{N_T}(t)]^T$ is transmitted simultaneously, where $s_j(t)$ is transmitted by the jth antenna. The total transmitted power is constrained to P,

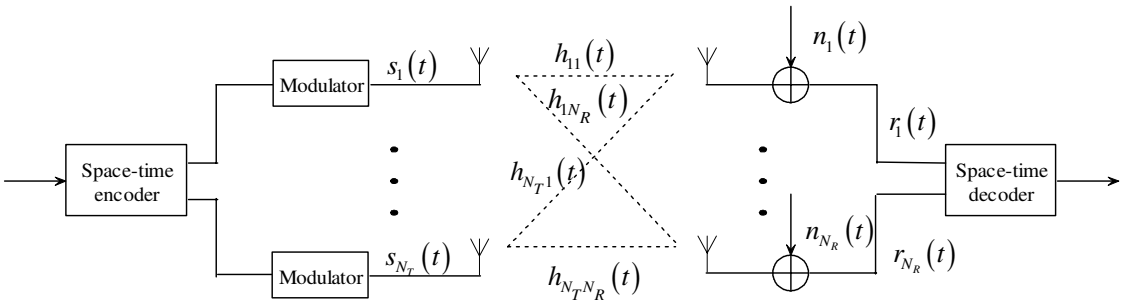


Fig. 2.1: Block diagram of a MIMO system.

which can be represented as

$$P = tr(\mathbf{R_{ss}}) \tag{2.1}$$

where $tr(\cdot)$ denotes the trace of a matrix, and $\mathbf{R_{ss}}$ is the covariance matrix of the transmitted signal. It is given by

$$\mathbf{R_{ss}} = \mathrm{E}[\mathbf{s}(t)\mathbf{s}^H(t)] \tag{2.2}$$

where $\mathrm{E}[\cdot]$ denotes the expectation and the superscript ^{H} denotes the Hermitian transpose of a vector/matrix. The signal at each receive antenna is a noisy superposition of the N_T transmitted signals corrupted by the channel fading. The $N_R \times 1$ received signal vector $\mathbf{r}(t) = [r_1(t) \ \cdots \ r_{N_R}(t)]^T$, where the ith element $r_i(t)$ refers to the signal received at antenna i, is given by

$$\mathbf{r}(t) = \mathbf{H}(t)\mathbf{s}(t) + \mathbf{n}(t) \tag{2.3}$$

where $\mathbf{n}(t)$ is the vector of additive, channel white noises at the receiver. Its components are i.i.d. complex, zero-mean, Gaussian random variables. The covariance matrix of $\mathbf{n}(t)$ is given by $\mathbf{R_{nn}} = \mathrm{E}[\mathbf{n}(t)\mathbf{n}^H(t)] = N_0\mathbf{I}_{N_R}$, where $\mathbf{I}_{N_R}$ is the $N_R \times N_R$ identity matrix. The receive branches have identical noise power of N_0. The symbol-wise constant channel fading matrix is described by an $N_R \times N_T$ complex matrix $\mathbf{H}(t) = [h_{ij}(t)]$. The component $h_{ij}(t)$ is the channel fading coefficient on the (i,j)th link, i.e., from the jth transmit to the ith receive antenna. The elements of $\mathbf{H}(t)$ are modeled by independent random variables. The channel fading processes are assumed to be independent of the additive noises.

When the channel matrix is unknown at the transmitter, the optimum distribution of transmitted signals $\mathbf{s}(t)$ achieving the channel capacity is Gaussian, and the elements of $\mathbf{s}(t)$ are i.i.d. Gaussian variables. Thus, the signals transmitted from each antenna have equal powers of $E_s = P/N_T$. We normalize the average energy of the constellation, contracting the elements in the constellation by a factor of $\sqrt{E_s}$. Then the received signals, $\mathbf{r}(t)$, can be expressed as

$$\mathbf{r}(t) = \sqrt{E_s}\mathbf{H}(t)\mathbf{v}(t) + \mathbf{n}(t) \tag{2.4}$$

where $\mathbf{s}(t) = \sqrt{E_s}\mathbf{v}(t)$. The elements in $\mathbf{v}(t)$ have unit energy.

Based on the theoretical work developed by Foschini [5] and Telatar [6], it is known that the capacity C for rapid fading MIMO channels can be obtained by

$$C = \mathrm{E}\left[W \log_2 \det(\mathbf{I}_r + \frac{E_s}{N_0}\mathbf{Q})\right] \tag{2.5}$$

where W is the bandwidth of each channel, r is the rank of $\mathbf{H}(t)$, and $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \begin{cases} \mathbf{H}(t)\mathbf{H}^H(t) & N_R < N_T \\ \mathbf{H}^H(t)\mathbf{H}(t) & N_R \geq N_T \end{cases} \tag{2.6}$$

Suppose $N_T = N_R = n$ and perfect CSI is known at the receiver, the *asymptotic* capacity is given by [6]

$$\lim_{n\to\infty} \frac{C}{Wn} = \frac{1}{\pi}\int_0^4 \log_2(1 + \frac{P}{N_0}\nu)\sqrt{\frac{1}{\nu} - \frac{1}{4}}d\nu \,. \tag{2.7}$$

By observing that $\log(1 + x) \geq \log(x)$, (2.7) can be bounded as

$$\lim_{n\to\infty} \frac{C}{Wn} \geq \frac{1}{\pi}\int_0^4 \log_2(\frac{P}{N_0}\nu)\sqrt{\frac{1}{\nu} - \frac{1}{4}}d\nu = \log_2(\frac{P}{N_0}) - 1 \ . \tag{2.8}$$

The bound in (2.8) shows that the capacity increases linearly with the number of antennas, and logarithmically with the SNR.

2.2 The Radio Channel Model

In a wireless communication environment, the channel is the space between the transmit and the receive antennas. The presence of reflecting objects and

scatterers in this space creates a constantly changing environment that dissipates the signal in amplitude, phase, and time. These effects result in multiple versions of the transmitted signal that arrive at the receiving antenna, displaced with respect to one another in time and spatial orientation. Assuming that the number of multipaths is large enough, the fading gain can then be modeled as a complex, symmetric Gaussian random variable. If there is no dominant path, then the absolute value of the complex Gaussian gain follows the Rayleigh distribution.

The Rayleigh distribution is frequently used to model the statistics of signals transmitted through radio channels such as cellular radio. The probability density function (PDF) of a random variable R with a Rayleigh distribution is given by

$$p_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \ . \tag{2.9}$$

Thus, the Rayleigh-distributed random variable R can be formed as $R = \sqrt{X_1^2 + X_2^2}$, where X_1 and X_2 are two zero-mean, statistically independent, Gaussian random variables, each having a variance of σ^2. As a generalization, consider the random variable

$$R = \sqrt{\sum_{i=1}^{n} X_i^2} \tag{2.10}$$

where the X_i, $i = 1, \cdots, n$, are statistically i.i.d., zero-mean, Gaussian random variables with variance σ^2. The random variable R has a Gamma distribution, whose PDF is given by

$$p_R(r) = \frac{r^{n-1}}{2^{(n-2)/2}\sigma^n \Gamma(n/2)} e^{-r^2/2\sigma^2} \tag{2.11}$$

where $\Gamma(p)$ is the gamma function, defined as $\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt$, where $p > 0$.

The Jakes spectrum is commonly used to model the Rayleigh fading process for the land-mobile cellular channel [49]. The Jakes power spectrum density (PSD) is defined as

$$S(\omega) = \begin{cases} \frac{2\sigma^2}{w_d \sqrt{1-(w/w_d)^2}} & |w| < w_d \\ 0 & |w| > w_d \end{cases} \tag{2.12}$$

where w_d is the maximum radian Doppler frequency. The corresponding autocorrelation of the Jakes spectrum is given by

$$R(\tau) = \mathrm{E}[x(t)x^*(t+\tau)] = \sigma^2 J_0(w_d\tau) \tag{2.13}$$

where $J_0(\tau)$ is the Bessel function of the first kind of order zero.

We will use the sum-of-sinusoids statistical simulation models proposed in [50] to simulate Rayleigh fading channels with Jakes PSD. This simulator has improved properties than others in the literature, which introduces random path gain, random initial phase, and conditional random Doppler frequency for all individual sinusoids. Thus, the autocorrelation and crosscorrelation of the quadrature components, and the autocorrelation of the complex-envelope of this simulator match the desired ones exactly, even if the number of sinusoids is as small as a single-digit integer. The normalized *lowpass* fading process $h(t)$ of the statistical sum-of-sinusoids simulation model is defined by

$$h(t) = x_c(t) + jx_s(t) \tag{2.14a}$$

$$x_c(t) = \frac{2}{\sqrt{M}} \sum_{n=1}^{M} \cos(\psi_n) \cdot \cos(w_d t \cos\alpha_n + \phi) \tag{2.14b}$$

$$x_s(t) = \frac{2}{\sqrt{M}} \sum_{n=1}^{M} \sin(\psi_n) \cdot \cos(w_d t \cos\alpha_n + \phi) \tag{2.14c}$$

with

$$\alpha_n = \frac{(2n-1)\pi + \theta}{4M} \tag{2.15}$$

where ψ_n, ϕ, and θ are statistically independent and uniformly distributed over $[-\pi, \pi)$ for all n. It has been shown in [50] that

$$R_{x_c x_c}(\tau) = R_{x_s x_s}(\tau) = J_0(w_d\tau) \tag{2.16a}$$

$$R_{x_c x_s}(\tau) = R_{x_s x_c}(\tau) = 0 \tag{2.16b}$$

$$R_{hh}(\tau) = 2J_0(w_d\tau)\,. \tag{2.16c}$$

It worth emphasizing that the autocorrelation and cross-correlation functions given by (2.16) do not depend on the number of sinusoids M, and they match the desired

second-order statistics exactly. When M approaches infinity, the envelope $|h(t)|$ is Rayleigh distributed and the phase $\theta_h(t) = \arctan \frac{x_s(t)}{x_c(t)}$ is uniformly distributed over $[-\pi, \pi)$.

Throughout the thesis, MIMO channels are modeled by spatially *independent*, Rayleigh fading processes. This is reasonable when antenna element spacing is considerably larger than the carrier wavelength, or the incoming wave incidence angle spread is relatively large, such as the down link in cellular mobile systems. Also, we focus on the case of *frequency non-selective* channels, where the transmitted signal bandwidth is narrow enough, so that the channel is non-selective.

2.3 Channel Estimation

Channel estimation is one of the most basic issues in communication theory over fading channels, since coherent receivers depend on some information of the current channel state to decode the transmitted signals. There are lots of works devoted to channel estimation and coding, but most of these papers deal with either channel estimation in uncoded systems, or design channel codes on fading channels with perfect CSI. Such simplifications facilitate the analysis, but do not give full insights into the effects of imperfect channel estimation. In this thesis, we will examine the effects of imperfect CSI on the receiver structure, performance and channel code design of MIMO systems. Before we examine these effects, a brief introduction of channel estimation methods is given.

2.3.1 Channel Estimation For SISO Systems

The channel estimation schemes can be classified into training, blind, and semi-blind techniques. For training estimation schemes, sequences of pilot symbols are transmitted to help estimate the channels [40]. With the known pilot symbols,

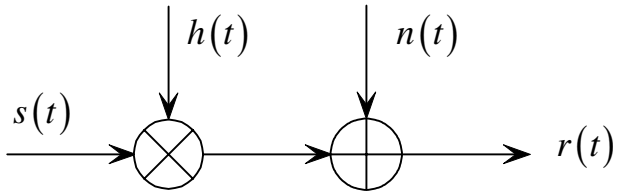


Fig. 2.2: The frequency non-selective SISO fading channel model.

good channel estimates can be obtained at the cost of spectrum efficiency. On the contrary, blind estimation schemes do not use any pilot symbols. The channel estimation is done based on the received noisy signals and the statistical information of the channels. Although blind schemes have high spectrum efficiency, their performances are worse than pilot-aided schemes when the length of the received signal sequence is not sufficiently long. Also, blind schemes have high computational complexity, and require the channels to be constant during estimation. Thus, blind schemes cannot be used to estimate time-varying channels [51]. Between these two extremes, semi-blind schemes are developed to achieve satisfactory estimates with high spectrum efficiency. In semi-blind schemes, short sequences of pilot symbols are transmitted at the beginning of the data burst. This helps obtain good channel estimates at the beginning. Then, detected data symbols are exploited to estimate the channel [52]. Here, we only focus on the non-blind estimation schemes to obtain satisfactory estimates for time-varying channels.

For flat SISO fading channels, the baseband channel model is given in Fig. 2.2, where $v(t)$ and $r(t)$ are the transmitted and received signals, respectively. The multiplicative channel gain $h(t)$ is introduced by the medium while the additive, white, Gaussian noise (AWGN) $n(t)$ arises from the electronic circuitry in the receiver. At the receiver, perfect sample timing is assumed, and thus the multiplicative channel gain is assumed to be piecewise constant for a symbol duration. The model in Fig. 2.2 can be expressed in discrete-time k, where k is the symbol index. The received signal over the kth symbol interval $[kT_s, (k+1)T_s)$, where T_s is the symbol duration time [53], is given by

$$r(k) = h(k)s(k) + n(k)\,. \tag{2.17}$$

In flat fading, if we have a good estimate of the complex channel fading coefficient $h(k)$, then we can perform coherent detection. To obtain a good estimate, one solution is to use the decision-feedback method [52]. This method belongs to semi-blind estimation schemes. It provides satisfactory estimation accuracy with high spectrum efficiency when data decisions are reliable. In [52], the author has provided the conceptual and analytical approach to the optimal receiver design for the nonselective Rayleigh fading channel, based on the criterion of minimum symbol error probability. It has been shown that the optimal minimum symbol error probability receiver has the detector-estimator structure. For the optimal detection of the kth symbol of the transmit signals, the totality received signal due to the entire transmitted data sequence outside the kth interval should be used to estimate the channel fading coefficient $h(k)$, based on all the possible values taken by the transmitted data sequence. Therefore, this optimal receiver is difficult to implement due to the exponential growth in complexity as a function of sequence length. Then, a suboptimal, realizable, decision-feedback receiver is provided. This decision-feedback receiver only uses the past received signals and the past decisions to estimate the channel fading gain $h(k)$ in the kth symbol interval. The decision-feedback channel estimation method is efficient in practice due to its low complexity. However, the decision-feedback channel estimation method assumes that all the data decisions are correct. If there exit errors in the decisions, the performance of the decision-feedback receiver degrades, which is especially true for low SNR.

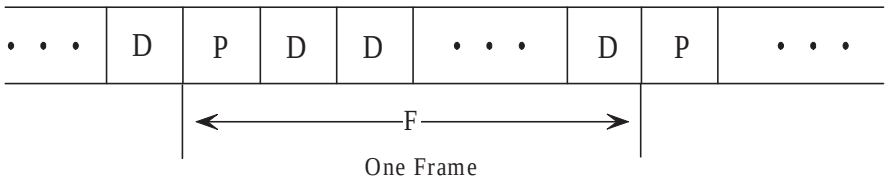


Fig. 2.3: Transmitted frame structure for SISO PSAM.

Another promising channel estimation method is to transmit known signals embedded in the data. The receiver picks out the known signals, estimates the complex fading coefficient, and compensates for it during demodulation. This

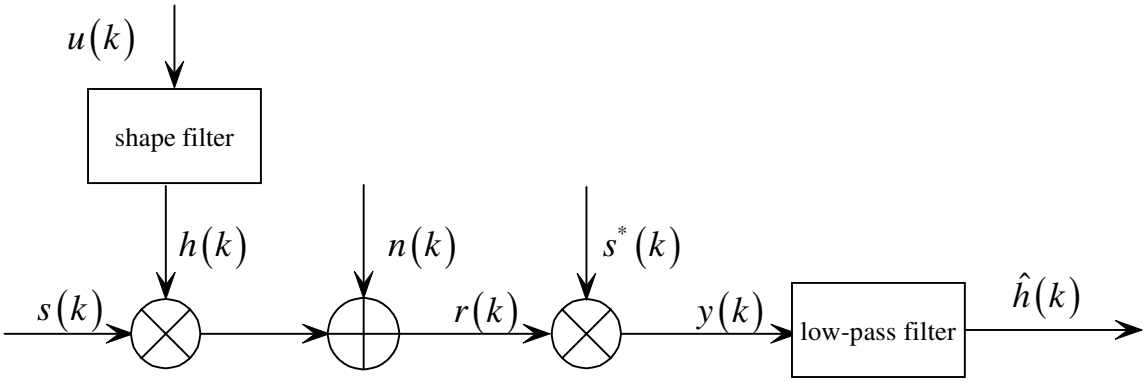


Fig. 2.4: Channel estimation for flat SISO fading channels.

method is called pilot-aided modulation. There are two pilot-aided modulation schemes. One is the pilot-tone assisted modulation (PTAM) scheme [54]. The other is the pilot symbol assisted modulation (PSAM) [40]. Compared with PTAM, PSAM is more preferable. In the PSAM scheme, the reference signal is introduced in the time domain. Known symbols are inserted periodically into the data sequence prior to pulse shaping. The resulting frame structure is shown in Fig. 2.3, where P denotes a pilot and D a data symbol. For rapid fading channels, PSAM is a popular scheme to estimate the channels. In PSAM schemes, the training symbols are separated from the information symbols in time (TDM training). This method can be employed in TDMA systems, such as GSM, where the training symbols are time multiplexed with the user's data.

Suppose $s(k)$ in (2.17) is a pilot symbol, which is known to the receiver. Then, we divide the received signal $r(k)$ by the transmit signal as

$$y(k) = r(k)/s(k) = h(k) + n(k)/s(k) \tag{2.18}$$

where we want to recover the fading coefficient $h(k)$ buried in the white noise $n(k)/s(k)$. The aim of any channel estimation procedure is to minimize an error criterion. The general criterion is the mean-square error (MSE). Here, we give a brief review on the mean-square estimation. Consider the problem of estimating a random variable X in terms of a random observation data vector $\mathbf{Y} = [Y_1 \cdots Y_N]^T$. Let us denote an estimate of the value taken by X as $\hat{x}$ when we know $\mathbf{Y} = \mathbf{y}$. In

general, $\hat{x}$ will not equal x, the actual value taken by X. An average measure of the error is provided by

$$\mathrm{E}[\|X - \hat{x}\|^2 | \mathbf{Y} = \mathbf{y}] \tag{2.19}$$

where $\|\cdot\|$ is the Frobenius norm. We define a minimum mean square error (MMSE) estimate $\hat{x}_0$ of x as one for which

$$\mathrm{E}[\|X - \hat{x}_0\|^2 | \mathbf{Y} = \mathbf{y}] \leq \mathrm{E}[\|X - \hat{x}\|^2 | \mathbf{Y} = \mathbf{y}] \tag{2.20}$$

for all estimates $\hat{x}$, determined in some way from $\mathbf{y}$. A major property of the MMSE estimate is contained in the following theorem [55].

Proposition 2.1 *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two jointly distributed random vectors, and let* $\mathbf{Y}$ *be measured as taking the value* $\mathbf{y}$*. Let* $\hat{\mathbf{x}}_0$ *be the MMSE estimate of* $\mathbf{X}$*. Then* $\hat{\mathbf{x}}_0$ *is also uniquely sepcified as the conditional mean of* $\mathbf{X}$ *given that* $\mathbf{Y} = \mathbf{y}$*, i.e.*

$$\hat{\mathbf{x}}_0 = \mathrm{E}[\mathbf{X} | \mathbf{Y} = \mathbf{y}] \, . \tag{2.21}$$

For Gaussian random variables, it can be shown that the MMSE estimate is equal to the linear MMSE estimate. For linear MMSE estimation, the estimate $\hat{x}$ of X given that $\mathbf{Y} = \mathbf{y}$ is given by

$$\hat{x} = \mathbf{w}^H \mathbf{y} \tag{2.22}$$

where $\mathbf{w}$ is the optimal weight vector such that the mean square value $\varepsilon = \mathrm{E}[\|X - \hat{x}\|^2]$ of the estimation error $X - \hat{x}$ is a minimum. Here we use the orthogonality principle to solve for the optimal weight vector $\mathbf{w}$ [39].

Proposition 2.2 *The mean square value* ε *is a minimum if the error* $X - \hat{x}$ *is orthogonal to the data* $\mathbf{y}$*:*

$$\mathrm{E}\big[(X - \mathbf{w}^H \mathbf{y}) y_i\big] = 0 \quad i = 1, \cdots, N \, . \tag{2.23}$$

The above proposition yields the optimum weight vector

$$\mathbf{w} = \mathbf{\Sigma}_{\mathbf{YY}}^{-1}\mathbf{\Sigma}_{X\mathbf{Y}} \tag{2.24}$$

where $\mathbf{\Sigma}_{\mathbf{YY}} = \mathrm{E}[\mathbf{Y}\mathbf{Y}^H]$ is the data autocorrelation matrix, and $\mathbf{\Sigma}_{X\mathbf{Y}} = \mathrm{E}[X^*\mathbf{Y}]$ is the crosscorrelation between the variable to be estimated and the data. Since $(X - \hat{x}) \perp y_i$ for every i, we conclude that $(X - \hat{x}) \perp \hat{x}$; hence

$$\varepsilon = \mathrm{E}[(X - \hat{x})\hat{x}] = \mathrm{E}[X^2] - \mathbf{\Sigma}_{X\mathbf{Y}}^H(\mathbf{\Sigma}_{\mathbf{YY}}^{-1})^H\mathbf{\Sigma}_{X\mathbf{Y}} \, . \tag{2.25}$$

The optimal weight vector $\mathbf{w}$ obtained above is the Wiener filter weight vector. It is well known that if the temporal correlation of the fading process is known, Wiener filter is the optimal filter. If the fading process is generated by an infinite impulse response (IIR) filter or an autoregressive (AR) filter, a Kalman filter can carry out the estimation recursively, which is computationally efficient.

2.3.2 Channel Estimation For MIMO Systems

For MIMO systems, channel estimation schemes have been mostly based on pilot-assisted approaches, assuming the quasi-static fading channel model. The effects of pilot assisted channel estimation on the achievable data rates over a frequency non-selective, quasi-static fading channel were analyzed in [56]. The lower bounds on the capacity derived in [56] show that the capacity is reduced if a part of channels are allocated to pilot symbols. The effects of the channel estimation errors are represented by the covariance of the effective noise.

Data-assisted channel estimation schemes for MIMO systems are few. It has been known that unlike SISO systems, it is impossible to obtain the estimate of the channel fading matrix based on the MIMO signal model given in (2.4), even if an estimate of the data vector $\mathbf{v}(t)$ is available. This is because the received signal at each receive antenna is the sum of the transmitted data symbols from all transmit antennas. The interantenna interference makes it impossible to solve the estimated

channel fading matrix with only the knowledge of the current data vector. Recently, it is shown that this problem can be solved by using a sequence of data decisions. In [57], an iterative method was derived to estimate channel parameters for an MIMO systems. However, they assume data decisions over the entire transmission have already been made, i.e., the whole data matrix is known to the receiver. In [58], a decision-directed maximum a posteriori probability (MAP) channel estimation scheme for MIMO time-varying fading channels is provided. To estimate the fading coefficient on a link at time t, a window of W channel measurements proceeding time t is used. The channel measurements are constructed by the received signals and the data symbol decisions. However, this method involves the inversion of the data decision matrix. Thus, the complexity increases rapidly with the window size W. Also, as mentioned, the performance degrades at low SNR due to error propagations.

In our thesis, we use the pilot-assisted channel estimation scheme for simple implementation and the satisfactory estimation accuracy. This scheme is first proposed in [59]. In the next chapter, we will build up the TDM STTC system with PSAM for MIMO channel estimation.

Chapter 3

Performance Analysis of STTC over i.i.d. Channels with Channel Estimation

---

In this chapter, we focus on the performance analysis of STTC over rapid fading channels with channel estimation. Here, i.i.d. channels are assumed. The PSAM channel estimation scheme is proposed for MIMO systems and its design parameters are analyzed. Then, the ML receiver incorporating the pilot channel measurements is derived. With the simple ML receiver metric, both analytical PEP and BEP results are obtained. Finally, simulation results are provided to verify the analysis.

3.1 Introduction

Developing effective techniques to evaluate the performance of STTC has attracted great research interests [20], [22], [23], [24], [26], [60], [61], [62]. Tarokh et al. [1] first evaluated the PEP of STTC over fading channels using a Chernoff bound. Based on this PEP bound, a design criterion for STTC over fading channels was

presented. Later, some efforts have been made to obtain the exact PEP, or tighter upper bounds on the PEP [22], [23], [24]. Most recently, researchers found that the distance spectrum should be considered to characterize the error performance of STTC effectively [20], [21], [60]. However, all these papers [1], [20], [22], [23], [24], [26], [60], [61], [62] only consider the case where perfect CSI is available to the receiver. There have been only a few works on STTC error performance analysis with imperfect CSI [41],[42], [43]. All these works on STTC with imperfect CSI, however, considered only the quasi-static fading case. To our knowledge, STTC performance analysis with imperfect CSI over rapid fading channels has not been considered so far. The works of [1], [22], [23] and [26] have considered the rapid fading case, but, as mentioned, they assumed perfect CSI. We also note that all the PEP results with perfect CSI in [22], [23], [26] are implicit, and require the computation of eigenvalues, residues or the numerical evaluation of an integral. The same is true for the PEP results with imperfect CSI in [41]-[43]. The asymptotic PEP result in [62] is explicit. However, their result is only an asymptotic one, valid for a large number of antennas over quasi-static fading channels.

In this chapter, we consider the error performance analysis of STTC with MPSK modulation over rapid, nonselective, Rayleigh fading channels with PSAM for channel estimation. The rapid fading scenario is assumed to be produced by perfect interleaving and de-interleaving. Channel estimation will be performed by exploiting the correlation of the fading process over time before de-interleaving. The choice of the pilot spacing and the Wiener filter length used in the PSAM scheme is a tradeoff between estimation accuracy, transmission rate/pilot overhead and receiver complexity. A direct derivation of the ML receiver with channel estimation is given. A new upper bound on the PEP is obtained in the presence of channel estimation errors. The bound shows explicitly the dependence of the PEP on the MSE of the channel estimates. The channel estimation MSE, in turn, is determined completely by the spectral model of the channel fading process and the channel estimator employed. In the limit of perfect CSI, or no channel estimation errors, our PEP bound reduces to a form similar to that in [1] for rapid fading,

but is tighter. Based on the newly derived PEP bound, an upper bound on the BEP is obtained. The BEP analysis is done using the method of dominant error events proposed in [63]. The choice of the maximum length of the dominant error events to be included in this analysis is critical in this method. We will examine in detail the convergence of the BEP upperbound as the maximum error event length increases. Our work here enables one to analyze the BEP of STTC under actual channel estimation conditions in rapid fading. The BEP upperbound is shown to be very tight in comparison with simulated BEP results. This, in turn, also indicates that the PEP upperbound is tight. The performance loss caused by channel estimation errors increases with the channel fade rate. For high fade rates, the diversity gain achieved by using a large number of transmit antennas can be significantly reduced by the increased estimation error variance.

3.2 The MIMO System with Rapid Fading

The general way to produce the rapid fading scenario is by using interleaving/deinterleaving techniques. Interleaving techniques are traditionally used to enhance the quality of transmission over the bursty wireless channel. This is usually accomplished by scrambling successive symbols of the transmitted sequence into different time slots. A channel is considered fully interleaved when the fading coefficients experienced by consecutive symbols of the received sequence are independent. This is achieved when adjacent symbols of the transmitted sequence are separated by more than the channel coherent time. A channel is considered partially interleaved when consecutive symbols of the received sequence are affected by the coherent fading. Interleaving improves the performance of wireless systems at the cost of increasing memory space requirements, system complexity, and time-delay. An $F \times K$ block interleaver consists of a rectangular array of F rows and K columns. The vertical dimension, F, of the array is called the interleaving depth. The transmitted sequence is usually fed into the array row by row

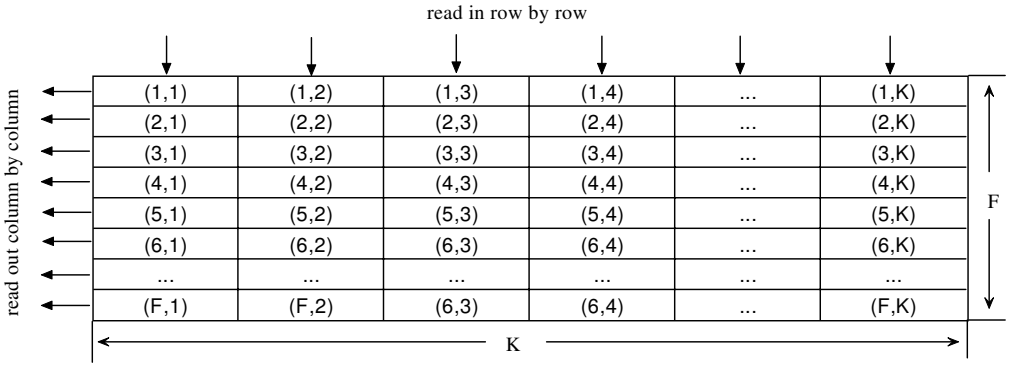


Fig. 3.1: A diagram of a block interleaver.

and shifted out column by column. Successive symbols of the transmitted sequence appear over the channel F symbols apart. The structure of the received sequence is restored by an inverse operation using an $F \times K$ array as a deinterleaver. The block interleaver/deinterleaver memory space requirements and time-delay is $2FK$ symbols for both. Here, we employ the TDMA technique to produce the rapid fading scenario. This can not only be implemented on the existing networks, but also reduce the memory space requirements and the time delay at the transmitter side for each user. In TDMA systems, the data from multiple users are multiplexed and transmitted. The required memory space for each user is only K symbols, and the delay time at the transmission side is F symbols.

We consider a TDM communication system using STTC with N_T transmit and N_R receive antennas. Referring to Fig. 3.2, the data from each user are encoded by a channel encoder and divided into N_T parallel streams. Each stream is pulse shaped and modulated. Here, we consider equal-energy, MPSK modulation. The modulated streams from all users are multiplexed into frames. In one frame, there are F time slots which can support U users, and U is less than F because some time slots are reserved for pilot symbols used for channel estimation. In each time slot, only one user is allowed to access the channel. The user transmits the N_T coded symbols simultaneously in the assigned time slot. Then he must wait for F time slots before he can transmit again in the next frame. Thus, F is the equivalent interleaving depth. Assuming the length of the data packet for a user is KN_T

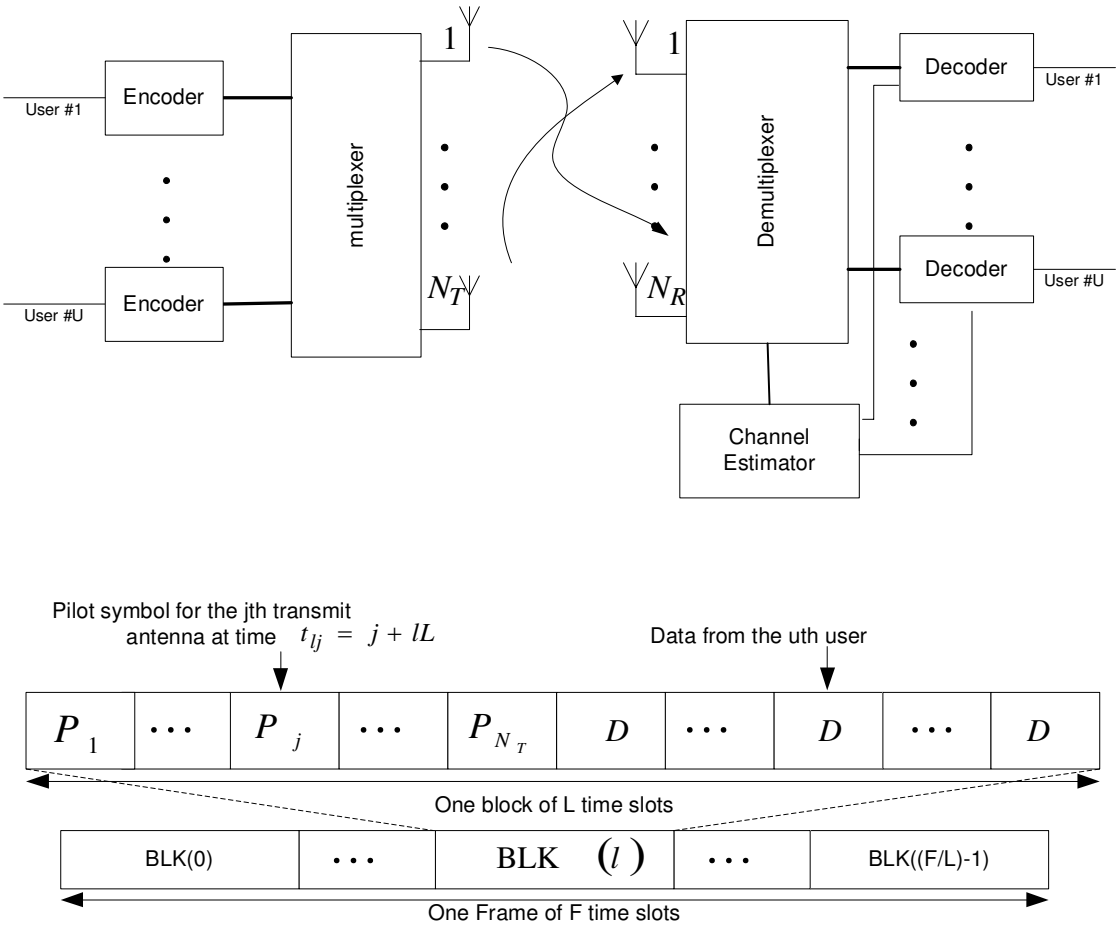


Fig. 3.2: Illustration of the TDM communication system.

symbols, it takes K frames for a user to transmit a packet. At the receiver, the received frames are demultiplexed to the detectors of the different users.

In the kth $(1 \leq k \leq K)$ frame, a codevector $\mathbf{v}(k,u) = [v_1(k,u) \ \cdots \ v_{N_T}(k,u)]^T$ from the uth user is transmitted in the uth $(1 \leq u \leq F)$ time slot, where $v_j(k,u)$ is the code symbol transmitted by the jth antenna. Here, the parameter u not only indicates the time slot index but also labels the user. The pair (k,u) corresponds to a time point t on the real time axis, which is given by

$$t = (k-1)F + u \, . \tag{3.1}$$

The unit of time is the duration of one code symbol. Then the output of the matched filter of the ith receive antenna sampled at the uth time slot in the kth

frame, $r_i(k,u)$, can be expressed as

$$r_i(k,u) = \sqrt{E_s}\sum_{j=1}^{N_T} h_{ij}(k,u)v_j(k,u)+n_i(k,u) \quad i = 1\cdots N_R,\ 1 \le k \le K,\ 1 \le u \le F. \tag{3.2}$$

Here, $\{n_i(k,u)\}$ is the additive, channel white noise sequence at receive antenna i. It is a sequence of i.i.d., complex, Gaussian random variables with $\mathrm{E}[n_i(k,u)] = 0$ and $\mathrm{E}[n_i(k,u)n_i^*(k',u')] = N_0\delta(t-t')$ where $\delta(x)$ is the Kronecker delta, t is given by (3.1), and t' similarly by $t' = (k'-1)F+u'$. The quantity $h_{ij}(k,u)$ is the fading coefficient experienced by the uth user's data at time t on the (i,j)th link. Each link (i,j) introduces slow, nonselective, Rayleigh fading. Thus, the fading coefficient process $\{h_{ij}(k,u)\}$ is a complex, Gaussian process with mean zero and autocorrelation function $\mathrm{E}[h_{ij}(k,u)h_{ij}^*(k',u')] = 2\sigma^2 R(t-t')$, where $R(\tau)$ is the normalized autocorrelation function of both the real and imaginary parts of the fading process, which are independent of one another. The quantity $2\sigma^2$ is the common covariance of the fading coefficients for the different links. We assume that both the additive noises and the fading processes are spatially independent for all links (i,j). We caution the reader that all the variables in (3.2) are functions of a single time variable t which is related to the parameters k and u through (3.1). We show the dependence on k and u in (3.1) for the sake of clarity. Combining the received signals of all the receive antennas at time t as $\mathbf{r}(k,u) = [r_1(k,u)\ r_2(k,u)\ \cdots\ r_{N_R}(k,u)]^T$, the matrix/vector representation of this MIMO system is

$$\mathbf{r}(k,u) = \sqrt{E_s}\mathbf{H}(k,u)\mathbf{v}(k,u) + \mathbf{n}(k,u)\,. \tag{3.3}$$

Here $\mathbf{n}(k,u) = [n_1(k,u)\ n_2(k,u)\ \cdots\ n_{N_R}(k,u)]^T$ is the noise vector for all the receive antennas. The symbolwise constant fading matrix is $\mathbf{H}(k,u)$, where $\mathbf{H}(k,u)$ is assumed to be independent of $\mathbf{n}(k,u)$ for all times t. Since our system incorporates multiplexing and de-multiplexing the data from all the users, the fading process experienced by the uth user's data on each link (i,j) is $\{h_{ij}(k,u)\}_{k=1}^K$, which is a complex, Gaussian process with correlation function $\mathrm{E}[h_{ij}(k,u)h_{ij}^*(l,u)] = 2\sigma^2 R(|k-l|F)$. For sufficiently long frame length F, this correlation approaches zero. Thus, in the limit of large F, the data transmitted by a given user experience

independent fading over time on the link, i.e., $\mathrm{E}[h_{ij}(k,u)h_{ij}^*(l,u)] = 2\sigma^2\delta(k-l)$. Under this independent fading scenario, both space and time diversities are achieved for each user, and this is called the rapid fading case [1]. All the papers that have considered STTC over rapid fading channels assume that perfect CSI is available to the receiver, i.e., the channel fading matrix $\mathbf{H}(k,u)$ is known to the receiver. This is, of course, too idealistic an assumption. Channel estimates are required in order for the receiver to perform coherent detection [52]. The details of channel estimation will be discussed in the next section.

3.3 PSAM Scheme for Channel Estimation

We estimate the channel using the PSAM scheme in which known symbols are inserted periodically into the data sequence. A description of the PSAM method for single-input single-output systems is provided in [40]. A PSAM scheme for MIMO systems without channel coding is presented in [59]. We use a modification of the scheme in [59] below.

The transmitted frame structure is shown in Fig. 3.2. A frame consists of multiple blocks. Each block has a pilot phase and a data phase. The length of a block is L time slots where F/L is an integer, and in each block N_T slots are assigned to pilots and the rest to data. During the pilot phase, N_T pilot symbols are transmitted one at a time in each time slot from each transmit antenna to allow estimation of all the channels associated with that antenna. Within each block, without loss of generality, let the pilot symbol for the jth transmit antenna be transmitted in the jth time slot during which only the jth transmit antenna is active while the other transmit antennas are silent. The pilots for a transmit antenna are thus spaced by L time slots. L is a design parameter that is chosen based on the channel fade rate, among other factors. Since channel estimation exploits the correlation of the slow Rayleigh fading process before de-multiplexing at the receiver, we first rewrite the received signal model of (3.2) in terms of time

t given by (3.1), and thus we have

$$r_i(t) = \sqrt{E_s}\sum_{j=1}^{N_T} h_{ij}(t)v_j(t) + n_i(t), \quad t = 1, \cdots, KF\,. \tag{3.4}$$

The received pilot symbols from the jth transmit antenna in the lth block sampled at the ith receive antenna, $r_i(t_{lj})$, are thus given by

$$r_i(t_{lj}) = \sqrt{E_s}h_{ij}(t_{lj})v_j^p(t_{lj}) + n_i(t_{lj})\,. \tag{3.5}$$

Here, $v_j^p(t_{lj})$ is the pilot symbol transmitted from the jth transmit antenna in the lth block, i.e., at time t_{lj}, where $t_{lj} = j + lL,\ l = 0, 1, \cdots, F/L, \cdots, ((KF/L) - 1)$. From (3.5), it can be seen that the energy of the pilot symbols is assumed to be equal to that of the data symbols. In practical systems, the energy of the pilot symbols could be different from that of the data symbols. High pilot symbol energy may be required to achieve good channel estimation. Thus, to be more general, we rewrite (3.5) as

$$r_i(t_{lj}) = \sqrt{E_p}h_{ij}(t_{lj})v_j^p(t_{lj}) + n_i(t_{lj}) \tag{3.6}$$

where E_p is the energy of the pilot symbols. The value of E_p may be equal to E_s, or be set to a value based on the requirements of the channel estimation accuracy. In Chapter 6, we will examine how to determine the optimal relationship between E_p and E_s. Since the pilot symbols are known, we define the normalized measurement sample $z_i(t_{lj})$ as

$$z_i(t_{lj}) = r_i(t_{lj})/(\sqrt{E_p}v_j^p(t_{lj})) = h_{ij}(t_{lj}) + n_i'(t_{lj})\,. \tag{3.7}$$

Each $z_i(t_{lj})$ is a noisy measurement on the channel coefficient $h_{ij}(t_{lj})$. Here $\{n_i'(t_{lj})\}$ is a set of i.i.d. complex, Gaussian random variables with mean zero and variance N_0/E_p, for all t_{lj} and all i.

We define $\mathcal{I} = \{\{\{z_i(t_{lj})\}_{l=0}^{(KF/L)-1}\}_{i=1}^{N_R}\}_{j=1}^{N_T}$ as the set of all channel measurements obtained from the pilot symbols. In theory, the MMSE estimate of $h_{ij}(t)$ at any time t is $\mathrm{E}[h_{ij}(t)|\mathcal{I}]$. However, due to the spatial independence among the channels and the de-correlation of each channel over time, the estimator will

use only the N nearest (in time) received pilot samples on the (i,j)th link to estimate $h_{ij}(t)$, where N is a design parameter that depends on the fade rate among other factors. The estimator uses $\lfloor N/2 \rfloor$ pilot symbols from the previous blocks, the pilot symbol from the current block where $h_{ij}(t)$ is located, and the pilot symbols from the $\lfloor (N-1)/2 \rfloor$ subsequent blocks. The pilot symbol from the current block is located at the time $t_j = j + \lfloor t/L \rfloor L$. Consequently, the N nearest pilot symbols for $h_{ij}(t)$ are located at times $\{t_{lj} = t_j + lL\}_{l=-\lfloor N/2 \rfloor}^{\lfloor (N-1)/2 \rfloor}$. Let the $N \times 1$ vector of noisy pilot measurements on the (i,j)th link be $\mathbf{z}_{ij}(t) = [z_i(t_j - \lfloor N/2 \rfloor)L) \ \cdots \ z_i(t_j + \lfloor (N-1)/2 \rfloor)L)]^T$. The MMSE estimate of $h_{ij}(t)$ thus reduces to

$$\hat{h}_{ij}(t) = \mathrm{E}\left[h_{ij}(t)|\mathcal{I}\right] \approx \mathrm{E}\left[h_{ij}(t)|\mathbf{z}_{ij}(t)\right] = \mathbf{w}^H \mathbf{z}_{ij}(t) \, . \tag{3.8}$$

Here, $\mathbf{w}$ is the Wiener filter weight vector that minimizes the mean square estimation error $\epsilon_{ij}(t) = \mathrm{E}\left[\left|h_{ij}(t) - \hat{h}_{ij}(t)\right|^2\right]$. By the well-known orthogonality principle, $\varepsilon_{ij}(t)$ is a minimum if the estimation error $e_{ij}(t) = h_{ij}(t) - \hat{h}_{ij}(t)$ is orthogonal to all the noisy fading measurements in $\mathbf{z}_{ij}(t)$. This yields the estimated channel coefficient $\hat{h}_{ij}(t)$ as

$$\hat{h}_{ij}(t) = \mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}^H (\mathbf{\Sigma}_{\mathbf{z}_{ij}}^{-1})^H \mathbf{z}_{ij}(t) \tag{3.9}$$

where the $N \times N$ matrix $\mathbf{\Sigma}_{\mathbf{z}_{ij}} = \mathrm{E}\left[\mathbf{z}_{ij}(t)(\mathbf{z}_{ij}(t))^H\right]$ is the autocorrelation matrix of $\mathbf{z}_{ij}(t)$, with the (m,n)th entry given by

$$\Sigma_{\mathbf{z}_{ij}}(m,n) = \begin{cases} 2\sigma^2 R\left((m-n)L\right) & m \neq n \\ 2\sigma^2 + N_0/E_p & m = n \end{cases} \tag{3.10}$$

while the $N \times 1$ vector $\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}} = \mathrm{E}\left[h_{ij}^*(t)\mathbf{z}_{ij}(t)\right]$ is the crosscorrelation vector of the fading coefficient $h_{ij}(t)$ with the N noisy fading measurements $\mathbf{z}_{ij}(t)$, with its nth entry given by

$$\Sigma_{h_{ij}\mathbf{z}_{ij}}(n) = 2\sigma^2 R\left((n - 1 - \lfloor N/2 \rfloor)L - \tau\right) \, . \tag{3.11}$$

Here τ is the time interval between the fading coefficient to be estimated, $h_{ij}(t)$, and the corresponding pilot symbol of the current block transmitted at time t_j, and is given by

$$\tau = t - t_j = t - j - \lfloor t/L \rfloor L \, .$$

Using the expression for t in (3.1), we get from the last equation,

$$\tau = u - j - \lfloor u/L \rfloor L \,. \tag{3.12}$$

The time interval τ can be seen to be a function of the user index u and the transmit antenna index j. The results in (3.9) through (3.11) show that the Wiener filter weight vector, $\mathbf{w} = \mathbf{\Sigma}_{\mathbf{z}_{ij}}^{-1}\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}$, depends on the time interval τ. The channel estimate $\hat{h}_{ij}(t)$ is a zero-mean, complex, Gaussian random variable with a variance of

$$\mathrm{E}\left[\left|\hat{h}_{ij}(t)\right|^2\right] = 2\hat{\sigma}_{ij}^2(\tau) = \mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}^H(\mathbf{\Sigma}_{\mathbf{z}_{ij}}^{-1})^H\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}} \,. \tag{3.13}$$

The corresponding mean square estimation error is given by

$$2\bar{\sigma}_{ij}^2(\tau) = \mathrm{E}\left[\left|h_{ij}(t) - \hat{h}_{ij}(t)\right|^2\right] = 2\sigma^2 - 2\hat{\sigma}_{ij}^2(\tau) \,. \tag{3.14}$$

We model the actual fading coefficient $h_{ij}(t)$ at any time t as

$$h_{ij}(t) = \hat{h}_{ij}(t) + e_{ij}(t) \tag{3.15}$$

where $e_{ij}(t)$ is the estimation error. The estimation error $e_{ij}(t)$ is also a zero-mean, complex, Gaussian random variable, and attains its minimum variance of $2\bar{\sigma}_{ij}^2(\tau)$. It can be seen from (3.13) and (3.14) that both the estimate variance $\hat{\sigma}_{ij}^2(\tau)$ and the estimation error variance $\bar{\sigma}_{ij}^2(\tau)$ depend on the time interval τ. Since the quantity τ is a function of the transmit antenna index j, the accuracy in estimating the channel coefficient $h_{ij}(t)$ varies from one transmit antenna to another. But both the estimate variance $\hat{\sigma}_{ij}^2(\tau)$ and the estimation error variance $\bar{\sigma}_{ij}^2(\tau)$ are independent of the receive antenna index i. Thus, we can drop i in the sequel, and because of the one-to-one relationship between τ and u for a fixed transmit antenna index j given in (3.12), these two variances will be written more briefly as $\hat{\sigma}_j^2(u)$ and $\bar{\sigma}_j^2(u)$, respectively.

At the receiver, the estimated channel fading processes $\{\hat{h}_{ij}(t)\}$ together with the received signal sequences are demultiplexed to the detectors of the different users. After de-multiplexing, the estimated channel fading process for each user u, $\{\hat{h}_{ij}(k,u)\}_{k=1}^K$, can be easily obtained from $\{\hat{h}_{ij}(t)\}$ by relating t to u and k as in

(3.1). Since the actual fading coefficients experienced by each user are all independent due to perfect interleaving, the corresponding estimated fading coefficients for the uth user, $\{\hat{h}_{ij}(k,u)\}_{k=1}^{K}$, is a sequence of zero-mean, complex, Gaussian random variables which are all mutually uncorrelated, with $\mathrm{E}\left[\hat{h}_{ij}(k,u)\hat{h}_{ij}^{*}(l,u)\right] = 2\hat{\sigma}_j^2(u)\delta(k-l)$.

3.4 The ML Receiver Structure

Since the data from different users are independent, the data detection for different users can proceed in parallel. In the following, we thus focus on one user and drop the user index u. Then the signal model (3.2) may be rewritten in terms of the frame index k alone, as

$$\mathbf{r}(k) = \sqrt{E_s}\mathbf{H}(k)\mathbf{v}(k) + \mathbf{n}(k) \quad k = 1, \cdots, K\ .$$

Suppose the entire transmitted signal sequence of a user is $\mathbf{v} = \left[\mathbf{v}^T(1)\ \cdots\ \mathbf{v}^T(K)\right]^T$. Then, the received signal, $\mathbf{r} = \left[\mathbf{r}^T(1)\ \cdots\ \mathbf{r}^T(K)\right]^T$, is given by

$$\mathbf{r} = \sqrt{E_s}\mathbf{H}\mathbf{v} + \mathbf{n} \tag{3.16}$$

where $\mathbf{H} = diag\left[\mathbf{H}(1)\ \cdots\ \mathbf{H}(K)\right]$ is the combined channel fading matrix and the combined noise vector is $\mathbf{n} = \left[\mathbf{n}^T(1)\ \cdots\ \mathbf{n}^T(K)\right]^T$. The estimated channel fading matrix is $\hat{\mathbf{H}} = \mathrm{E}\left[\mathbf{H}|\mathcal{I}\right]$, and it clearly has the structure $\hat{\mathbf{H}} = diag\left[\hat{\mathbf{H}}(1)\ \cdots\ \hat{\mathbf{H}}(K)\right]$. Here, the (i,j)th element of $\hat{\mathbf{H}}(k)$ is $\hat{h}_{ij}(k)$ given in (3.8). All the elements in $\hat{\mathbf{H}}$ are independent, zero-mean, complex, Gaussian random variables with a variance of $\mathrm{E}\left[\hat{h}_{ij}(k)\hat{h}_{ij}^{*}(l)\right] = 2\hat{\sigma}_j^2\delta(k-l)$, as we have discussed. The estimation error matrix is $\mathbf{E} = diag\left[\mathbf{E}(1)\ \cdots\ \mathbf{E}(K)\right]$, where $\mathbf{E}(k) = \mathbf{H}(k) - \hat{\mathbf{H}}(k)$. The elements $\{e_{ij}(k)\}$ of $\mathbf{E}$ are also independent, zero-mean, complex, Gaussian random variables, each with variance $\mathrm{E}\left[e_{ij}(k)e_{ij}^{*}(l)\right] = 2\bar{\sigma}_j^2\delta(k-l)$.

The estimated channel fading matrix $\hat{\mathbf{H}}$ and the associated estimation error covariance matrix summarize all the information concerning the channel matrix

$\mathbf{H}$ given by the set of pilot channel measurements $\mathcal{I}$. Based on the set of all pilot channel measurements $\mathcal{I}$ and the entire received sequence $\mathbf{r}$, the ML receiver detects the transmitted signal sequence $\mathbf{v}$ by computing the likelihood function $Q_l = p(\mathbf{r}, \mathcal{I}|\mathbf{v} = \mathbf{v}_l)$ for each possible value $\mathbf{v}_l$ of $\mathbf{v}$, and deciding that $\mathbf{v} = \mathbf{v}_j$ if $Q_j = \max_l Q_l$. The likelihood function may be rewritten as

$$Q_l = p(\mathbf{r}|\mathcal{I}, \mathbf{v} = \mathbf{v}_l)\, p(\mathcal{I}|\mathbf{v} = \mathbf{v}_l) \ .$$

Since the pilot channel measurements $\mathcal{I}$ are independent of the data sequence $\mathbf{v}_l$, we have $p(\mathcal{I}|\mathbf{v} = \mathbf{v}_l) = p(\mathcal{I})$ and we can thus drop the term $p(\mathcal{I}|\mathbf{v} = \mathbf{v}_l)$ in the expression of Q_l. The ML receiver is reduced to computing the conditional probability density function $p(\mathbf{r}|\mathcal{I}, \mathbf{v} = \mathbf{v}_l)$. Given $\mathcal{I}$ and $\mathbf{v}$, $\mathbf{r}$ is a complex, Gaussian random vector with a mean of

$$\boldsymbol{\eta} = \mathrm{E}[\mathbf{r}|\mathcal{I}, \mathbf{v}] = \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v} \tag{3.17}$$

and a covariance matrix of

$$\boldsymbol{\Gamma} = \mathrm{E}\left[(\mathbf{r} - \boldsymbol{\eta})(\mathbf{r} - \boldsymbol{\eta})^H|\mathcal{I}, \mathbf{v}\right] = E_s\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right] + N_0\mathbf{I}_{N_RK} \ . \tag{3.18}$$

We now show that the conditional covariance matrix $\boldsymbol{\Gamma}$ is proportional to an identity matrix for equal-energy MPSK modulation. We first define $\boldsymbol{\alpha} = \mathbf{E}\mathbf{v}$. Due to the block-diagonal structure of the estimation error matrix $\mathbf{E}$, we can easily show that $\boldsymbol{\alpha} = \left[\boldsymbol{\alpha}^T(1) \ \cdots \ \boldsymbol{\alpha}^T(K)\right]^T$, where $\boldsymbol{\alpha}(k) = \mathbf{E}(k)\mathbf{v}(k) = [\alpha_1(k) \ \cdots \ \alpha_{N_R}(k)]^T$ for each k. Each element $\alpha_i(k)$ of $\boldsymbol{\alpha}(k)$ is now given by

$$\alpha_i(k) = \sum_{j=1}^{N_T} e_{ij}(k)v_j(k) \ . \tag{3.19}$$

Due to the independence of the elements $\{e_{ij}(k)\}$ of $\mathbf{E}$, we have from (3.19) that

$$\mathrm{E}\left[\alpha_i(k)\alpha_{i'}^H(l)\right] = \mathrm{E}\left[\left|\sum_{j=1}^{N_T} e_{ij}(k)v_j(k)\right|^2\right]\delta(k-l)\delta(i-i') \ . \tag{3.20}$$

Next, noting that the estimation error variances, $\bar{\sigma}_j^2$'s, are independent of receive antenna index i, as mentioned before, (3.20) reduces to

$$\mathrm{E}\left[\alpha_i(k)\alpha_{i'}^H(l)\right] = \left(2\sum_{j=1}^{N_T}\bar{\sigma}_j^2\left|v_j(k)\right|^2\right)\delta(k-l)\delta(i-i') \ . \tag{3.21}$$

The result in (3.21) is also independent of receive antenna index i. Using (3.21) in the first term on the rightmost side of (3.18) gives

$$\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right] = \mathrm{E}\left[\boldsymbol{\alpha}\boldsymbol{\alpha}^H\right] = diag\left[\boldsymbol{\Lambda}(1)\ \cdots\ \boldsymbol{\Lambda}(K)\right] \tag{3.22}$$

where $\boldsymbol{\Lambda}(k) = 2\left(\sum_{j=1}^{N_T}\bar{\sigma}_j^2\left|v_j(k)\right|^2\right)\mathbf{I}_{N_R}$ for $k = 1, \cdots, K$. The diagonal elements of the matrix $\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right]$ in (3.22) have different values in general. However, for equal-energy MPSK modulation, we have $\left|v_j(k)\right|^2 = 1$, for all k and j. This leads to the result from (3.22), that

$$\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right] = 2\left(\sum_{j=1}^{N_T}\bar{\sigma}_j^2\right)\mathbf{I}_{N_RK}\,.$$

Thus, the covariance matrix $\boldsymbol{\Gamma}$ in (3.18) can now be simplified to

$$\boldsymbol{\Gamma} = \left(N_0 + 2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2\right)\mathbf{I}_{N_RK} = \overline{N}_0\mathbf{I}_{N_RK}\,. \tag{3.23}$$

It can be seen that $\boldsymbol{\Gamma}$ is proportional to an identity matrix. Here, we have defined $\overline{N}_0$ by

$$\overline{N}_0 = N_0 + 2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2 \tag{3.24}$$

which is a *constant* dependent only on the statistical information of channel estimation error. We can think of $\overline{N}_0$ as the variance of the *effective noise* at each receive antenna. It is made up of the additive channel noise and the channel estimation errors.

Having obtained the conditional mean $\boldsymbol{\eta}$ and the conditional covariance matrix $\boldsymbol{\Gamma}$ of the Gaussian random vector $\mathbf{r}$, its conditional probability density function can be expressed as

$$p(\mathbf{r}|\mathcal{I}, \mathbf{v} = \mathbf{v}_l) = \frac{1}{(\pi)^{N_RK}|\boldsymbol{\Gamma}|}\exp\left\{-(\mathbf{r}-\boldsymbol{\eta})^H\boldsymbol{\Gamma}^{-1}(\mathbf{r}-\boldsymbol{\eta})\right\}.$$

Using the expression for $\boldsymbol{\Gamma}$ in (20), the ML receiver can be seen to reduce to minimizing the metric

$$m(\mathbf{r}, \mathbf{v}) = \|\mathbf{r} - \boldsymbol{\eta}\|^2 = \|\mathbf{r} - \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v}\|^2\,. \tag{3.25}$$

Thus, the receiver only computes the metric $m(\mathbf{r}, \mathbf{v}_l)$ for each possible value $\mathbf{v}_l$ of $\mathbf{v}$, and decides that $\mathbf{v} = \mathbf{v}_j$ if $m(\mathbf{r}, \mathbf{v}_j) = \min_l m(\mathbf{r}, \mathbf{v}_l)$. The metric (3.25) uses the channel estimate $\hat{\mathbf{H}}$ as if it were the actual channel fading matrix $\mathbf{H}$. The fact that the ML receiver reduces to minimizing the simple metric (3.25) is due to the fact that $\mathbf{\Gamma}$ is proportional to the identity matrix as shown in (3.23). This property of $\mathbf{\Gamma}$ is attributed to the equal-energy MPSK modulation and to the PSAM scheme for channel estimation. The structure of our PSAM channel estimation scheme leads to the channel estimation error variances, $\bar{\sigma}_j^2$'s, being independent of the receive antenna index i. For general non-equal-energy modulation schemes, the diagonal elements of $\mathbf{\Gamma}$ are not equal. The decision metric would then have to take into account the effect of the covariance matrix $\mathbf{\Gamma}$. Although the extension of the ML receiver to non equal-energy modulation is straightforward, we will not consider it here. We focus on equal-energy MPSK modulation here, and derive an explicit upper bound on the PEP.

Reference [64] presents a similar receiver derivation but it considers the *quasi-static* fading case.

3.5 Error Performance Analysis

3.5.1 The PEP Upper Bound

Since the ML metric (3.25) is a Euclidean distance metric given the pilot channel measurements $\mathcal{I}$, the conditional PEP $P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c)$ that the receiver decides in favor of an erroneous codeword $\mathbf{v}_e$ when the actual codeword sent is $\mathbf{v}_c$ and $\mathbf{v}_e$ is the only other alternative, is given by [65]

$$\begin{aligned} P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) &= P\left(\|\mathbf{r} - \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v}_e\|^2 - \|\mathbf{r} - \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v}_c\|^2 < 0|\mathcal{I}, \mathbf{v} = \mathbf{v}_c\right) \\ &= Q\left(\sqrt{\frac{E_s\|\hat{\mathbf{H}}\mathbf{v}_{ce}\|^2}{2\overline{N}_0}}\right) \end{aligned} \tag{3.26}$$

where $\mathbf{v}_{ce} = \mathbf{v}_c - \mathbf{v}_e$ is the code difference vector, each of whose elements $\{v_{cej}(k)\}_{j,\,k=1}^{N_T,\,K}$ is a code difference symbol $v_{cej}(k) = v_{cj}(k) - v_{ej}(k)$. We first write the quantity $\|\hat{\mathbf{H}}\mathbf{v}_{ce}\|^2$ as

$$\|\hat{\mathbf{H}}\mathbf{v}_{ce}\|^2 = \sum_{i=1}^{N_R} \sum_{k\in\kappa} \left| \sum_{j=1}^{N_T} \hat{h}_{ij}(k) v_{cej}(k) \right|^2 . \tag{3.27}$$

Here κ is the set of frame indices k where $\mathbf{v}_c(k) \neq \mathbf{v}_e(k)$. We define

$$g_i(k) = \sum_{j=1}^{N_T} \hat{h}_{ij}(k) v_{cej}(k) . \tag{3.28}$$

From (3.28), $g_i(k)$ is a linear combination of independent, Gaussian random variables $\{\hat{h}_{ij}(k)\}_{j=1}^{N_T}$. It is then obvious that $g_i(k)$ is a complex, Gaussian random variable with a mean of zero and a variance of $\mathrm{E}\left[|g_i(k)|^2\right] = 2\sigma_g^2(k) = 2\sum_{j=1}^{N_T} \hat{\sigma}_j^2 d_j^2(k)$ that is independent of receive antenna index i. The quantity $d_j(k) = |v_{cej}(k)|$ is the Euclidean distance between the erroneous code symbol $v_{ej}(k)$ and the actual transmitted code symbol $v_{cj}(k)$ from the jth transmit antenna in the kth frame. Also, $g_i(k)$ is independent of $g_{i'}(k')$, for $i \neq i'$ or $k \neq k'$ or both, due to the independence of the estimated channel coefficients $\hat{h}_{ij}(k)$. The conditional PEP (3.26) can now be expressed as

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathcal{I}, \mathbf{v} = \mathbf{v}_c) = Q\left(\sqrt{\frac{E_s}{2\overline{N}_0} \sum_{i=1}^{N_R} \sum_{k\in\kappa} |g_i(k)|^2} \right) . \tag{3.29}$$

From the conditional PEP (3.29), we need to average over all possible realizations of $\{g_i(k)\}_{i,\,k=1}^{N_R,\,K}$ to obtain the desired PEP $P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c)$. In general, this is an average over a weighted chi-square distribution. Although there are numerical methods to calculate this PEP exactly [22], [23], such methods do not lead to explicit results and would not provide any insight into code design. This motivates us to obtain a simple, explicit, tight upper bound on the average PEP. First, using the bound $Q(x) < e^{-x^2/2}/2$, the right hand side in (3.29) can be bounded as

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathcal{I}, \mathbf{v} = \mathbf{v}_c) < \frac{1}{2} \prod_{i=1}^{N_R} \prod_{k\in\kappa} \exp\left(-\frac{E_s}{4\overline{N}_0} |g_i(k)|^2 \right) . \tag{3.30}$$

The right hand side of (3.30) is now a product of independent random variables, and the average can be taken term by term. Noting that the average of each term

over the quantity $g_i(k)$ is

$$\mathrm{E}\left[\exp\left(-\frac{E_s}{4\overline{N}_0}\left|g_i(k)\right|^2\right)\right] = \left(1+\frac{2\sigma_g^2(k)E_s}{4\overline{N}_0}\right)^{-1}$$

for all i [65, eq. (2.1-109)], we obtain the PEP upper bound [1]

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v}=\mathbf{v}_c)|_{ICSI} < \frac{1}{2}\prod_{k\in\kappa}\left(1+\frac{E_s\sum_{j=1}^{N_T}\hat{\sigma}_j^2 d_j^2(k)}{2\left(N_0+2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2\right)}\right)^{-N_R}. \qquad (3.31)$$

The result (3.31) is our new upper bound on the PEP over rapid Rayleigh fading channels with imperfect CSI at the receiver. The effect of channel estimation on the error performance is clear. First, the channel estimation errors introduce more noise at the receiver as shown by the effective noise term $\overline{N}_0 = N_0 + 2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2$. Second, the Euclidean distances between code symbols are *weighted* by the different variances of the estimated channel coefficients associated with the different transmit antennas. Thus, the structure and performance of codes are expected to be affected by the quality of the channel estimates. Optimum code design with imperfect CSI should take into consideration the statistical information of the channel estimates. Finally, our result is explicit. Both the variances $\hat{\sigma}_j^2$'s in (3.13) of the estimated channel coefficients and the channel estimation error variances $\bar{\sigma}_j^2$'s in (3.14) are completely determined by the channel estimator and the channel fading spectral model. Thus, the PEP upperbound (3.31) is easy to compute once the channel fading model and the channel estimator are given. Our results are the first to show explicitly the effect of channel estimation error on the PEP performance of STTC over rapid fading channels. This result promises to ease performance analysis and code design.

From (3.31), in the limit of perfect channel estimation, where we have $\bar{\sigma}_j^2 = 0$, or $\hat{\sigma}_j^2 = \sigma^2$ for all j, the PEP bound reduces to that of the perfect CSI case. Thus, for perfect CSI (PCSI), we have

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v}=\mathbf{v}_c)|_{PCSI} < \frac{1}{2}\prod_{k\in\kappa}\left(1+\frac{2\sigma^2 E_s}{4N_0}\left(\sum_{j=1}^{N_T} d_j^2(k)\right)\right)^{-N_R}. \qquad (3.32)$$

[1] Imperfect CSI is abbreviated as ICSI in equations and figures throughout this thesis.

Since $\|\mathbf{v}_{ce}(k)\|^2 = \sum_{j=1}^{N_T} d_j^2(k)$ by definition, the new PEP upper bound (3.32) for perfect CSI channels is similar to the well-known bound in [1, eq. (17)] given by

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v} = \mathbf{v}_c)|_{PCSI} < \prod_{k\in\kappa} \left(\frac{2\sigma^2 E_s \|\mathbf{v}_{ce}(k)\|^2}{4N_0} \right)^{-N_R}.$$

However, our bound (3.32) is tighter.

3.5.2 The Estimated BEP Upperbound

We apply here the PEP upperbound (3.31) to obtain an upperbound on the BEP, which is ultimately the performance measure of interest. Since the PEP result (3.31) is in a product form, the upper bound on the BEP can be obtained by using, for instance, the transfer function technique in [60]. Alternatively, an accurate estimate of the BEP P_b can be obtained more simply through the method of dominant error events [23], [63]. This method estimates the BEP P_b as

$$P_b \approx \frac{1}{n} \sum_{\mathbf{v}_c} \sum_{\mathbf{v}_e} a(\mathbf{v}_c \to \mathbf{v}_e) P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v} = \mathbf{v}_c) P(\mathbf{v}_c) \tag{3.33}$$

where n is the number of input bits per code sequence under consideration, and $a(\mathbf{v}_c \to \mathbf{v}_e)$ is the number of bit errors associated with the error event of deciding in favor of $\mathbf{v}_e$ when $\mathbf{v}_c$ was sent. In (3.33), we only sum over erroneous codewords $\mathbf{v}_e$ with error event lengths from the minimum value L_{pmin} which depends on the code structure, up to a certain maximum value L_{pmax} which we choose. The choice of L_{pmax} is critical in this BEP estimation method. If L_{pmax} is too small, the estimated BEP may not be accurate enough because it does not take into account enough dominant error events. On the other hand, we do not need a very large L_{pmax}, because the BEP computed via (3.33) converges rapidly as L_{pmax} increases. Thus, a proper L_{pmax} should be chosen. We will examine the choice of L_{pmax} in the next section. Our work here is an extension of the method of using dominant error events to estimate the BEP to the case with channel estimation errors.

Since each term $P(\mathbf{v}_c \rightarrow \mathbf{v}_e|\mathbf{v}_c)$ in (3.33) can be upper-bounded using the result in (3.31), we can express the estimated BEP upperbound as

$$P_b < \frac{1}{n}\sum_{\mathbf{v}_c}\sum_{\mathbf{v}_e}\left(\frac{a(\mathbf{v}_c \rightarrow \mathbf{v}_e)P(\mathbf{v}_c)}{2}\prod_{k\in\kappa}\left(1+\frac{E_s\sum_{j=1}^{N_T}\hat{\sigma}_j^2 d_j^2(k)}{2\left(N_0+2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2\right)}\right)^{-N_R}\right). \tag{3.34}$$

In the next section, we will show the tightness of the upperbound (3.34) when computed using the choice of L_{pmax} given by $L_{pmax} = L_{pmin} + \nu - 1$, where ν is the memory order of the code and L_{pmin} is the shortest error event length. This work will serve also to show the tightness of our PEP upperbound (3.31).

3.6 Performance Results

We use the codes of [1] and [2] (abbreviated here as TSC and FVY codes, respectively) as examples to demonstrate the validity of our performance analysis. Codes with different numbers of states and using different MPSK modulation schemes will be considered. We assume each channel fading process has a Jake's power spectrum so that $R(\tau) = J_0(2\pi f_d \tau)$, although our analysis is not restricted to this specific model. The PSAM scheme is used to estimate the channels. In the numerical examples below, we consider a user whose time slot is located immediately after the pilot phase in a block. The energy of the pilot symbols is assumed to be equal to that of the data symbols, i.e., $E_p = E_s$.

Using the PSAM scheme to estimate the channels, we first need to get an idea of the pilot spacing L and the Wiener filter length N for optimum performance. To obtain the optimum performance, we should minimize the performance loss caused by estimation errors. We measure the effect of estimation errors by the total estimation error variance $\bar{\sigma}^2$, which is the sum of the estimation error variances from all transmit antennas, i.e., $\bar{\sigma}^2 = \sum_{j=1}^{N_T}\bar{\sigma}_j^2$. As can be seen in (3.31), minimizing the total estimation error variance $\bar{\sigma}^2$ minimizes the PEP and, hence, the BEP in (3.34). In Fig. 3.3 and Fig. 3.4, we plot the total estimation error variance $\bar{\sigma}^2$

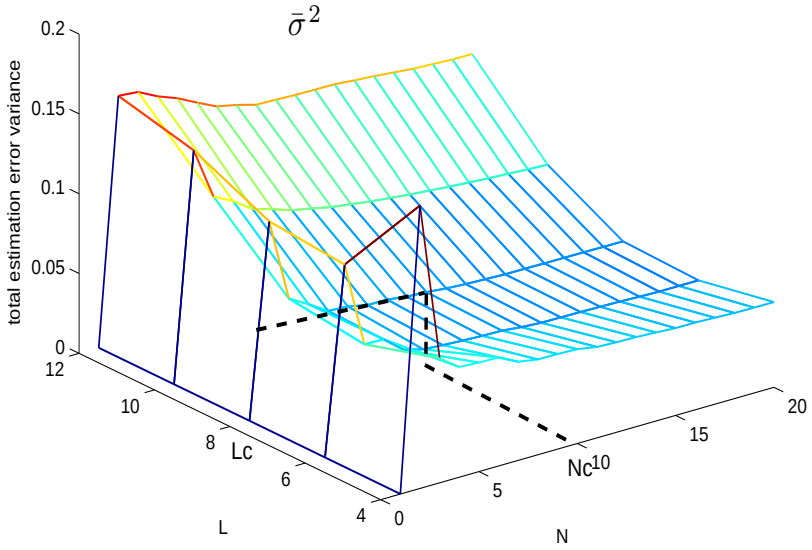


Fig. 3.3: The surface of the total estimation error variance $\bar{\sigma}^2$ as a function of the pilot spacing L and the Wiener filter length N, with two transmit antennas and $f_dT = 0.05$ at $E_s/N_0 = 15\ dB$.

as a function of the pilot spacing L and the Wiener filter length N for SNR of $E_s/N_0 = 15\ dB$, and for two values of the normalized fade rate, i.e., $f_dT = 0.05$ and $f_dT = 0.01$, respectively. We assume $N_T = 2$ transmit antennas. To be fair, we have to account for the energy spent on the pilot symbols. Thus, the effective energy per data symbol in the PSAM scheme is $E'_s = E_s(L - N_T)/L$. The fraction of overhead is N_T/L. It can be observed that for a given Wiener filter length N, the total estimation error variance $\bar{\sigma}^2$ increases rapidly with L for L greater than a certain value. On the other hand, one would like to use as large a value of L as possible to minimize the faction of overhead N_T/L, and to maximize the data transmission rate. This means that as a tradeoff we should choose a suitable compromise value for the pilot spacing L. Similarly, one would in general like to have a longer Wiener filter length N to minimize the total estimation error variance $\bar{\sigma}^2$. However, as can be observed in Fig. 3.3 and Fig. 3.4, for each pilot spacing L, the total estimation error variance $\bar{\sigma}^2$ decreases very slowly for values of N

greater than a certain value. The decrease in the total estimation error variance $\bar{\sigma}^2$ obtained by increasing N beyond this value is not sufficiently significant to justify the increase in filter complexity. Thus, again, as a tradeoff, we also choose a suitable compromise value for N. For each value of SNR E_s/N_0, fade rate f_dT and number of transmit antennas N_T, we use a plot like that in Fig. 3.3 or Fig. 3.4 to determine values L_c and N_c, which are the most suitable ones to be used for the pilot spacing L and the Wiener filter length N, respectively, based on the above considerations. For the case of $N_T = 2$, $E_s/N_0 = 15\ dB$, and $f_dT = 0.05$ in Fig. 3.3, we choose $L_c = 8$ and $N_c = 10$, respectively. Decreasing the fade rate to $f_dT = 0.01$ in Fig. 3.4, we find that we can increase L_c to 16 and N_c to 12, respectively. Our investigations show that L_c and N_c are not very sensitive to SNR, varying only slightly as SNR changes. We find that the values of L_c and N_c chosen at a SNR of $E_s/N_0 = 15\ dB$ work well for the entire SNR range of about 8 dB to about 30 dB.

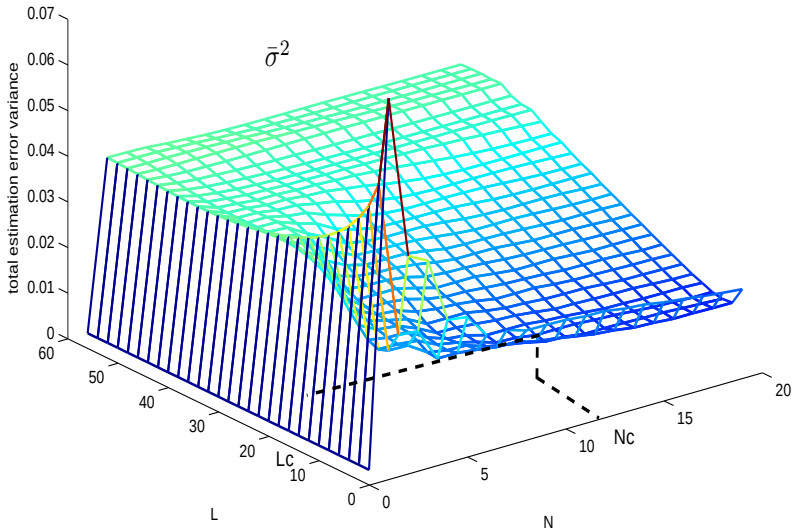


Fig. 3.4: The surface of the total estimation error variance $\bar{\sigma}^2$ as a function of the pilot spacing L and the Wiener filter length N, with two transmit antennas and $f_dT = 0.01$ at $E_s/N_0 = 15\ dB$.

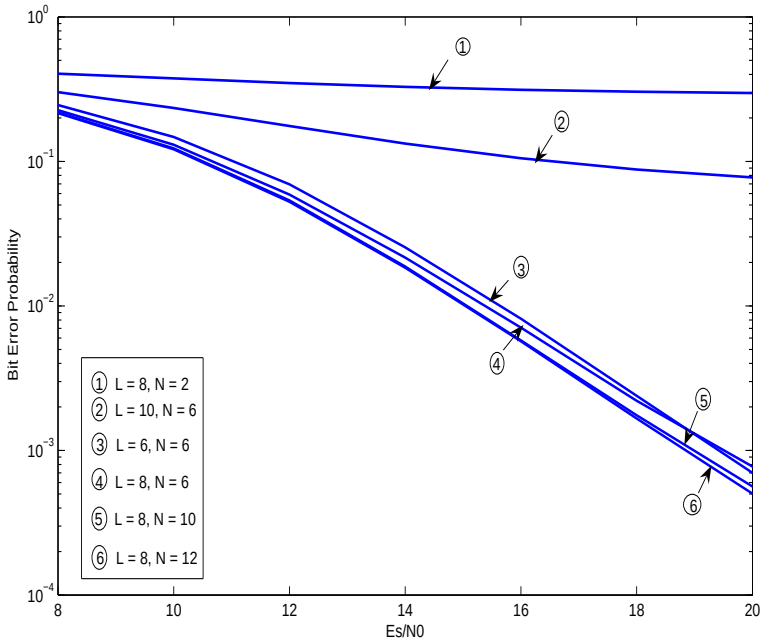


Fig. 3.5: The simulated BEP performances of the 8-state QPSK TSC code of [1] with different channel estimation parameters L and N, using two transmit and one receive antenna, and $f_dT = 0.05$.

The choice of L and N discussed here is verified in Fig. 3.5, where we show the simulated BEP performance of the 8-state QPSK TSC code with two transmit and one receive antenna and a normalized fade rate of $f_dT = 0.05$. Through all the simulated results, we take the data packet length as $K = 100$. The simulations show that the performance is near optimum with $L = 8$ and $N = 10$. The performance with $L = 8$ and $N = 12$ is nearly the same as that with $N = 10$ above, but this is achieved at the cost of higher receiver complexity. When the pilot spacing is too large, for instance, when L is increased to 10, there is a high error floor caused by the large channel estimation errors. In fact, we can obtain satisfactory performance close to optimum with the lowest complexity using $L = 8$ and $N = 6$. Taking into account channel estimation accuracy, transmission rate, fraction of overhead, and filter design complexity, we will hereafter choose the pilot spacing as $L = 8$ and the Wiener filter length as $N = 6$ for the case of two transmit antennas to estimate channels whose normalized fade rate f_dT is less than or equal to 0.05, unless stated

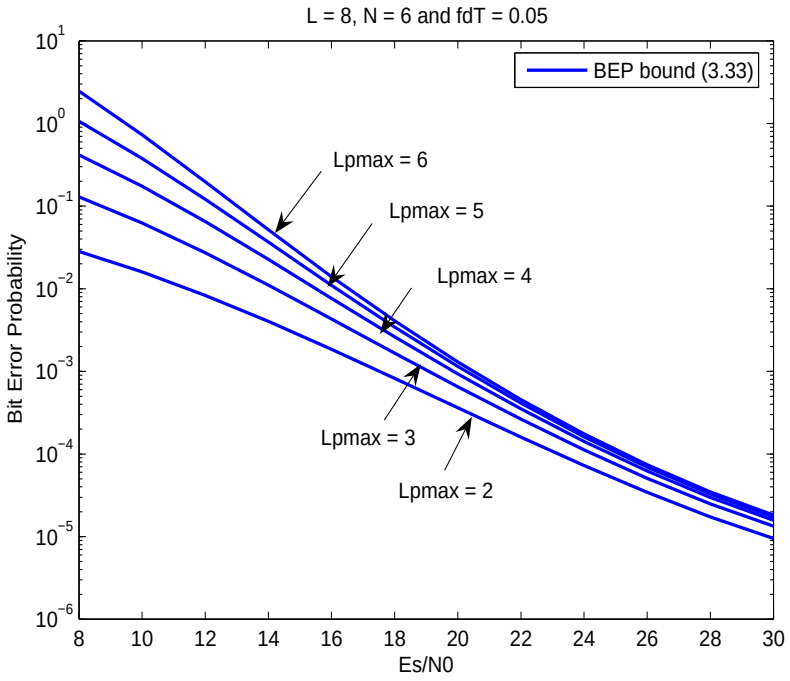


Fig. 3.6: Convergence of the BEP upperbound (3.34) for the 8-state QPSK TSC code of [1] with two transmit and one receive antenna, where $L = 8$, $N = 6$ and $f_dT = 0.05$.

otherwise.

Fig. 3.6 shows the BEP of the 8-state QPSK TSC code of [1], computed using the BEP upperbound (3.34) with various values of L_{pmax}. The results show that the upperbound (3.34) converges rapidly as the value of L_{pmax} used increases, especially at high SNR. This is also true for the 8-state QPSK FVY codes of [2], as shown in Fig. 3.7. Thus, we do not need a very large L_{pmax} to get an accurate estimate of the BEP upperbound, especially at high SNR. In general, the BEP upperbound (3.34) has practically converged when $L_{pmax} = L_{pmin} + \nu - 1$. Therefore, a larger L_{pmax} is needed for codes with a larger number of states. For QPSK modulation, the 8-state STTC with the linear structure of [66] has $L_{pmin} = 2$, and the 32-state STTC has $L_{pmin} = 3$.

Next, we will show that the BEP upperbound (3.34) is tight when computed using the choice of L_{pmax} given by $L_{pmax} = L_{pmin} + \nu - 1$. In Fig. 3.8 and Fig. 3.9, we consider TSC/FVY codes with different numbers of states under

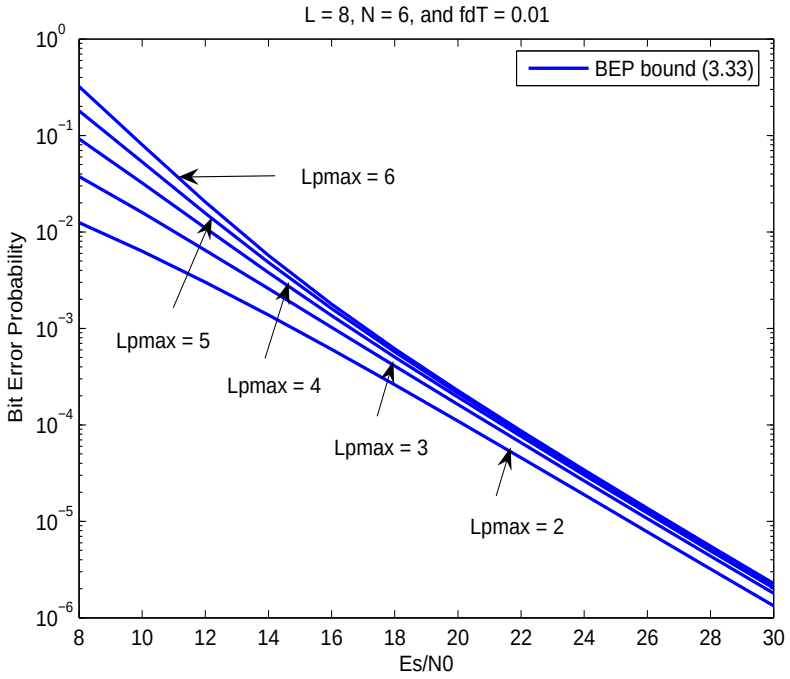


Fig. 3.7: Convergence of the BEP upperbound (3.34) for the 8-state QPSK FVY code of [2] with two transmit and one receive antenna, where $L = 8$, $N = 6$ and $f_dT = 0.01$.

two modulation schemes, namely, QPSK and 8PSK. Two normalized fade rates are considered, namely, $f_dT = 0.05$ and $f_dT = 0.01$. The simulation results show that the computed BEP upperbound (3.34) is tight when computed using $L_{pmax} = L_{pmin} + \nu - 1$. Thus, we conclude that the BEP upperbound (3.34) has practically converged by choosing $L_{pmax} = L_{pmin} + \nu - 1$, and it provides the desired accuracy.

In Fig. 3.10, we use our BEP bound (3.34) to analyze the performance loss, due to increasing fade rates, of STTC with PSAM channel estimation over rapid fading channels, compared with the case of perfect CSI. In this case, we choose the channel estimation parameters L and N for different fade rates, as discussed above. It is shown that at a very low fade rate, i.e., $f_dT = 0.001$, the performance loss is about 1 dB compared with that of perfect CSI. For a higher fade rate of $f_dT = 0.005$, the loss increases to about 2 dB, because the higher fade rate leads to a greater estimation error variance.

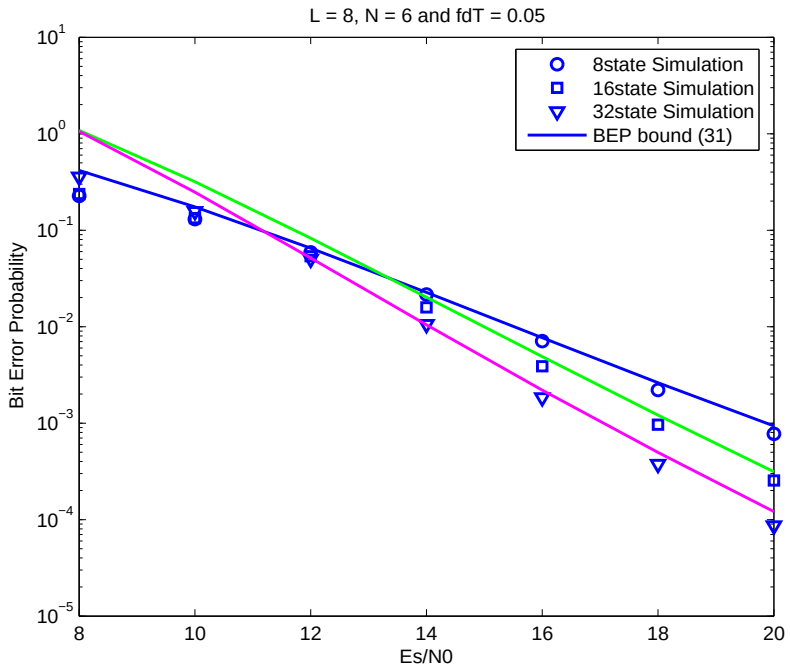


Fig. 3.8: The BEP analysis and simulation results for the QPSK TSC codes of [1] under imperfect CSI with $f_dT = 0.05$, using two transmit and one receive antenna, with $L = 8$ and $N = 6$.

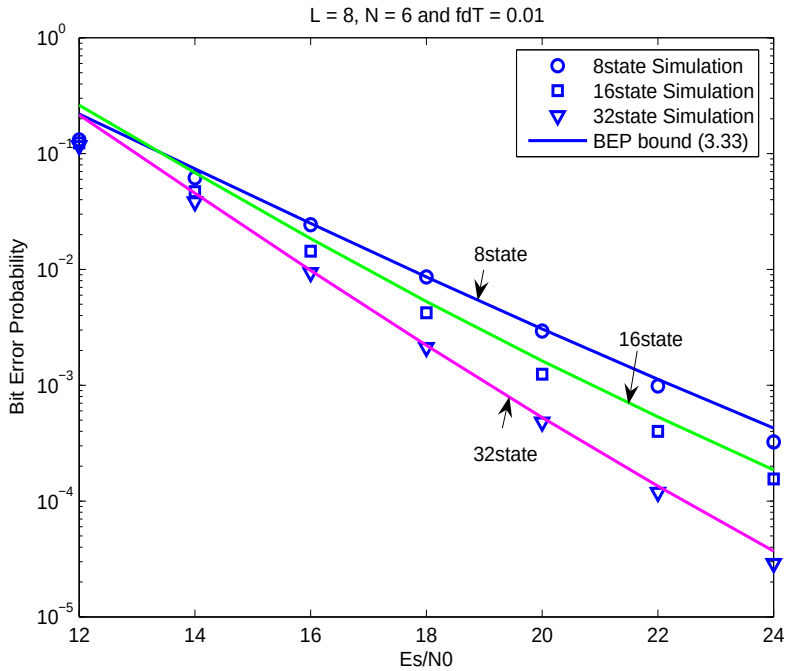


Fig. 3.9: The BEP analysis and simulation results for the 8PSK FVY codes of [2] under imperfect CSI with $f_dT = 0.01$, using two transmit and one receive antenna, with $L = 8$ and $N = 6$.

In our PSAM scheme, the number of transmit antennas N_T determines the pilot duration for channel estimation. A larger N_T leads to a greater decorrelation between the channel coefficients to be estimated and the corresponding pilot symbol measurements. The decorrelation increases with the fade rate, and leads to a larger total estimation error variance $\bar{\sigma}^2$. Thus, a larger N_T can lead to a greater performance loss, especially for high fade rates. While the diversity gain of STTC increases with N_T for the perfect CSI case, it can be reduced or even lost in the imperfect CSI case. This motivates us to analyze the performance loss of STTC with different numbers of transmit antennas. In Fig. 3.11, we examine the performance of two STTC with imperfect CSI. One is the QPSK 8-state FVY code of [2] with two transmit antennas, while the other is the QPSK 8-state code of [3] (abbreviated as CVYL code) with four transmit antennas. Channel estimation parameters L and N are again chosen as in Fig. 3.10. It can be seen that for a small fade rate of $f_dT = 0.001$, the diversity gain achieved by the CVYL code using $N_T = 4$ transmit antennas is still maintained. However, for a high fade rate of $f_dT = 0.05$, the CVYL code performs worse than the FVY code with $N_T = 2$ transmit antennas. The diversity gain achieved by the larger number of transmit antennas is thus lost due to the increased channel estimation error variance caused by a high fade rate.

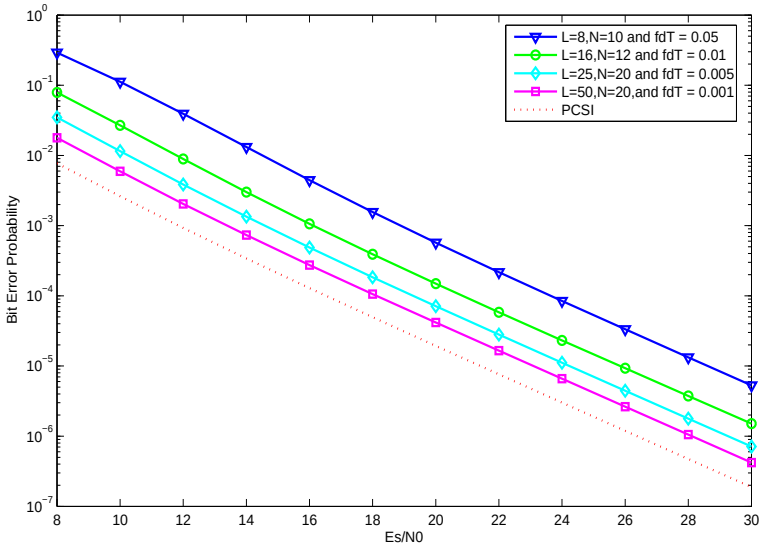


Fig. 3.10: The BEP analysis of the QPSK 8state FVY code of [2] with two transmit and one receive antenna for the perfect CSI case, and the imperfect CSI case.

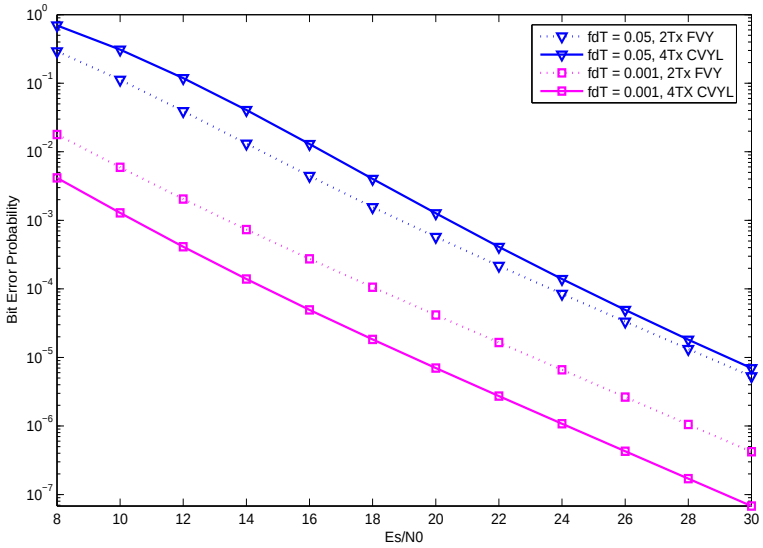


Fig. 3.11: The comparison of the BEP of the QPSK 8state FVY code of [2] with two transmit antennas, with the QPSK 8state CVYL code of [3] with four transmit antennas, under imperfect CSI.

3.7 Summary

In this chapter, we used a PSAM scheme to estimate the Rayleigh, rapid fading MIMO channels. The pilot spacing and the Wiener filter length are chosen to achieve good estimation accuracy while maintaining low receiver complexity. Based on the channel estimates obtained, we derived a ML receiver with imperfect CSI for STTC with MPSK modulation. Explicit, tight upperbounds on the PEP and the BEP of the ML receiver are obtained. The BEP bound enables us to easily evaluate the performance of STTC over rapid fading channels with channel estimation errors. Using the BEP result, we observed that the performance loss of STTC caused by channel estimation errors increases with the channel fade rate. More importantly, it is shown that in contrast to the perfect CSI case, increasing the number of transmit antennas in the imperfect CSI case may lead to a reduced or even negative transmit diversity advantage if the fade rate is high, because of the increased total estimation error variance.

Chapter 4

Code Design of STTC over i.i.d. Channels with Channel Estimation

4.1 Introduction

In the previous chapter, we have presented the receiver structure and performance analysis of STTC with imperfect CSI at the receiver. Based on the analysis results, we here will focus on code design of STTC over rapid fading channels with channel estimation. Research on space-time block codes with channel estimation has been done in [67], but we will consider here STTC. Although a closed-form expression for the PEP of space-time codes with imperfect CSI was presented in [43] for the quasi-static fading case, the complicated and implicit nature of the expression made the implementation of code design difficult. The PEP upperbound obtained in chapter 3, which is published in [68], is the only closed-form, explicit result so far for STTC over rapid fading channels with imperfect CSI. The exact PEP and tighter PEP bounds are further derived in [69]. These PEP upperbounds promise to ease the design and search for STTC over rapid fading channels with imperfect CSI.

In this chapter, we propose a new code design criterion for STTC using MPSK modulation over rapid, nonselective, Rayleigh fading channels with imperfect CSI at the receiver, based on the explicit PEP upperbound obtained in chapter 2. This PEP upperbound shows explicitly the dependence of the code performance on the mean square error (MSE) of the channel estimates. Using these simple and explicit PEP expressions, our code search is practical to implement. The PEP results lead to a union bound on the node error event probability. Our design criterion uses the node error event probability as the performance measure, which takes into account the number of error event paths. This is necessary because for general STTC, the geometrical uniform property is not always satisfied, and considering only the minimum PEP as a code design criterion is not sufficient [21]. More importantly, our criterion incorporates the statistical information concerning the channel estimates, and this leads to codes which have better performance under actual channel estimation conditions. It is found that the effect of channel estimation errors on code design can be measured by the *maximum estimation variance difference*. This maximum estimation variance difference increases with the channel fade rate and the number of transmit antennas. This implies that with an increase in the channel fade rate or the number of transmit antennas, the codes based on our new criterion can achieve more performance gains than those of existing STTC which are designed under the perfect CSI assumption. Since the effect of channel estimation depends on the channel fade rate, based on the knowledge of the channel fade rate at the transmitter, we consider two channel conditions for code design. When the channel fade rate is known to the transmitter, new STTC with imperfect CSI are obtained by using a novel iterative search algorithm. The BEP simulation results demonstrate the advantages of our new STTC under actual channel estimation situations. On the other hand, when the channel fade rate is unknown to the transmitter, robust codes are proposed based on the knowledge of the distribution of the channel fade rates, and shown to have the best average performance.

4.2 Code Design with Channel Estimation

It has been shown in (3.31) the PEP upperbound of STTC over rapid fading channels with imperfect CSI is given by

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v}=\mathbf{v}_c)|_{ICSI} < \frac{1}{2}\prod_{k\in\kappa}\left[b\left(\mathbf{v}_c(k),\mathbf{v}_e(k)\right)\right]^{-N_R} \tag{4.1}$$

where κ is the set of frame indices k where $\mathbf{v}_c(k) \neq \mathbf{v}_e(k)$, and the *branch distances* $b(\mathbf{v}_c(k),\mathbf{v}_e(k))$'s are defined by,

$$b(\mathbf{v}_c(k),\mathbf{v}_e(k))|_{ICSI} = 1 + \frac{E_s\sum_{j=1}^{N_T}\hat{\sigma}_j^2 d_j^2(k)}{2\left(N_0 + 2E_s\sum_{j=1}^{N_T}\bar{\sigma}_j^2\right)} \tag{4.2}$$

for all $k \in \kappa$, and $d_j(k) = |v_{cej}|$, where v_{cej} is the jth element of $\mathbf{v}_{ce}(k) = \mathbf{v}_c(k) - \mathbf{v}_e(k)$.

Based on the explicit PEP result (4.1) obtained with imperfect CSI, we next present a new code design criterion for STTC over rapid fading channels with imperfect CSI.

4.2.1 Code Construction

For space-time trellis codes, the encoder maps binary data to modulation symbols, where the mapping function is described by a trellis diagram. As in [66], the encoder of a STTC is implemented as a set of feedforward shift registers. Let us consider an encoder of space-time trellis coded MPSK modulation with multiple transmit antennas as shown in Fig. 4.1. The input information sequence of a user, denoted by $\mathbf{u}$ is given by

$$\mathbf{u} = \left[\mathbf{u}^T(1)\ \mathbf{u}^T(2)\ \cdots\ \mathbf{u}^T(k)\ \cdots\right]^T$$

where $\mathbf{u}(k) = [u^1(k)\ \cdots\ u^m(k)]^T$ is a group of $m = \log_2 M$ information bits transmitted by the user in the kth frame. Thus, m binary input sequences $\mathbf{u}^1\ \cdots\ \mathbf{u}^m$

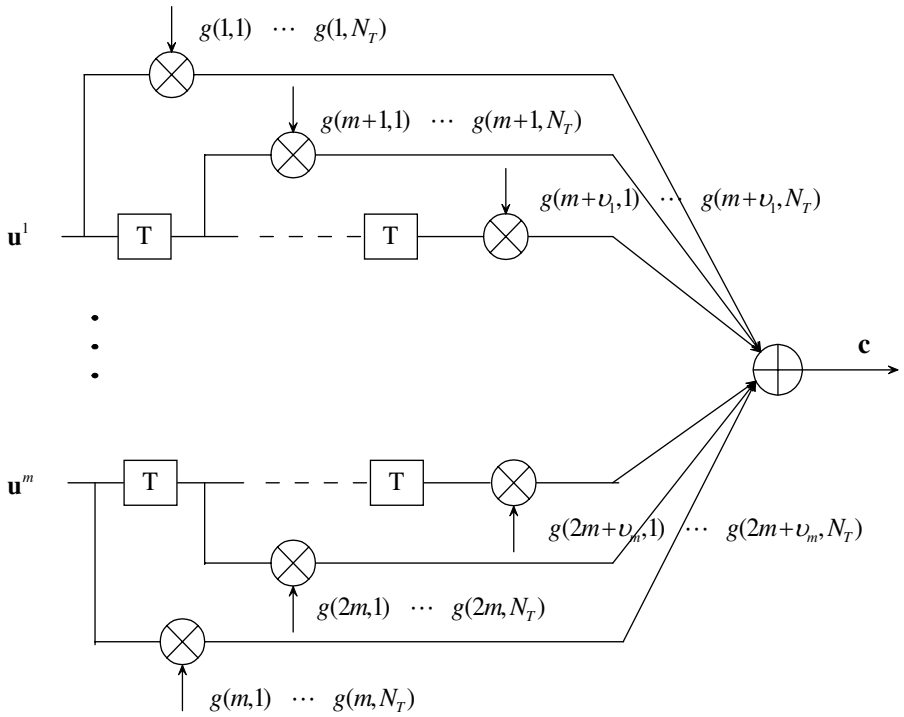


Fig. 4.1: Encoder for STTC.

are fed into the encoder. At the ith branch, $\mathbf{u}^i = [u^i(1) \ \cdots \ u^i(k) \ \cdots]^T$ is passed to the shift registers. The space-time trellis coded MPSK can achieve a bandwidth efficiency of m bits/s/Hz. The total memory order of the encoder, denoted by ν, is given by

$$\nu = \sum_{i=1}^{m} \nu_i \tag{4.3}$$

where ν_i, $i = 1, \cdots, m$ is the memory order for the ith encoder branch. The value of ν_i for MPSK constellations is determined by

$$\nu_i = \lfloor (\nu + i - 1)/m \rfloor \, . \tag{4.4}$$

The total number of states for the trellis encoder is 2^ν.

Since the memory orders for all branches, $\{\nu, \nu+1, \cdots, \nu+m-1\}$, are m consecutive numbers, there is only one number among them, denoted as $\nu + i_0 - 1$, that can be divided by m. For $i < i_0$, we have $\nu_i = \nu_1 = (\nu + i_0 - 1)/m - 1$, while for $i \geq i_0$, we have $\nu_i = \nu_2 = \nu_1 + 1$. If $i_0 = 1$, i.e., ν can be divided by m, we have $\nu_i = \nu_2$, for all i. Let $\mathbf{u}_k$ be the sub-sequence of $\mathbf{u}$, which affects the output code

symbols of the user concerned in the kth frame. It is given by

$$\mathbf{u}_k = \begin{cases} [u^1(k) \ \cdots \ u^m(k) \ \cdots \ u^{i_0}(k-\nu_2) \ \cdots \ u^m(k-\nu_2)]^T & i_0 \neq 1 \\ [u^1(k) \ \cdots \ u^m(k) \ \cdots \ u^1(k-\nu_2) \ \cdots \ u^m(k-\nu_2)]^T & i_0 = 1 \, . \end{cases}$$

The dimension of $\mathbf{u}_k$ is $P \times 1$, where $P = m\nu_2 + (m - i_0 + 1) = m + \nu$. These input bits are combined by the encoder, which can be described by using the generator matrix $\mathbf{G}_{P \times N_T}$. Each element g_{ij} of the generator matrix takes on a value in the set $\{0, 1, \cdots M-1\}$. The output code vector in the kth frame, $\mathbf{c}(k) = [c_1(k) \ \cdots \ c_{N_T}(k)]^T$, is obtained as the linear, modulo-M sum of the current and delayed binary inputs, and can be expressed as

$$\mathbf{c}(k) = \mathbf{G}^T \mathbf{u}_k \bmod M \, . \tag{4.5}$$

The encoder output symbols $\mathbf{c}(k)$ are modulated on to the transmitted code symbols $\mathbf{v}(k)$ via the mapping $v_i(k) = e^{j2\pi c_i(k)/M}$, for each transmit antenna i, as in [1].

4.2.2 The New Design Criterion

We will use the node error event probability as the performance measure in code design, instead of the PEP. The node error event probability $P(e)$ can be union bounded tightly at high SNR as in [65] by

$$P(e) = \sum_{\mathbf{v}_c} P(\mathbf{v} = \mathbf{v}_c) P(e|\mathbf{v} = \mathbf{v}_c) \leq \sum_{\mathbf{v}_c} P(\mathbf{v} = \mathbf{v}_c) \sum_{\mathbf{v}_e} P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) \tag{4.6}$$

where $P(e|\mathbf{v} = \mathbf{v}_c)$ is the conditional node error event probability given that the correct codeword sent is $\mathbf{v}_c$, and the PEP $P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c)$ is given by (4.1). Define Λ by

$$\Lambda = \max_{\mathbf{v}_c} \frac{1}{2} \sum_{\mathbf{v}_e} \prod_{k \in \kappa} \left[b\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right) \right]^{-N_R} \tag{4.7}$$

which is the maximum conditional node error probability. The union bound on the node error event probability on the right side of (4.6) can be further bounded as

$$P(e) \leq \Lambda \sum_{\mathbf{v}_c} P(\mathbf{v} = \mathbf{v}_c) = \Lambda \, . \tag{4.8}$$

From (4.8), the maximum conditional node error probability Λ, which is determined by the branch distances $b\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$'s, should be minimized to reduce the node error probability $P(e)$. Therefore, we arrive at the following design criterion.

Design Criterion: To minimize the node error event probability, one should minimize Λ, i.e., minimize the maximum sum of the products of the inverse of the branch distances, $\sum_{\mathbf{v}_e} \prod_{k \in \kappa} \left[b\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)\right]^{-1}$, over all possible codewords.

Here the kth branch distance $b\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$ is computed by (4.2) based on knowledge of the channel fading model and the channel estimator used. The above criterion is the first to incorporate the statistical information of the channel estimates in STTC design over rapid fading channels. The effect of channel estimation on code design is reflected by the branch distances $b\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$ through the variances $2\hat{\sigma}_j^2$'s of the estimated channel coefficients and the variances $2\bar{\sigma}_j^2$'s of the estimation errors, which can be obtained from (3.13) and (3.14), respectively. The kth branch distance $b(\mathbf{v}_c(k), \mathbf{v}_e(k))$ in (4.2) can be easily computed once the channel fading model and the channel estimator are given. With this explicit criterion, STTC design with imperfect CSI over rapid fading channels is easy to implement in practice. As will be seen, the loss caused by the channel estimation errors can be minimized by our proposed design criterion by exploiting the statistical information of the channel estimates.

4.2.3 The Optimally Distributed Euclidean Distances

According to (4.2), given the different channel estimation variances $2\hat{\sigma}_j^2$'s associated with each transmit antenna, the code symbol Euclidean distances $d_j(k)$'s should be optimally distributed among all transmit antennas to minimize the PEP. The general analytical study of the optimally distributed Euclidean distances for STTC is difficult, because the distribution depends on the structure of STTC, and the modulation scheme. Here, we examine the optimal distribution of the symbol-wise Euclidean distances for a particular example. It illustrates how to optimally

distribute the symbol-wise Euclidean distances when the statistical information of the channel estimates is available.

Consider two sets of codewords $\mathcal{C}_1$ and $\mathcal{C}_2$ with different code generator matrices. Assume the all-zeros bit sequence is transmitted. At high SNR, only the shortest error events are important. The vector-wise Euclidean distances of a shortest error event path for the codeword set $\{\mathcal{C}_i\}_{i=1}^2$ are given by $\{A_i^2(k)\}_{k\in\kappa_{min}}$, where κ_{min} is the frame index set of the shortest error event, and $A_i^2(k) = \sum_{j=1}^{N_T} d_j^2(k)$. Assume that both $\mathcal{C}_1$ and $\mathcal{C}_2$ have the *same* Euclidean distances on the segments of the shortest error event paths, i.e., $A_1^2(k) = A_2^2(k) = A^2(k)$ for all $k \in \kappa_{imin}$. Therefore, according to (4.1), the PEP upperbounds of $\mathcal{C}_1$ and $\mathcal{C}_2$ are approximately the same at high SNR, if perfect CSI is available to the receiver. However, for the imperfect CSI case, the performance of the codes may be different if we change the distribution of the symbol Euclidean distances $d_j^2(k)$'s among all the transmit antennas. Thus, the PEP optimization problem can be outlined as

1. maximize $D(\hat{\sigma}) = \prod_{k\in\kappa} \sum_{j=1}^{N_T} \hat{\sigma}_j^2 d_j^2(k)$
2. subject to
   (a) $\{d_j(k)\}_{k\in\kappa} \subset \Delta$ where Δ is the set of all the possible Euclidean distances that can be assumed in the MPSK constellation, and each element Δ_i in Δ is given by $\Delta_i = 2\sin(i\pi/M)$, $i = 1, \cdots, M/2$, and
   (b) $\sum_{j=1}^{N_T} d_j^2(k) = A^2(k)$

To maximize $D(\hat{\sigma})$, we note first that

$$\sum_{j=1}^{N_T} \hat{\sigma}_j^2 d_j^2(k) \le \sum_{j=1}^{N_T} \hat{\sigma}_{max}^2 d_j^2(k) \le \hat{\sigma}_{max}^2 A^2(k) \tag{4.9}$$

where $\hat{\sigma}_{max} = \max_j \hat{\sigma}_j$. The antenna with the largest estimate variance $2\hat{\sigma}_{max}^2$ has the highest channel estimation accuracy. We denote this antenna as the j_mth transmit antenna. The equality sign in (4.9) holds, i.e., the maximum of $\sum_{j=1}^{N_T} \hat{\sigma}_j^2 d_j^2(k)$

is attained only when

$$d_j(k) = \begin{cases} 0 & j \neq j_m \\ A^2(k) & j = j_m \, . \end{cases} \tag{4.10}$$

Therefore, the PEP attains its minimum when the vector-wise Euclidean distance $A^2(k)$ is all assigned to the j_mth transmit antenna at each transmission time. However, this is only suitable for $A^2(k) \leq 2$, since 2 is the maximum symbol-wise Euclidean distance that can be achieved by the normalized MPSK constellation. For general STTC systems with imperfect CSI, the symbol-wise Euclidean distance of the transmit antenna with higher estimation accuracy, should be larger.

4.2.4 The Effect of Channel Estimation on Code Design

As mentioned, it is the *differences* of the channel estimate variances that lead to the importance of channel estimation information in code design with imperfect CSI. It is intuitively clear that if the differences between the variances of the channel estimates are larger, the effect of channel estimation on code design will be greater. Thus, we measure the effect of channel estimation on code design by using the *maximum estimation variance difference*. The maximum estimation variance difference $\triangle\hat{\sigma}^2_{max}$ is the difference between the maximum channel estimate variance $\hat{\sigma}^2_{max}$ and the minimum channel estimate variance $\hat{\sigma}^2_{min}$ among all the transmit antennas, i.e., $\triangle\hat{\sigma}^2_{max} = \hat{\sigma}^2_{max} - \hat{\sigma}^2_{min}$, where $\hat{\sigma}^2_{max} = \max_j \hat{\sigma}^2_j$ and $\hat{\sigma}^2_{min} = \min_j \hat{\sigma}^2_j$ for all $j = 1, \cdots, N_T$.

Here, we focus on the effect of the channel fade rate f_dT on the maximum estimation variance difference $\triangle\hat{\sigma}^2_{max}$ by fixing L, N with their robust values. The robust values are used because choosing L and N adaptively on-line greatly increases the complexity of the receiver. Also, the measurement of channel parameters such as fade rate requires additional overhead, and cannot be done frequently, in most cases. The robust value is defined in a statistical sense, which indicates that this value can work well with high probability when the channel fade rate is variable.

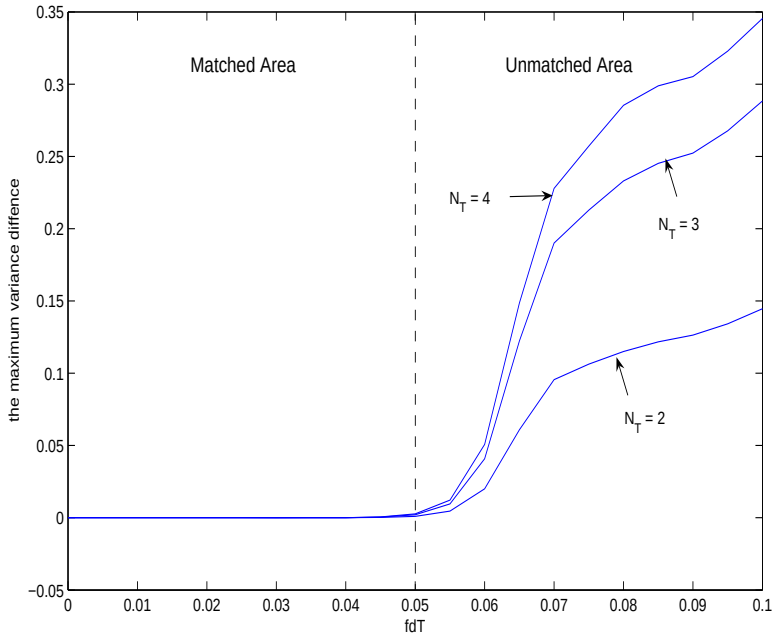


Fig. 4.2: The maximum estimation variance difference as a function of the channel fade rate f_dT for different numbers of transmit antennas, with fixed parameters $L = 8$ and $N = 6$ at $E_p/N_0 = 30$ dB.

Due to the physical limit on the speeds of mobiles and the symbol-wise constant channel model, the channel fade rate caused by this mobility can be assumed to have a maximum value of, say, $f_{d,\,max}T$. In practice, the channel fade rate f_dT can be assumed less than a certain threshold value f_cT with a high probability. From our work in chapter 3, the optimum values of L and N for a high fade rate can work for the lower fade rates, while the converse is not true. Thus, we take the optimum values of L and N at the fade rate of $f_dT = f_cT$ as the robust values. These values can work well with high probability for fade rates that are less than the threshold value f_cT.

For the purpose of illustration, we assume here the threshold channel fade rate is $f_cT = 0.05$, and the maximum channel fade rate is $f_{d,\,max}T = 0.1$. The robust values for the pilot spacing and the Wiener filter length are $L = 8$ and $N = 6$, respectively, which are determined at the threshold channel fade rate of $f_cT = 0.05$, as shown in the previous chapter. In Fig. 4.2, we show the maximum

estimation variance difference $\triangle\hat{\sigma}^2_{max}$ as a function of f_dT, for different numbers of transmit antennas. A high pilot SNR of $E_p/N_0 = 30$ dB is used. The values of the channel estimate variances become steady with the increase of SNR. Thus, the dependence of the channel estimate variances on SNR can be dropped at high SNR. Fig. 4.2 shows that the steady maximum estimation variance difference $\triangle\hat{\sigma}^2_{max}$ will increase with the fade rate f_dT. Also, $\triangle\hat{\sigma}^2_{max}$ increases with N_T, but the increase in $\triangle\hat{\sigma}^2_{max}$ is smaller for larger values of N_T. Since robust values of L and N are chosen according to the threshold channel fade rate of f_cT, the channel estimation accuracy is only guaranteed for those fade rates that are less than f_cT. Thus, Fig. 4.2 can be seen to have two parts. One is the *matched* region ($0 \leq f_dT \leq f_cT$), where accurate channel estimation can be obtained; the other is the *unmatched* region ($f_cT < f_dT \leq f_{d,\,max}T$), where the channel estimation accuracy is poor. The maximum estimation variance difference $\triangle\hat{\sigma}^2_{max}$ is very small in the matched region, while it becomes much larger in the unmatched region. With the larger $\triangle\hat{\sigma}^2_{max}$, there is greater effect of channel estimation on code design. It is expected that codes designed with imperfect CSI can achieve greater performance gains for high fade rates or a large number of transmit antennas, compared with those designed assuming perfect CSI. This will be verified in our numerical results later. Based on the knowledge of the channel fade rate at the transmitter, we will consider two channel conditions for code design in the following sections.

Table 4.1: The proposed code generator matrices G^T for perfect and imperfect CSI, and the known generator matrices in the literature using QPSK modulation scheme.

N_T		Codes	State No.	QPSK
2	New	ICSI $f_dT = 0.05$	8	[2 0 3 0 2; 2 2 2 1 3]
			16	[2 1 0 2 3 2; 1 2 2 0 2 3]
			32	[1 2 2 0 3 0 2; 2 3 2 2 2 3 1]
		ICSI $f_dT = 0.1$	8	[2 1 3 1 2; 2 1 2 3 3]
		PCSI	8	[2 0 1 2 2; 2 2 2 1 0]
			16	[2 0 2 3 1 2; 1 2 0 2 2 3]
			32	[2 3 2 2 2 3 1; 3 2 2 0 1 0 2]
	Known	TSC	8	[0 0 2 1 2; 2 1 0 0 2]
			16	[0 0 1 2 2 0; 1 2 2 0 0 2]
			32	[2 0 1 2 1 2 0; 3 2 2 2 0 1 2]
		FVY	8	[0 2 1 1 2; 2 2 3 2 0]
			16	[0 2 0 1 2 0; 2 0 1 2 2 2]
			32	[0 1 2 1 3 2 1; 1 3 0 2 0 1 2]
		ZQWL	8	[2 1 0 2 1; 1 2 2 1 1]
			16	[0 0 1 2 2 0; 1 2 2 0 0 2]
			32	[0 0 1 2 2 3 2; 1 2 1 2 0 3 2]
3	New	ICSI $f_dT = 0.05$	8	[2 0 3 0 2; 2 2 2 1 3; 2 1 2 1 1]
		ICSI $f_dT = 0.1$	8	[2 0 3 0 2; 2 2 2 1 3; 2 0 2 3 3]
		PCSI	8	[2 0 1 2 2; 2 2 2 1 0; 2 3 2 0 3]

4.2.5 Code Design for Known Fade Rates

When the channel fade rate f_dT is known to the transmitter, the variances $\hat{\sigma}_j^2$'s of the channel estimates, which are functions of f_dT, can be easily computed using (3.13). With the variances of the channel estimates $\hat{\sigma}_j^2$'s calculated for a given fade rate, the optimum codes for that fade rate with channel estimation can be obtained

using the proposed design criterion. We term these codes *the ICSI codes* for the given fade rate. In the limit of perfect channel estimation, *the PCSI codes* are obtained. While we will present in section IV an iterative search algorithm for the design of codes using our design criterion, we consider here first the performance of some codes we obtained. In Tab. 4.1, we present the ICSI codes for $f_dT = 0.05$ and $f_dT = 0.1$ and the PCSI codes, with QPSK modulation and two or three transmit antennas. The ICSI codes for $f_dT = 0.05$ and the PCSI codes with 8PSK modulation are listed in Tab. 4.2

Table 4.2: The proposed code generator matrices G^T for perfect and imperfect CSI, and the known generator matrices in the literature using 8PSK modulation scheme.

N_T		Codes	State No.	8PSK
2	New	ICSI $f_dT = 0.05$	8	[4 4 2 2 4 3; 4 0 2 0 1 4]
			16	[3 4 4 2 4 3 2; 1 4 2 4 0 2 1]
			32	[3 2 4 4 4 2 4 2; 4 0 3 4 0 2 1 1]
		PCSI	8	[4 2 4 2 4 3; 4 2 0 0 1 4]
			16	[2 4 4 4 3 4 2; 4 1 4 0 4 2 2]
			32	[4 4 3 4 2 4 2 4; 4 2 0 3 2 0 4 1]
	Known	TSC	8	[0 0 0 4 2 5; 4 2 1 0 0 0]
			16	[0 0 0 4 2 5 1; 4 2 1 4 2 1 5]
			32	[0 0 0 4 2 5 2 3; 4 2 1 4 2 1 2 7]
		FVY	8	[0 2 4 2 4 2; 1 4 6 4 0 1]
			16	[0 1 2 4 2 6 0; 4 5 1 2 1 4 5]
			32	[3 1 1 0 0 3 6 1; 4 1 2 4 6 1 1 0]

To show the effect of channel estimation on code design, we compare the performances of our proposed ICSI codes with some PCSI codes available in the literature over rapid fading channels with channel estimation. Fig. 4.3 shows the simulated BEP of STTC with $N_T = 2$ transmit antennas and $N_R = 1$ receive antenna at the channel fade rate of $f_dT = 0.05$. On each link, the fading process has a Jake's

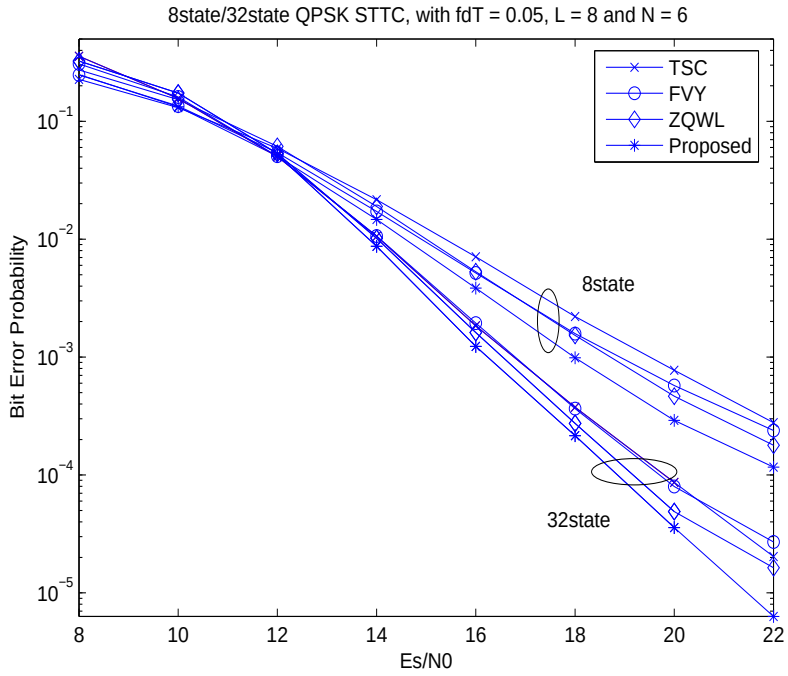
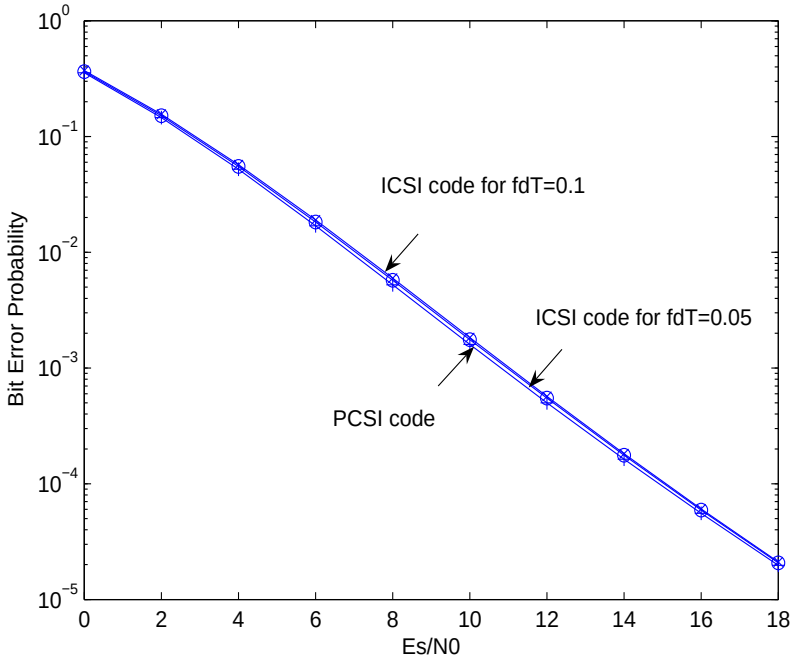
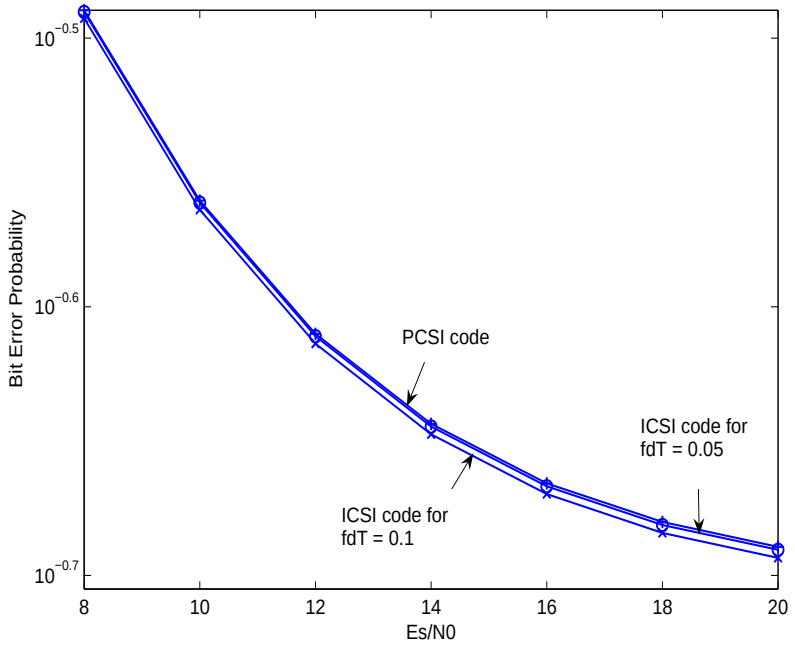


Fig. 4.3: Comparison of simulated BEP comparison of STTCs with two transmit and one receive antenna, where the channel fade rate is $f_dT = 0.05$, which is estimated with $L = 8$, and $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p = (E_s)_{PSAM}$.

power spectrum. The fading coefficients are estimated using the PSAM scheme with the pilot spacing $L = 8$, and the Wiener filter length $N = 6$. It can be seen that the newly obtained 8-state/32-state QPSK, ICSI codes for $f_dT = 0.05$ outperform those STTC of [1], [2], [9] (abbreviated as TSC, FVY, and ZQWL codes, respectively, after the authors' names) in the presence of imperfect CSI with channel estimation errors. Our codes have better performance, especially at high SNR. The performance gain achieved by our 8-state ICSI code over the ZQWL code is about 1 dB. In fact, the modulation scheme, the number of transmit antennas, and the channel conditions affect the performance gain achieved by our proposed ICSI codes. The performance gain decreases as the number of code states increases, but increases as the number of modulation levels increases. The larger the maximum estimation variance difference, the greater the performance gains that can be achieved. Furthermore, code design with imperfect CSI is based on the variances of the channel estimates. Thus, the optimal ICSI codes are those that have design pa-



(a) Perfect CSI is available to the receiver

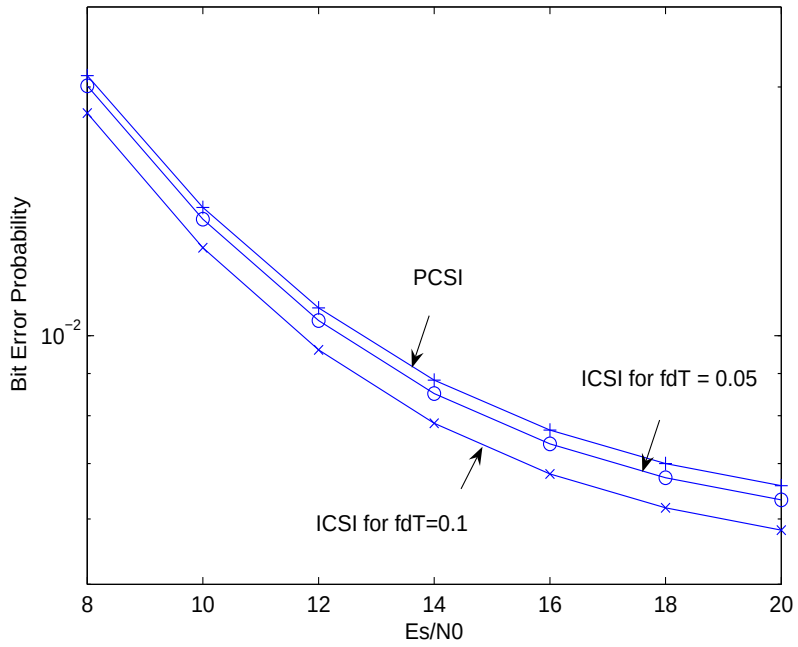


(b) The channel fade rate is $f_dT = 0.1$

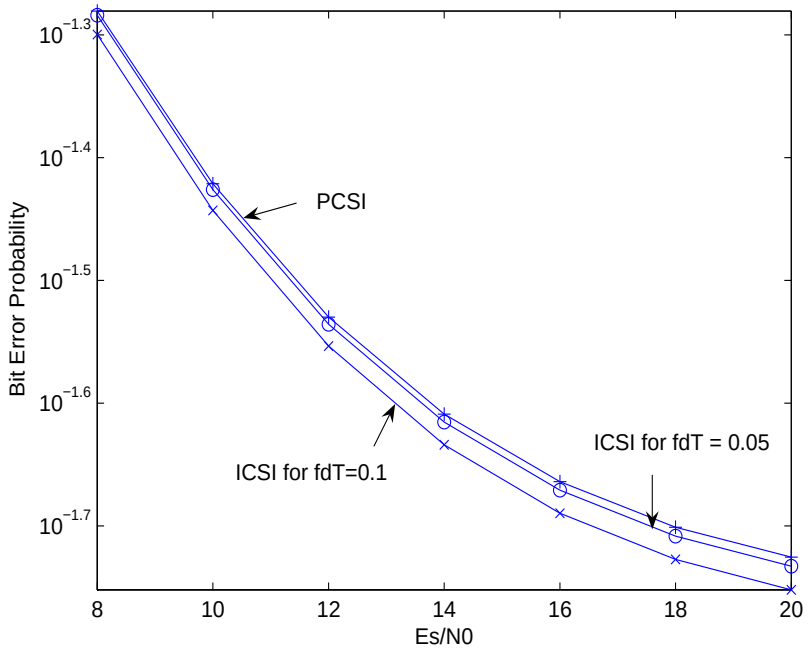
Fig. 4.4: The BEP performance comparison of the 8-state QPSK STTCs using two transmit and one receive antenna under different channel situations. For the imperfect CSI, the channel gains are estimated with $L = 8$, and $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p/N_0 = 30$ dB.

rameters matched to the channel conditions. In Fig. 4.4, we demonstrate adaptive coding is needed when the maximum estimation variance difference is sufficiently high. It can be seen from Fig. 4.4(a) that the PCSI code has the best performance when perfect CSI is available to the receiver. However, when the channel situation changes to a high fade rate of $f_dT = 0.1$ and only ICSI is available, the ICSI code for $f_dT = 0.1$ is best as shown in Fig. 4.4(b). The high BEP in Fig. 4.4(b) is due to the fact of the large mean square estimation errors since the high fade rate of $f_dT = 0.1$ falls into the unmatched region defined in Fig. 4.4(b). We choose this unmatched case of $f_dT = 0.1$ because the maximum estimation variance is large for this case, which can satisfy the condition for adaptive code design. The performance improvement shown in Fig. 4.4 using adaptive coding is not very large. This is reasonable for i.i.d. channels, where the maximum estimation variance difference is not so large because the statistics of the different links are the same, and, thus, the effect of channel estimation on code design is not obvious. However, it is an indication of the usefulness of adaptive code design. As the variances of the channel estimates change with the channel conditions, different optimum codes should be used, if the effect of channel estimation on code design is sufficiently large. For non-identical channels, the maximum estimation variance difference is usually large when the imbalance among the links is high. Thus, for non-identical channels, the performance gain achieved by the optimal ICSI codes matched to the channel conditions is great. The details of code design for non-identical channels with imperfect CSI will be discussed in Chapter 5.

It is also worth noting that the performance gains achieved by taking into account the statistical information of the channel estimates may not be obvious when the maximum estimation variance difference is small. In our investigation, we found that if the maximum estimation variance difference falls in the range of $0 \leq \triangle\hat{\sigma}^2_{max} \leq 0.15$, the ICSI codes obtained for different fade rates have similar performances. However, when $\triangle\hat{\sigma}^2_{max}$ becomes larger than 0.15, significant performance gains can be obtained especially for high SNR. In Fig. 4.5, two situations with greater $\triangle\hat{\sigma}^2_{max}$ are assumed. One has $\triangle\hat{\sigma}^2_{max} = 0.2$ with $\hat{\sigma}^2_1 = 0.2$, and



(a) $\hat{\sigma}_1^2 = 0.3, \hat{\sigma}_2^2 = 0.7$



(b) $\hat{\sigma}_1^2 = 0.2, \hat{\sigma}_2^2 = 0.4$

Fig. 4.5: The BEP performance of the 8-state QPSK STTCs using two transmit and one receive antenna under different channel situations. For the imperfect CSI case, channel estimation variances and the variance of the channel estimation errors are fixed to $\sum_{j=1}^{N_T} \bar{\sigma}_j^2 = 0.1$.

$\hat{\sigma}_2^2 = 0.4$. The other has $\triangle\hat{\sigma}_{max}^2 = 0.4$ with $\hat{\sigma}_1^2 = 0.3$, and $\hat{\sigma}_2^2 = 0.7$. As shown in Section 3.6, the variances of the channel estimates $2\hat{\sigma}_j^2$ for $j = 1, \cdots, N_T$ depend on SNR, the vales taken by the PSAM parameters L and N, and the channel fade rate. For i.i.d. channels, the large value for $\triangle\hat{\sigma}_{max}^2$ can be obtained for high channel fade rates with a large number of transmit antennas, such as in the unmatched region shown in the Fig. 4.2. Here, we assume these large values of $\triangle\hat{\sigma}_{max}^2$ to show the effect of channel estimation on code design more clearly. For both cases, the total estimation error variance is assumed to be $\sum_{j=1}^{N_T} \bar{\sigma}_j^2 = 0.1$. The results in Fig. 4.5 clearly show that the performance gains achieved by the ICSI code for $f_dT = 0.1$ is much better than the ICSI code for $f_dT = 0.05$ and the PCSI code. This is because when $f_dT = 0.1$, the maximum estimation variance difference is the largest among the three codes, which is close to the assumed channel situation. Therefore, the ICSI code for $f_dT = 0.1$ performs best and achieves significant performance gains. Comparing Fig. 4.5(a) and Fig. 4.5(b), we can see that the performance gain increases with the maximum estimation variance difference.

The above results indicate that adaptive code design based on knowledge of the channel fade rate should be used to achieve the optimal performance. Since the channel fade rates fall in the range of $[0, f_{d,max}T]$, the range can be divided into Q segments. Let $\triangle_{max}$ be the maximum estimation variance difference at the fade rate of $f_{d,max}T$, $Q = \triangle_{max}/0.15$ and $\triangle f = f_{d,max}/Q$. For the fade rates falling into a segment $[(i-1)\triangle f,\ i\triangle f]$ where $i = 1, \cdots, Q$, the differences among the associated $\triangle\hat{\sigma}_{max}^2$'s are less than 0.15, and the performances of the ICSI codes designed for each fade rate are similar. Thus, the same code can be used for all fade rates in each segment for efficiency. However, if the fade rates vary over different segments, adaptive coding should be employed by assigning the corresponding optimum codes for different segments.

4.2.6 Robust Code Design for Unknown Fade Rates

In wireless communication, the knowledge of channel fade rate obtained at the transmitter may be outdated or imperfect because of the time-variance of the channel fade rates and the feedback errors. This motivates us to look for robust codes which can perform well when the channel fade rate is unknown to the transmitter. The robust codes are to be understood from a statistical viewpoint. That means although the robust codes may perform worse than the code for a certain channel fade rate, the robust codes do have the best average performance due to the variance of the channel fade rates. The average performance is obtained by averaging the conditional performance at each fade rate over the probability distribution of the channel fade rates. For the purpose of illustration, we assume that there is a finite number of discrete fade rates. We assume here that a probability model for the channel fade rate in which the probability of each possible channel fade rate is known to the transmitter and the receiver. The set of all the possible values for the channel fade rate is denoted as $\chi = \{x_i\}_{i=1}^{q}$, where q is the cardinality of χ. The probability of the channel fade rate taking the value of x_i is denoted as $P(x_i) = P(f_dT = x_i)$. To achieve the optimum average performance, the robust code can be designed based on *the average variance vector* $\overline{\mathbf{\Omega}}$ of the channel estimates. Given a channel fade rate, the conditional variance of the channel estimate associated with the jth transmit antenna $\hat{\sigma}_j^2$ can be calculated by (3.13). We define the conditional variance vector $\mathbf{\Omega}|_{f_dT=x}$ of the channel estimates at the fade rate of $f_dT = x$ as $\mathbf{\Omega}|_{f_dT=x} = [\hat{\sigma}_1^2 \cdots \hat{\sigma}_{N_T}^2]^T|_{f_dT=x}$. The average variance vector $\overline{\mathbf{\Omega}}$ of the channel estimates can be computed by averaging $\mathbf{\Omega}|_{f_dT=x}$ over $P(x)$. The elements $\overline{\Omega}_j$'s ($j = 1, \cdots, N_T$) in $\overline{\mathbf{\Omega}}$ can be viewed as the *equivalent estimation variances*. Based on the proposed code design criterion, robust codes can be obtained, using the calculated average variance vector $\overline{\mathbf{\Omega}}$ as the statistical information of the channel estimates, for the imperfect CSI case and the fade rates are unknown to the transmitter.

In general, robust codes are designed based on the average variance vector $\overline{\mathbf{\Omega}}$.

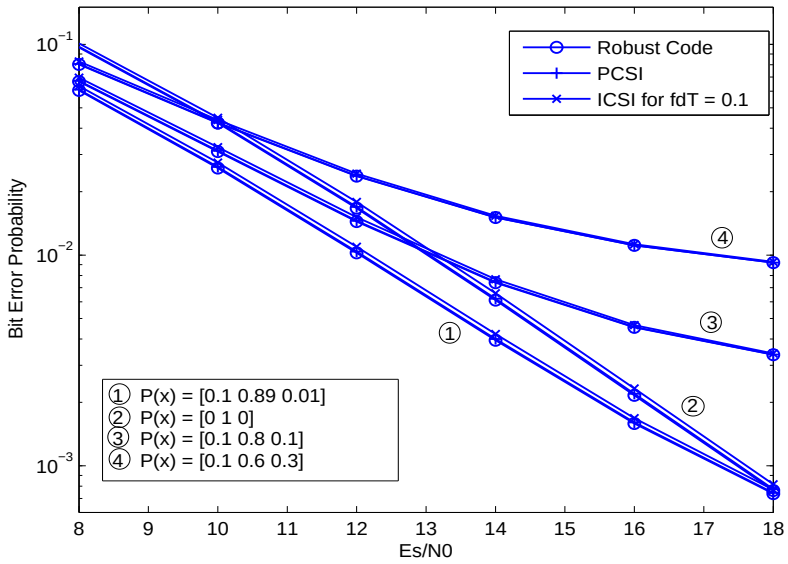


Fig. 4.6: The performance analysis of the QPSK 8-state PCSI, ICSI and robust code over the channel with time-variant fade rates, where $N_T = 2$, $N_R = 1$, $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, $E_p = (E_s)_{PSAM}$ *and* $P(x)$ *is the assumed probability distribution of the channel fade rates.*

However, this can be simplified, if the channel has a *dominant* fade rate, which occurs with a probability much higher than the sum of the probabilities of other fade rates. In this situation, the average variance vector can be approximated by the variance vector of the dominant fade rate, because the contribution from the variances of the channel fade rate with the highest probability is the largest. Thus, the variance vector of the dominant fade rate can be used to design codes instead of the average variance vector. In other words, the robust code becomes the same as the code of the dominant fade rate. In Fig. 4.6, we show the performances of the 8-state QPSK PCSI code, the ICSI code for $f_dT = 0.1$, and the robust code over channels with time-varying fade rates, when only the distribution of the channel fade rates is available at the transmitter. We assume that the channel fade rate has three possible states, namely $f_dT = 0.01$, $f_dT = 0.05$ and $f_dT = 0.1$, respectively. For the four different distributions assumed in Fig. 4.6, the dominant fade rate is $f_dT = 0.05$. The robust code is thus the ICSI code for the fade rate of $f_dT = 0.05$. For all the four time-variant channel conditions, the robust code performs best, and

the PCSI code is a bit worse than the robust code because of the small maximum estimation variance difference. The ICSI code for $f_dT = 0.1$ is the worst. Although the ICSI code for $f_dT = 0.1$ is the best when the channel fade rate is $f_dT = 0.1$, it performs no better than the robust code because of the small probability of the fade rate f_dT being equal to 0.1 that we assumed. It is found that with the increase of the probability of $f_dT = 0.1$, the performance gap between the ICSI code for $f_dT = 0.1$ and the robust code is reduced. From Fig. 4.4 and Fig. 4.6, we conclude that the robust code is a good tradeoff between performance and complexity of adaptive coding for unknown fade rate channels.

4.3 Iterative Code Search Algorithm

As is well known, the complete code search is highly complex because it requires a search over $M^{P\times N_T}$ matrices over all possible code sequences. The number of possible code sequences is exponentially increasing with the sequence length. Thus, it is important to keep the code search complexity low. To do this, we rewrite the union bound on the error event probability (4.6) as

$$P(e) \leq \sum_{\mathbf{v}_c} P(\mathbf{v}=\mathbf{v}_c) \sum_{L_p=L_{pmin}}^{\infty} \sum_{\mathbf{v}_e\in\mathcal{V}_{L_p}} P(\mathbf{v}_c \rightarrow \mathbf{v}_e|\mathbf{v}=\mathbf{v}_c)\,. \tag{4.11}$$

by categorizing the error events according to their lengths. Here, L_p is the error event length, which varies from the shortest error event length L_{pmin} to infinity. $\mathcal{V}_{L_p}$ is the set of the error events whose error event lengths are equal to L_p. Without considering all the error events at one time, we focus on the error events with lengths up to a certain value L_{pmax}, where $L_{pmin} \leq L_{pmax} < \infty$. Since for a small L_{pmax}, the number of possible error events is quite small, we can first obtain sub-optimal codes with relatively low complexity. Then, the sub-optimal codes can be refined by increasing L_{pmax} step by step. The final optimal codes can be obtained via an iterative search. This iterative search algorithm can guarantee the optimality of the codes for high SNR, since the error events with the shortest error event lengths

determine the PEP for high SNR. Also, it can achieve good performance at low SNR by increasing the error event length step by step.

Suppose the generator matrix set to be searched is Θ. We assume initially Θ includes all the $M^{P\times N_T}$ possible generator matrices. At the first iteration, our algorithm searches over Θ to choose those matrices that have the minimum Λ given by (4.7) over the shortest error events, by setting $L_{pmax} = L_{pmin}$. The matrices chosen form a new set Θ_n. The old set Θ is replaced by Θ_n. Increasing L_{pmax} by one and repeating the search process, the set Θ can be further refined. The iterations continue until there is no more change to the set Θ. This way, we finally obtain the optimal generator set G_{opt}. In addition to using the iterative search algorithm, the complexity of the search can be further reduced by using some rules to discard unsatisfactory codes efficiently at the initial step. First, the one-to-one code mapping property should be guaranteed, i.e., different input bits at a state produce different output code bits. Second, detection ambiguity should be avoided by ensuring that the code labels converging to the same state are different. Combined with the criterion presented, the final search algorithm can be summarized as follows.

Search Algorithm

1. Initially, the maximum error event length is set to $L^1_{pmax} = L_{pmin}$. Search for generator matrices which have minimum Λ given by (4.7) over the set Θ. Those matrices found will comprise the set Θ_n. Update Θ by letting $\Theta = \Theta_n$.

2. During the search, there are some generator matrices discarded directly without computing Λ because they do not satisfy the following rules:

   (a) All code labels diverging from the same state should be different.

   (b) The code labels converging to the same state should be different.

3. At the lth step, we set $L^l_{pmax} = L^{l-1}_{pmax} + 1$ and search the current set Θ_n by looking for matrices that minimize Λ. Then, update Θ by letting $\Theta = \Theta_n$.

4. Repeat step 3) until there are no more changes to the generator matrix set Θ, which is then the final optimal generator set G_{opt}.

Actually, the initial size of the set Θ can be reduced by using a row-wise search idea proposed in [70], which is used in the code design for 8PSK because of the high search complexity. The proposed new codes (in Tab. 4.1) are chosen from the final optimal generator set G_{opt}. Most of the optimum codes obtained from our new search algorithm have better BEP performance under actual channel estimation conditions than the existing ones in the literature that are designed under the perfect CSI assumption. The superiority of the optimal codes obtained from our search is more obvious at high SNR. It is because the iterative algorithm favors the generator matrices whose performances are better when the error events have small error event lengths. The error events with small lengths dominate the performance at high SNR. The performance gains achieved by our proposed codes increase with the channel fade rate and the number of transmit antennas. This will be shown in the next section.

4.4 Code Search Results and Performances

Based on the code design criterion and the iterative code search algorithm, new STTC for both perfect and imperfect CSI with two or three transmit antennas are obtained, as shown in Tab. 4.1. As mentioned, they are termed ICSI and PCSI codes, respectively. There are two kinds of ICSI codes. One is the ICSI code for a given channel fade rate. The other is the robust ICSI code. Under simplified situations, the robust code can be reduced to the ICSI code corresponding to the dominant fade rate. Since the perfect CSI case is too idealistic, the codes that have optimal performance under real channel estimation conditions have more practical importance. In this section, numerical results of the proposed codes are presented. In the simulations, we assume $\sigma^2 = 0.5$, and each channel fading process has a

Jake's power spectrum so that $R(\tau) = J_0(2\pi f_d \tau)$. We take the user whose time slot is located immediately after the pilot phase in a block as the user of interest. The total number of frames is $K = 100$. We account for the energy spent on the pilot symbols by taking the effective energy per data symbol in the PSAM scheme as $(E_s)_{PSAM} = E_s(L - N_T)/L$, where E_s is the data symbol energy for the case of PCSI when no pilots are transmitted. Two situations for the pilot symbol energy are considered. In one case, the pilot symbol energy is assumed to be equal to the data symbol energy, i.e., $E_p = (E_s)_{PSAM}$. In the other case, a high pilot symbol energy is used to obtain steady variances of the channel estimates, where we assume $E_p/N_0 = 30$ dB. The robust values of the pilot spacing and the Wiener filter length are $L = 8$ and $N = 6$, respectively.

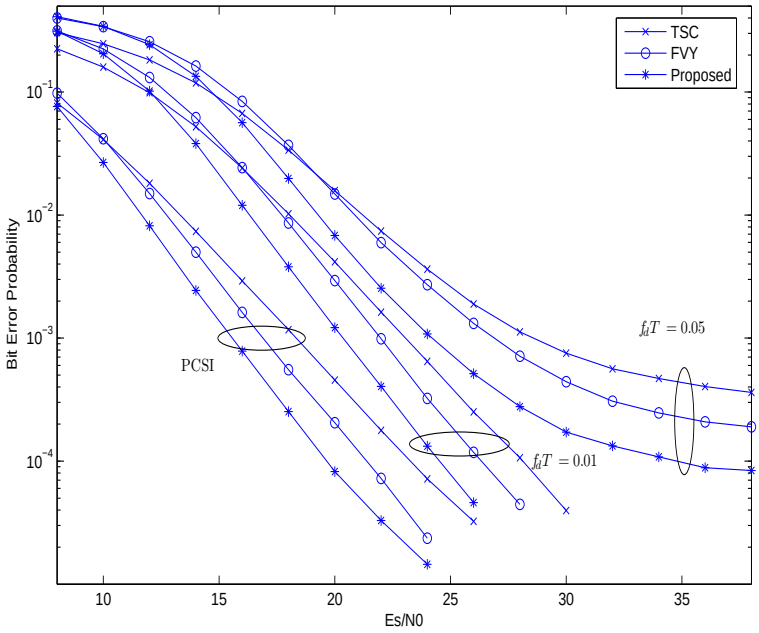


Fig. 4.7: The performance gains of the proposed 8-state 8PSK ICSI code compared with the FVY in [2], and the TSC in [1] over different channel fade rates, with two transmit and one receive antenna. For imperfect CSI, the channel is estimated using $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p = (E_s)_{PSAM}$.

Fig. 4.7 verifies that more performance gains can be achieved by our ICSI codes with an increase in the channel fade rate. Three codes, namely the FVY code, the TSC code, and our ICSI code are examined under different channel conditions.

Both perfect CSI and imperfect CSI channel conditions are considered. For the perfect CSI case, the channel fading coefficients are independent from symbol to symbol, and known to the receiver. For the imperfect CSI case, the channel fade rates considered are $f_dT = 0.01$ and $f_dT = 0.05$, respectively. The channel fading coefficients are estimated using the PSAM scheme. The performance gain of our ICSI code over the FVY code is great. For instance, a 1.5 dB gain is achieved under the perfect CSI condition at a BEP of $10^{-3.5}$; in the imperfect CSI condition, a 2.6 dB gain is obtained for $f_dT = 0.01$, and a gain of up to 4.3 dB for $f_dT = 0.05$, at the same BEP. Thus, the higher the channel fade rate, the greater the performance gain achieved by our codes designed under actual channel estimation conditions, by incorporating the statistical information of the channel estimates in the code design.

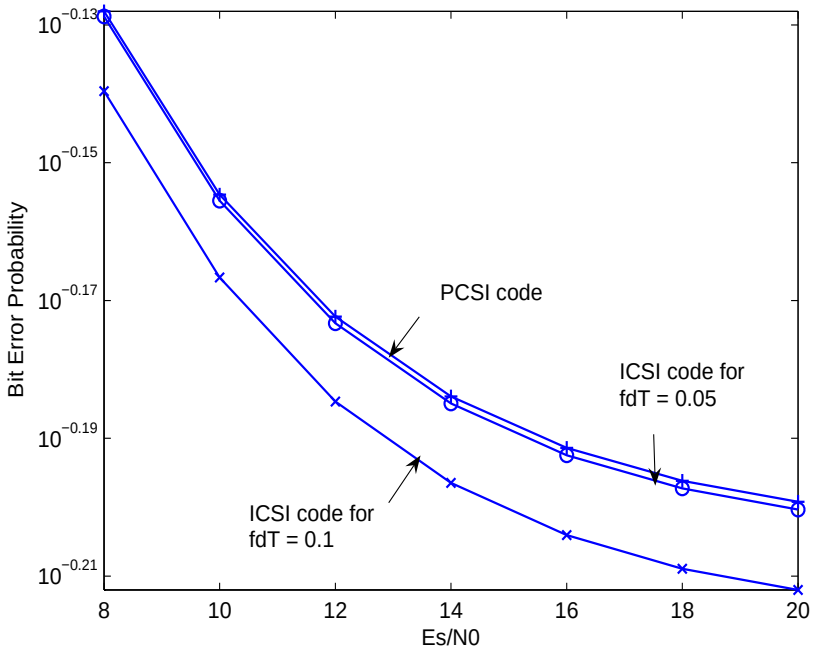


Fig. 4.8: The BEP performance of the proposed 8-state QPSK PCSI code and the ICSI codes under channel estimation, using three transmit and one receive antenna. The channel fade rate is $f_dT = 0.1$, which is estimated with $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p/N_0 = 30$ dB.

Not limiting the importance of using the statistical information of channel estimates to high fade rate conditions, we also found that with an increase in the

transmit antenna number, more performance improvements could be achieved by using ICSI codes obtained using our code design criterion. Due to the complexity of the code search for larger numbers of transmit antennas, we give only one example of the codes designed with $N_T = 3$. These codes are obtained using the sub-optimal search method in [3]. Based on the code generators obtained for $N_T = 2$, we only search for the remaining column corresponding to the third transmit antenna of the generator matrix $\mathbf{G}$. It can be seen in Fig. 4.8 that the performance gains achieved by the ICSI codes compared to the PCSI code is greater in the case when $N_T = 3$ than in the case when $N_T = 2$, which was shown in Fig. 4.4(b). The greater improvement is due to the fact the maximum estimation variance difference increases with the number of transmit antennas, as shown in Fig. 4.2. The BEP in Fig. 4.8 is high because the high fade rate of $f_dT = 0.1$ falls in the unmatched region in Fig. 4.2. This unmatached scenario is used to produce the maximum possible $\triangle\hat{\sigma}^2_{max}$, which can lead to the maximum possible performance gain. Our results demonstrate that the performance gain achieved by adaptive coding increases as the number of transmitter antennas increases, as well as when the channel fade rate increases.

From the above results, we conclude that greater performance gains can be achieved by exploiting the statistical information of the channel estimates when the number of transmit antennas is large, or when the fade rate is high. These results also verify the analysis of the effect of channel estimation on code design. With an increase in the maximum estimation variance difference, there are greater performance improvements that can be obtained by using the proposed design criterion incorporating channel estimation. Our results here can be straightforwardly extended to the case of independent, non-identical Rayleigh fading channels. In that case, the incorporation of the statistical information of the channel estimates is even more important, as will be shown in the next chapter.

4.5 Summary

In this chapter, we presented a new, practical code design criterion for STTC over rapid, nonselective, Rayleigh fading channels with imperfect CSI. The codes designed via this criterion achieve performance gains over those designed under the perfect CSI assumption, when the receiver operates with imperfect CSI. We optimally distribute the symbol-wise Euclidean distances between codewords to minimize the error event probability under actual channel estimation conditions, by taking into account the statistical information of the channel estimates. The effect of channel estimation on code design can be measured by the maximum estimation variance difference, which, in turn, depends on the channel fade rate, the values taken on by the PSAM parameters, and SNR. The larger the maximum estimation variance difference, the more the performance improvement that can be achieved by using our code design criterion. When the channel fade rate is known to the transmitter at each time, the codes whose design parameters match those of channels and estimators perform best. Thus, adaptive coding can be used to achieve the optimal performance, when the maximum variance is sufficiently large. However, for i.i.d. channels, the improvement achieved by adaptive coding is not much because the statistics of the different links are nearly the same. Therefore, the maximum estimation variance difference is usually small, and the effect of channel estimation on code design is not so much. When only the knowledge of the distribution of the channel fade rates is known, robust codes can be obtained based on the average variance vector of the channel estimates. An iterative code search algorithm is developed to facilitate the search for new codes. This algorithm reduces the code search complexity by iteratively increasing the maximum error event length considered, one step at a time. Our investigation shows that with increasing channel fade rate or with a larger number of transmit antennas, greater performance gains can be achieved by exploiting the statistical information of the channel estimates.

Chapter 5

STTC over Non-identically Distributed Channels with Channel Estimation

5.1 Introduction

In this chapter, we will examine the performance analysis and code design of STTC over independent, non-identically distributed (i.n.i.d.), rapid, fading channels with imperfect channel estimation. Most works on STTC assume that fading channels are i.i.d. and perfect CSI is available at the receiver. In many situations, this assumption may not hold. First, the fading processes may have different variances, or different channel fade rates depending on the individual scattering environment on each propagation path. For example, the channel multipath intensity profile of IMT-2000 channel models [45] and JTC channel models [46] is variable, i.e., the mean square fading gain of each diversity branch is different. Second, in practice, perfect CSI may not be obtained due to channel estimation errors, especially for time-varying channels. These factors motivate us to extend our work in Chapters 3 and 4 and consider the general i.n.i.d. MIMO channels.

First, the ML receiver is derived for STTC over non-selective, i.n.i.d. rapid, Rayleigh fading channels with channel estimation. Due to the i.n.i.d. fading channels, the receiver requires in its signal detection function, both the channel estimates and the second order statistical information of the estimates, which can be obtained from the channel estimator. The exact PEP, PEP bounds and union bounds on the BEP of STTC for the ML receiver are obtained. The new PEP bounds are explicit and simple to compute, which is helpful in providing insights into suppressing the performance loss caused by imperfect channel estimation. Based on the performance analysis, a design criterion and new STTC over i.n.i.d. channels with imperfect channel estimation are proposed. To minimize the performance loss caused by channel estimation errors, the symbol-wise Euclidean distances among codewords should be optimally distributed based on the statistical information of the channel estimates. Under i.n.i.d. channel conditions, our newly proposed codes perform better than the existing codes in the literature which are designed assuming i.i.d. fading channels and perfect CSI at the receiver. When prior knowledge of the channel and estimator is available at the transmitter, the codes whose design parameters are matched to the channel and estimator parameters should be used to achieve the optimal performance.

5.2 The System Model

For the i.n.i.d. case, we still employ the TDM STTC system as shown in Fig. 3.2. However, unlike the i.i.d. case, the statistics of the fading processes of the links are non-identical. Here, we give a brief review of the system model. In the TDM STTC system, the data streams and the pilot symbols are multiplexed into frames. In one transmission frame, there are F time slots, and F/L blocks, where L is the pilot spacing and F/L is a positive integer. Each block has two phases. One is the data phase; the other is the pilot phase. The signal models for the two phases are given below.

5.2.1 The Data Phase

During the data phase, only one user is allowed to access the channel in each time slot. The user transmits the N_T coded symbols simultaneously in the assigned time slot. The outputs of the receiver matched filters sampled at the uth time slot in the kth frame, $\mathbf{r}(k,u)$, can be expressed as,

$$\mathbf{r}(k,u) = \sqrt{E_s}\mathbf{H}(k,u)\mathbf{v}(k,u) + \mathbf{n}(k,u) \quad 1 \le k \le K;\ 1 \le u \le F\,. \tag{5.1}$$

The pair (k,u) corresponds to a time point t on the real time axis, which is given by $t = (k-1)F + u$. $\mathbf{H}(k,u) = [h_{ij}(k,u)]$ is the $N_R \times N_T$ channel fading matrix. On the (i,j)th link, $\{h_{ij}(k,u)\}$ is a complex, Gaussian sequence with mean zero and autocorrelation function $\mathrm{E}[h_{ij}(k,u)h_{ij}^*(k',u')] = 2\sigma_{ij}^2 R_{ij}(t-t')$, where $R_{ij}(\tau)$ is the normalized autocorrelation function of both the real and imaginary parts of the fading process on the (i,j)th link, which are independent of one another. The variances σ_{ij}^2's and the correlation functions $R_{ij}(\tau)$'s for different links are different in general. We assume that both the additive noises and the fading processes are spatially independent, and are independent of one another.

5.2.2 The Pilot Phase

During the pilot phase, N_T pilot symbols are transmitted one at a time in each time slot from each transmit antenna to allow estimation of all the channels associated with that antenna. The pilot for the jth transmit antenna is transmitted in the jth time slot in each block, i.e., at time $t_{lj} = (l-1)L + j$, where $l = 0, \cdots, ((KF/L) - 1)$ and $j = 1, \cdots, N_T$. The received pilot symbol at the ith receive antenna is

$$r_i(t_{lj}) = \sqrt{E_p}h_{ij}(t_{lj})v_j^p(t_{lj}) + n_i(t_{lj}), \quad 1 \le i \le N_R\,, \tag{5.2}$$

where E_p is the pilot energy. The normalized received pilot symbol $z_i(t_{lj}) = r_i(t_{lj})/(E_p v_j^p(t_{lj}))$ is used as the channel measurement. Similar to the i.i.d. case,

let $\mathcal{I} = \{\{\{z_i(t_{lj})\}_{l=0}^{(KF/L)-1}\}_{i=1}^{N_R}\}_{j=1}^{N_T}$ be the set of all channel measurements obtained from the pilot symbols. The estimator only uses the nearest N noisy pilot measurements on the (i,j)th link, i.e., $\mathbf{z}_{ij}(t) = [z_i(t_j - \lfloor N/2 \rfloor)L) \ \cdots \ z_i(t_j + \lfloor (N-1)/2 \rfloor)L)]^T$, to estimate $h_{ij}(t)$. The pilot symbol from the current block is located at the time $t_j = j + \lfloor t/L \rfloor L$.

5.2.3 The Statistics of the Channel Estimates

Using the PSAM channel estimation scheme, the estimate $\hat{h}_{ij}(t)$ of $h_{ij}(t)$ is a complex, Gaussian, random variable with mean zero and variance

$$\mathrm{E}\left[\hat{h}_{ij}(t)\hat{h}_{ij}^*(t)\right] = 2\hat{\sigma}_{ij}^2(\tau) = \mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}^H (\mathbf{\Sigma}_{\mathbf{z}_{ij}}^{-1})^H \mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}} \ . \tag{5.3}$$

For the i.n.i.d. fading channels, the autocorrelation matrix of the channel measurements, $\mathbf{\Sigma}_{\mathbf{z}_{ij}} = \mathrm{E}[\mathbf{z}_{ij}(t)\mathbf{z}_{ij}^H(t)]$, and the crosscorrelation vector between the channel measurements and the fading coefficient to be estimated, $\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}} = \mathrm{E}[\mathbf{z}_{ij}(t)h_{ij}^*(t)]$, are given by different expressions from those obtained in the i.i.d. case in (3.10) and (3.11). Here, the (m,n)th entry of $\mathbf{\Sigma}_{\mathbf{z}_{ij}}$, is given by

$$\mathbf{\Sigma}_{\mathbf{z}_{ij}}(m,n) = \begin{cases} 2\sigma_{ij}^2 R_{ij}\left((m-n)L\right) & m \neq n \\ 2\sigma_{ij}^2 + N_0/E_p & m = n \ . \end{cases} \tag{5.4}$$

The nth entry of $\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}$ is given by

$$\mathbf{\Sigma}_{h_{ij}\mathbf{z}_{ij}}^H(n) = 2\sigma_{ij}^2 R_{ij}\left((n-1-\lfloor N/2 \rfloor)L - \tau\right) \ . \tag{5.5}$$

Here τ is the time interval between $h_{ij}(t)$, and the corresponding pilot symbol of the current block transmitted at time t_j, and is given by

$$\tau = t - t_j = u - \lfloor u/L \rfloor \times L - j \ . \tag{5.6}$$

The estimation error $e_{ij}(t) = h_{ij}(t) - \hat{h}_{ij}(t)$ is also a zero-mean, complex, Gaussian random variable, and attains its minimum variance of [39]

$$2\bar{\sigma}_{ij}^2(\tau) = 2\sigma_{ij}^2 - 2\hat{\sigma}_{ij}^2(\tau) \ . \tag{5.7}$$

It can be seen from (5.3) to (5.7) that both the estimate variance $\hat{\sigma}_{ij}^2(\tau)$ and the estimation error variance $\bar{\sigma}_{ij}^2(\tau)$ depends on the link index (i,j), which varies from one link to another. These two variances are also functions of the time interval τ. Similar to the i.i.d. case, the variances will be written more briefly as $\hat{\sigma}_{ij}^2(u)$ and $\bar{\sigma}_{ij}^2(u)$, respectively, because of the one-to-one relationship between τ and u for a fixed transmit antenna index j. At the receiver, after de-multiplexing, the estimated channel fading process for each user u, $\{\hat{h}_{ij}(k,u)\}_{k=1}^K$, can be easily obtained. Due to perfect interleaving, the corresponding estimated fading coefficients for the uth user, $\{\hat{h}_{ij}(k,u)\}_{k=1}^K$, is a sequence of zero-mean, complex, Gaussian random variables which are all mutually uncorrelated, with $\mathrm{E}\left[\hat{h}_{ij}(k,u)\hat{h}_{ij}^*(l,u)\right] = 2\hat{\sigma}_{ij}^2(u)\delta(k-l)$. In the following, we focus on one particular user and drop the user index u for simplicity.

5.3 Performance Analysis

By incorporating the pilot channel measurements into the receiver, we next derive the ML receiver structure under channel estimation.

5.3.1 The ML Receiver

Suppose the entire sequence of transmitted signals over K frames for the user concerned is $\mathbf{v} = [\mathbf{v}^T(1)\cdots\mathbf{v}^T(K)]^T$. Then, the received signal, $\mathbf{r} = \left[\mathbf{r}^T(1)\cdots\mathbf{r}^T(K)\right]^T$, is given by

$$\mathbf{r} = \sqrt{E_s}\mathbf{H}\mathbf{v} + \mathbf{n} \tag{5.8}$$

where $\mathbf{H} = diag\left[\mathbf{H}(1)\ \cdots\ \mathbf{H}(K)\right]$ is the combined channel fading matrix and the combined noise vector is $\mathbf{n} = \left[\mathbf{n}^T(1)\ \cdots\ \mathbf{n}^T(K)\right]^T$. The estimated channel fading matrix is $\hat{\mathbf{H}} = \mathrm{E}\left[\mathbf{H}|\mathcal{I}\right] = diag\left[\hat{\mathbf{H}}(1)\ \cdots\ \hat{\mathbf{H}}(K)\right]$. The estimation error matrix is $\mathbf{E} = diag\left[\mathbf{E}(1)\ \cdots\ \mathbf{E}(K)\right]$, where $\mathbf{E}(k) = \mathbf{H}(k) - \hat{\mathbf{H}}(k)$.

Based on the set of all pilot channel measurements $\mathcal{I}$ and the entire received sequence $\mathbf{r}$, the ML receiver detects the transmitted signal sequence $\mathbf{v}$ by computing the conditional probability density function $p(\mathbf{r}|\mathcal{I}, \mathbf{v} = \mathbf{v}_l)$. Conditioned on $\mathcal{I}$ and $\mathbf{v}$, $\mathbf{r}$ is a complex, Gaussian random vector with mean

$$\boldsymbol{\eta} = [\boldsymbol{\eta}^T(1) \cdots \boldsymbol{\eta}^T(k) \cdots \boldsymbol{\eta}^T(K)]^T = \mathrm{E}[\mathbf{r}|\mathcal{I}, \mathbf{v}] = \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v} . \tag{5.9}$$

The ith element of $\boldsymbol{\eta}(k)$ is $\eta_i(k) = \sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}(k)$, where $\hat{\mathbf{h}}_i(k)$ is the ith column of $\hat{\mathbf{H}}^T(k)$. The conditional covariance matrix of $\mathbf{r}$ is given by

$$\boldsymbol{\Gamma} = \mathrm{E}\left[(\mathbf{r} - \boldsymbol{\eta})(\mathbf{r} - \boldsymbol{\eta})^H|\mathcal{I}, \mathbf{v}\right] = E_s\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right] + N_0\mathbf{I}_{N_RK} . \tag{5.10}$$

In Appendix A, it is shown that the covariance matrix $\boldsymbol{\Gamma}$ has a block diagonal structure, and can be expressed as

$$\boldsymbol{\Gamma} = \mathbf{I}_K \otimes \overline{\mathbf{N}}_0 \tag{5.11}$$

where $\otimes$ denotes the Kronecker product. The $N_R \times N_R$ *effective noise matrix* is $\overline{\mathbf{N}}_0 = diag\left[\overline{N}_{0_1} \cdots \overline{N}_{0_{N_R}}\right]$. The diagonal element $\overline{N}_{0_i}$ is the variance of the effective noise at the ith receive antenna, and can be computed as

$$\overline{N}_{0_i} = N_0 + 2E_s \sum_{j=1}^{N_T} \bar{\sigma}_{ij}^2 . \tag{5.12}$$

From (5.12), it can be seen that the variances of the channel estimation errors increase the total noise power at each receive antenna, and lead to the effective noise being different from one receive antenna to another because of the non-identical channel statistics. The unequal elements in $\overline{\mathbf{N}}_0$ result in $\boldsymbol{\Gamma}$ not being proportional to an identity matrix. Note that both $\overline{\mathbf{N}}_0$ and $\boldsymbol{\Gamma}$ are independent of the signal sequence $\mathbf{v}$, and this results from the equal-energy MPSK modulation. Given the conditional mean $\boldsymbol{\eta}$ and the conditional covariance matrix $\boldsymbol{\Gamma}$ of the Gaussian random vector $\mathbf{r}$, its conditional probability density function can be expressed as

$$p(\mathbf{r}|\mathcal{I}, \mathbf{v}) = \frac{1}{(\pi)^{N_RK}|\boldsymbol{\Gamma}|} \exp\left\{-(\mathbf{r} - \boldsymbol{\eta})^H\boldsymbol{\Gamma}^{-1}(\mathbf{r} - \boldsymbol{\eta})\right\} \tag{5.13}$$

where $|\boldsymbol{\Gamma}|$ denotes the determinant of $\boldsymbol{\Gamma}$. Recall that $\boldsymbol{\Gamma}$ is independent of the signal sequence $\mathbf{v}$, and thus $\frac{1}{(\pi)^{N_RK}|\boldsymbol{\Gamma}|}$ is a common term for all hypothesized transmitted

signal sequences. Taking the logarithm of $p(\mathbf{r}|\mathcal{I}, \mathbf{v})$, and dropping the common term, the ML decision metric can be reduced to

$$\begin{aligned} m(\mathbf{r}, \mathbf{v}) &= (\mathbf{r} - \boldsymbol{\eta})^H \boldsymbol{\Gamma}^{-1} (\mathbf{r} - \boldsymbol{\eta}) \\ &= \sum_{k=1}^{K} \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} |r_i(k) - \sqrt{E_s} \hat{\mathbf{h}}_i^T(k) \mathbf{v}(k)|^2 \,. \end{aligned} \tag{5.14}$$

Thus, the ML receiver detects the transmitted signal sequence $\mathbf{v}$ by computing $m(\mathbf{r}, \mathbf{v})$ for each possible value $\mathbf{v}_l$ of $\mathbf{v}$, and deciding that $\mathbf{v} = \mathbf{v}_j$ if $m_j = \min_l m(\mathbf{r}, \mathbf{v}_l)$. In this i.n.i.d. channel case, channel estimation accuracy plays an important role in determining the weight on the signals received at each receive antenna. In the limit of perfect channel estimation, where we have $\bar{\sigma}_j^2 = 0$ and $\overline{N}_{0_i} = N_0$, the receiver (5.14) reduces to the known receiver for the perfect CSI case.

5.3.2 The exact PEP and the PEP Bounds

Based on the ML metric derived in (5.14), the conditional PEP $P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c)$ that the receiver decides in favor of an erroneous codeword $\mathbf{v}_e$ when the actual codeword sent is $\mathbf{v}_c$ and $\mathbf{v}_e$ is the only other alternative, is given by

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) = P\left(X > 0|\mathcal{I}, \mathbf{v} = \mathbf{v}_c\right) \,. \tag{5.15}$$

Here X is given by

$$X = \sum_{k=1}^{K} \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} \Big\{ |r_i(k) - \sqrt{E_s} \hat{\mathbf{h}}_i^T(k) \mathbf{v}_c(k)|^2 - |r_i(k) - \sqrt{E_s} \hat{\mathbf{h}}_i^T(k) \mathbf{v}_e(k)|^2 \Big\} \,.$$

It is shown in Appendix B that X is a conditional Gaussian random variable given $\mathcal{I}$, with conditional mean and variance given by

$$\mathrm{E}\left[X|\mathcal{I}, \mathbf{v} = \mathbf{v}_c\right] = -E_s \sum_{k \in \kappa} \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} |\hat{\mathbf{h}}_i^T(k) \mathbf{v}_{ce}(k)|^2 \tag{5.16}$$

and

$$\mathrm{Var}\left[X|\mathcal{I}, \mathbf{v} = \mathbf{v}_c\right] = 2E_s \sum_{k \in \kappa} \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} |\hat{\mathbf{h}}_i^T(k) \mathbf{v}_{ce}(k)|^2 \tag{5.17}$$

where $\mathbf{v}_{ce}(k) = \mathbf{v}_c(k) - \mathbf{v}_e(k)$ is the code difference vector, and κ is the set of frame indices k where $\mathbf{v}_c(k) \neq \mathbf{v}_e(k)$. The conditional PEP in (5.15) can thus be computed as

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) = Q\left(\sqrt{\frac{E_s}{2}\sum_{k\in\kappa}\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}|\hat{\mathbf{h}}_i^T(k)\mathbf{v}_{ce}(k)|^2}\right) . \tag{5.18}$$

We define

$$\hat{\mathbf{h}}(k) = \text{vec}(\hat{\mathbf{H}}^T(k)) = [\hat{\mathbf{h}}_1^T(k)\cdots\hat{\mathbf{h}}_{N_R}^T(k)]^T \tag{5.19a}$$

$$\mathbf{B}^H(k) = (\overline{\mathbf{N}}_0)^{-\frac{1}{2}} \otimes \mathbf{v}_{ce}^*(k) \tag{5.19b}$$

$$\mathbf{A}(k) = \mathbf{B}^H(k)\mathbf{B}(k) \tag{5.19c}$$

where vec$(\cdot)$ is the vectorization operator. Then, (5.18) can be rewritten as $P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) = Q\left(\sqrt{\frac{E_s}{2}\mathcal{D}}\right)$, where $\mathcal{D} = \sum_{k\in\kappa}\hat{\mathbf{h}}^H(k)\mathbf{A}(k)\hat{\mathbf{h}}(k)$. In Appendix C, we show that the characteristic function of $\mathcal{D}$ is

$$\psi_{\mathcal{D}}(\jmath\omega) = \text{E}[e^{\jmath\omega D}] = \prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1 - 2\jmath\omega(\overline{N}_{0_i})^{-1}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right)^{-1} \tag{5.20}$$

where $d_j(k) = |v_{cej}(k)|$, and $v_{cej}(k)$ is the jth element of $\mathbf{v}_{ce}(k)$. The average of the conditional PEP in (5.18) over all realizations of $\hat{h}_{ij}(k)$ can be expressed as

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v} = \mathbf{v}_c) = \text{E}_{\hat{\mathbf{h}}}\left[Q\left(\sqrt{\frac{E_s}{2}\mathcal{D}}\right)\right] = \text{E}_{\hat{\mathbf{h}}}\left[\frac{1}{\pi}\int_0^{\pi/2} e^{-\frac{E_s\mathcal{D}}{4\sin^2\theta}}d\theta\right] . \tag{5.21}$$

Interchanging the integration and expectation operations, and using the result obtained in (5.20), we have the final average PEP as follow:

$$P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v} = \mathbf{v}_c) = \frac{1}{\pi}\int_0^{\pi/2}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(\frac{\sin^2\theta}{\sin^2\theta + \frac{E_s}{2\overline{N}_{0_i}}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}\right)d\theta\,. \tag{5.22}$$

The result (5.22) is the exact PEP of STTC over i.n.i.d. fading channels with imperfect CSI. Since $\sin^2\theta \geq 0$, it can be dropped in the denominator of the integrand in (5.22), which gives the first PEP upperbound (abbreviated as PUB1) with imperfect CSI as

$$PUB1|_{ICSI} = \frac{1}{\pi}\int_0^{\frac{\pi}{2}}\prod_{k\in\kappa}\left(\frac{\sin^{2N_R}\theta}{\prod_{i=1}^{N_R}\left[\frac{\gamma_i}{2}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right]}\right)d\theta \tag{5.23}$$

where $\gamma_i = E_s/\overline{N}_{0_i}$ is the *effective SNR* at the ith receive antenna with imperfect channel estimation, and $\overline{N}_{0_i}$ is given by (5.12). The SNR per receive antenna for the perfect CSI case is $\gamma = E_s/N_0$. Using the identity: $\int_0^{\frac{\pi}{2}} \sin^{2n}\theta = \frac{\pi}{2}\frac{(2n-1)!!}{2n!!}$, $n > 0$ [71, sect. 3.621 eq. (3)], $PUB1|_{ICSI}$ can be reduced to

$$\begin{aligned} PUB1|_{ICSI} &= \frac{1}{2}\frac{(2q-1)!!}{(2q)!!}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(\frac{\gamma_i}{2}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right)^{-1} \\ &= \frac{1}{2}\frac{(2q-1)!!}{(2q)!!}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(\frac{\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}{2(\gamma^{-1}+2\sum_{j=1}^{N_T}\bar{\sigma}_{ij}^2)}\right)^{-1}. \end{aligned} \quad (5.24)$$

Here, $q = N_R \times |\kappa|$ is the total diversity order achieved over rapid fading channels, where $|\cdot|$ denotes the cardinality of the set, $(2q)!! = 2q(2q-2)\cdots 2$, and $(2q-1)!! = (2q-1)(2q-3)\cdots 1$. Alternatively, applying the inequality: $\frac{\sin^2\theta}{\sin^2\theta+a} \le \frac{1}{1+a}$, $a > 0$ to (5.22), we have

$$\begin{aligned} PUB2|_{ICSI} &= \frac{1}{2}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1+\frac{\gamma_i}{2}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right)^{-1} \\ &= \frac{1}{2}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1+\frac{\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}{2(\gamma^{-1}+2\sum_{j=1}^{N_T}\bar{\sigma}_{ij}^2)}\right)^{-1}. \end{aligned} \quad (5.25)$$

On the other hand, with the inequality: $\frac{\sin^2\theta}{\sin^2\theta+a} \ge \frac{\sin^2\theta}{1+a}$, $a > 0$ in (5.22), we have the lower bound on the PEP (abbreviated as PLB) given by

$$\begin{aligned} PLB|_{ICSI} &= \frac{1}{2}\frac{(2q-1)!!}{2q!!}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1+\frac{\gamma_i}{2}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right)^{-1} \\ &= \frac{1}{2}\frac{(2q-1)!!}{2q!!}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1+\frac{\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}{2(\gamma^{-1}+2\sum_{j=1}^{N_T}\bar{\sigma}_{ij}^2)}\right)^{-1}. \end{aligned} \quad (5.26)$$

The PEP bounds obtained above are easy to compute once the channel fading model and the channel estimator are given. They also provide new criteria for code design because they clearly indicate that the Euclidean distances between code symbols are *weighted* by the different variances of the estimated channel coefficients on the different links. The details of code design with imperfect CSI over i.n.i.d. channels will be given in the following section.

For a simplified semi-i.n.i.d. scenario, the ML receiver and the PEP results obtained above can be shown to reduce to the same as those for i.i.d. channels. The

semi-i.n.i.d. channels assume that all the fading links associated with a transmit antenna have the same distribution; but the distribution changes from one transmit antenna to another. Such a situation is common for the downlink of a cellular system, where the antennas on the base station are separated far enough, whereas the antennas on a mobile handset are closely mounted. Mathematically, we have $\sigma_{ij}^2 = \sigma_{i'j}^2$, and $R_{ij}(\tau) = R_{i'j}(\tau)$ for all $i, i' = 1, \cdots, N_R$. In other words, σ_{ij}^2 and $R_{ij}(\tau)$ are *independent* of the receive antenna i, which render the channel estimate variances $\hat{\sigma}_{ij}^2$ and the estimation error variances $\bar{\sigma}_{ij}^2$ independent of i. Therefore, we can drop the subscript i in the variances, and the covariance matrix $\mathbf{\Gamma}$ in (5.11) is reduced to

$$\mathbf{\Gamma} = \overline{N}_0 \mathbf{I}_{N_R K} \tag{5.27}$$

where $\overline{N}_0 = N_0 + 2E_s \sum_{j=1}^{N_T} \bar{\sigma}_j^2$. It can now be seen that the effective noise matrix $\overline{\mathbf{N}}_0$ for the full-i.n.i.d. case is reduced to being proportional to an identity matrix. The ML receiver can then be seen to reduce to minimizing the metric

$$m(\mathbf{r}, \mathbf{v}) = \|\mathbf{r} - \boldsymbol{\eta}\|^2 = \|\mathbf{r} - \sqrt{E_s}\hat{\mathbf{H}}\mathbf{v}\|^2 \,. \tag{5.28}$$

This metric is the same as that obtained for the case of i.i.d. fading channels in (3.25), and depends only on the estimated channel matrix $\hat{\mathbf{H}}$. This result shows that in this semi-i.n.i.d. case, the ML receiver reduces to the same form as in the i.i.d. case. The corresponding exact PEP expression and PEP bounds for the semi-i.n.i.d. case are, thus, the same as those in [68] obtained for i.i.d. channels. Due to the independence of $\hat{\sigma}_{ij}^2$ and $\bar{\sigma}_{ij}^2$ on i, these PEP results can be easily obtained from (5.22) through (5.26) by replacing $\hat{\sigma}_{ij}^2$ and $\bar{\sigma}_{ij}^2$ by $\hat{\sigma}_j^2$ and $\hat{\sigma}_j^2$, respectively.

5.3.3 The Upper Bounds on the BEP

Since the PEP results obtained above are in product form, the transfer function approach in [65] can be used to derive the union bound on the BEP, which is ultimately the performance measure of interest. For a geometrically uniform STTC

over Rayleigh fading channels, the error probability does not depend on the transmitted code sequences. Consequently, we can assume without loss of generality that the all-zero code sequence is transmitted. In the kth frame, for a transition from state a to state b, the branch in the state diagram of a STTC is labeled by

$$L_{a,b}(\theta) = N^{w_{a,b}} D_{a,b}(\theta) \tag{5.29}$$

where the exponent of the factor N indicates the number of error bits, i.e., the Hamming weight of the input bits associated with this transition. The factor $D_{a,b}(\theta)$ indicates the product distance between the output codevector associated with this transition and the actual transmitted codevector. From the expression of the exact PEP in (5.22), $D_{a,b}(\theta)$ is given by

$$D_{a,b}(\theta) = \begin{cases} \prod_{i=1}^{N_R} \frac{\sin^2\theta}{\sin^2\theta + \frac{E_s}{2N_{0_i}} \sum_{j=1}^{N_T} \hat{\sigma}_{ij}^2 d_j^2(k)} & \text{if the transition } a \rightarrow b \text{ exists} \\ 0 & \text{otherwise .} \end{cases} \tag{5.30}$$

The transfer function $T(D_{a,b}(\theta), N)$ of a STTC with 2^ν states, whose branch labels are given by (5.29), can be determined by solving the nodal equations [26],

$$T(D_{a,b}(\theta), N) = [\xi_1, \xi_2, \cdots, \xi_{2^\nu - 1}] \begin{bmatrix} L_{1,0}(\theta) \\ L_{2,0}(\theta) \\ \vdots \\ L_{2^\nu-1,0}(\theta) \end{bmatrix} \tag{5.31}$$

where

$$\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{2^\nu-1} \end{bmatrix} \tag{5.32}$$

$$= \begin{bmatrix} 1 - L_{1,1}(\theta) & -L_{2,1}(\theta) & \cdots & -L_{2^\nu-1,1}(\theta) \\ -L_{1,2}(\theta) & 1 - L_{2,2}(\theta) & \cdots & -L_{2^\nu-1,2}(\theta) \\ \vdots & \vdots & \ddots & \vdots \\ -L_{1,2^\nu-1}(\theta) & -L_{2,2^\mu-1}(\theta) & \cdots & 1 - L_{2^\nu-1,2^\nu-1}(\theta) \end{bmatrix}^{-1} \begin{bmatrix} L_{0,1}(\theta) \\ L_{0,2}(\theta) \\ \vdots \\ L_{0,2^\nu-1}(\theta) \end{bmatrix} .$$

With the transfer function obtained in (5.31), the union bound on the BEP P_b is given by [65]

$$P_b \leq \frac{1}{\pi}\int_0^{\pi/2} \frac{1}{m}\frac{\partial T(D_{a,b}(\theta), N)}{\partial N}|_{N=1} d\theta \tag{5.33}$$

where $m = \log_2 M$ is the number of bits per MPSK symbol. Thus, for geometrically uniform STTC over Rayleigh fading channels, the union bound on the BEP can be numerically computed as in (5.33). Alternatively, for general STTC, a simple, accurate estimate of the BEP P_b can be obtained more simply through the method of dominant error events. As shown in eq. (3.33) in Chapter 3, the BEP P_b can be estimated as

$$P_b \approx \frac{1}{n}\sum_{\mathbf{v}_c}\sum_{\mathbf{v}_e} a(\mathbf{v}_c \to \mathbf{v}_e) P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) P(\mathbf{v}_c)\,. \tag{5.34}$$

Since each term $P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v}_c)$ in (5.34) can be upper-bounded by PUB2 in (5.25),the estimated BEP can be upperbounded as

$$P_b < \frac{1}{n}\sum_{\mathbf{v}_c}\sum_{\mathbf{v}_e}\left(\frac{a(\mathbf{v}_c \to \mathbf{v}_e)P(\mathbf{v}_c)}{2}\prod_{k\in\kappa}\prod_{i=1}^{N_R}\left(1 + \frac{\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}{2\left(\gamma^{-1} + 2\sum_{j=1}^{N_T}\bar{\sigma}_{ij}^2\right)}\right)^{-1}\right). \tag{5.35}$$

In (5.35), we only sum over erroneous codewords $\mathbf{v}_e$ with error event lengths from the minimum value L_{pmin} which depends on the code structure, up to a certain maximum value L_{pmax} which we choose. Using PUB2, the BEP upperbound is shown in [68] to be tight with $L_{pmax} = L_{pmin} + \nu - 1$, where ν is the memory order of the code and L_{pmin} is the shortest error event length.

5.4 Code Design with Channel Estimation

In this section, we will discuss the code design of STTC with imperfect channel estimation over i.n.i.d. fading channels. From the above performance analysis, it can be seen that the effects of imperfect channel estimation are also reflected by the variances of the channel estimates and the channel estimation errors, which

is similar to the observation for the i.i.d. case. The effects of channel estimation on code design can be measured by the maximum variance difference among the channel estimates. Thus, the design criterion for STTC with channel estimation for the i.n.i.d. case is same as that for the i.i.d. case in principle. However, the computation of the variances of the channel estimates is different between the i.i.d. and the i.n.i.d. case. For the i.n.i.d. case, the differences among the variances of the channel estimates are attributed to two factors. One is the inherent non-identical statistics of the fading processes among the links. The other is that the correlation functions between the pilot symbols and the fading coefficient to be estimated are unequal among different transmit antennas. For the i.i.d. case, only the second factor matters.

We apply here the simple, explicit performance bounds to design new STTC, taking into account the statistical information of the channel estimates. We will use the node error event probability $P(e)$ as the performance measure in code design [72], which can be union bounded tightly at high SNR by

$$P(e) = \sum_{\mathbf{v}_c} P(\mathbf{v} = \mathbf{v}_c) P(e|\mathbf{v} = \mathbf{v}_c) \leq \Lambda \tag{5.36}$$

where $P(e|\mathbf{v} = \mathbf{v}_c) \leq \sum_{\mathbf{v}_e} P(\mathbf{v}_c \to \mathbf{v}_e|\mathbf{v} = \mathbf{v}_c)$ is the conditional node error event probability given that the correct codeword sent is $\mathbf{v}_c$. The maximum conditional node error probability $\Lambda = \max_{\mathbf{v}_c} P(e|\mathbf{v} = \mathbf{v}_c)$ is upperbounded by

$$\Lambda < \max_{\mathbf{v}_c} \frac{1}{2} \sum_{\mathbf{v}_e} \prod_{k \in \kappa} \prod_{i=1}^{N_R} \left[b_i \left(\mathbf{v}_c(k), \mathbf{v}_e(k) \right) \right]^{-1} . \tag{5.37}$$

where the $b_i \left(\mathbf{v}_c(k), \mathbf{v}_e(k) \right)$'s are the kth branch distances at the ith receive antenna. The value of $b_i \left(\mathbf{v}_c(k), \mathbf{v}_e(k) \right)$ can be obtained from the PEP upperbounds. Here, we use PUB2 in (5.25) as it has the same trend as the exact PEP. Thus, $b_i \left(\mathbf{v}_c(k), \mathbf{v}_e(k) \right)$ is given by

$$b_i \left(\mathbf{v}_c(k), \mathbf{v}_e(k) \right) = 1 + \frac{\sum_{j=1}^{N_T} \hat{\sigma}_{ij}^2 d_j^2(k)}{2(\gamma^{-1} + 2 \sum_{j=1}^{N_T} \bar{\sigma}_{ij}^2)} . \tag{5.38}$$

To minimize the node error probability $P(e)$, the upperbound on the maximum conditional node error probability Λ should be minimized. Therefore, we arrive at the following design criterion.

Design Criterion

To minimize the node error event probability, one should minimize Λ*, i.e., minimize the maximum sum of the products of the inverses of the branch distances,* $\sum_{\mathbf{v}_e} \prod_{i=1}^{N_R} \prod_{k \in \kappa} b_i^{-1}\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$*, over all possible codewords.*

The branch distances $b_i^{-1}\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$'s can be easily computed based on knowledge of the channel fading model and the channel estimator used. The proposed criterion incorporates the statistical information of the channel estimates. The performance degradation caused by the channel estimation errors can be minimized by optimally distributing the symbol-wise Euclidean distances based on the variances of the channel estimates. Thus, the codes designed using this criterion are more robust to channel estimation errors. Compared with the design criterion for the i.i.d. case with channel estimation, here $b_i^{-1}\left(\mathbf{v}_c(k), \mathbf{v}_e(k)\right)$'s are dependent on the receive antenna index because the variances of the channel estimates are different from one link to another. It has been shown that the variances of the channel estimates depend on the channel and estimator parameters. The channel parameters include the channel fade rates and the variances of the fading processes. The estimator parameters refer to the pilot spacing and the Wiener filter length. These parameters can be conveyed to the transmitter by a feedback link. Then, the corresponding optimal codes are designed by using parameters matched to these parameters. Using the design criterion given above, new STTC for non-identical channels with imperfect channel estimation can be obtained by using our reduced-complexity, iterative search algorithm proposed in Chapter 4. For different channel conditions, different optimal codes may be obtained. It will be shown that our proposed codes have improved error performances compared with those existing codes in the literature that are designed assuming i.i.d. channels and perfect CSI at the receiver. The performance improvement achieved by our codes over existing codes increases as the variance differences among all the fading links become greater. When the channel and the estimation parameters are known at the transmitter, the optimal codes should be used to achieve the best performance.

5.5 Numerical Results

In the numerical results, we assume the TDM PSAM system has $N_T = 2$ transmit and $N_R = 1$ receive antenna. On the (i, j)th link, the channel fading process has a Jake's power spectrum so that $R_{ij}(\tau) = J_0(2\pi f_{dij}\tau)$, where the channel fade rate is denoted by f_{dij}. The variance of the fading process of the first transmit antenna is denoted by σ_1^2, and that of the second is σ_2^2. The PSAM scheme is used to estimate the channels with pilot spacing $L = 8$ and Wiener filter length $N = 6$, for all channel fade rates that are lower than $f_dT = 0.05$, as suggested in Chapter 3. Without loss of generality, we consider the user whose time slot is located immediately after the pilot phase in the first block. To be fair, we account for the energy spent on the pilot symbols by taking the energy per data symbol in the PSAM scheme as $(E_s)_{PSAM} = E_s(L - N_T)/L$, where E_s is the data symbol energy for the perfect CSI case when no pilots are transmitted.

In Fig. 5.1, we examine the effects of the differences among the channel fade rates and the variances on the maximum variance difference $\triangle\hat{\sigma}^2_{max}$. The maximum variance difference $\triangle\hat{\sigma}^2_{max}$ is given by $\hat{\sigma}_2^2 - \hat{\sigma}_1^2$, which reflects the effect of channel estimation on code design, as shown in Chapter 4. The x-axis represents the difference between the channel fade rates, i.e., $f_{d2}T - f_{d1}T$. Without loss of generality, we assume $f_{d2}T > f_{d1}T$, and $f_{d1}T = 0.01$. Thus, the range of the channel fade rates considered is $[0.01,\ 0.05]$, during which channel estimates with satisfactory accuracy can be obtained using $L = 8$ and $N = 6$. Compared with the difference among the channel fade rates, the effects of the different variances can be seen to be more important. Thus, we examine here the i.n.i.d. channels by focusing on the different variances. The channel fade rates on each link are assumed to be equal.

Using the proposed iterative search algorithm in Chapter 4, we obtained new 8-state QPSK STTC's over i.n.i.d. channels with imperfect CSI, as shown in Tab. 5.1. These proposed codes are termed *ICSI* codes. The three optimal ICSI codes in Tab. 5.1 are designed for channels with different variances σ_1^2 and σ_2^2, and with a

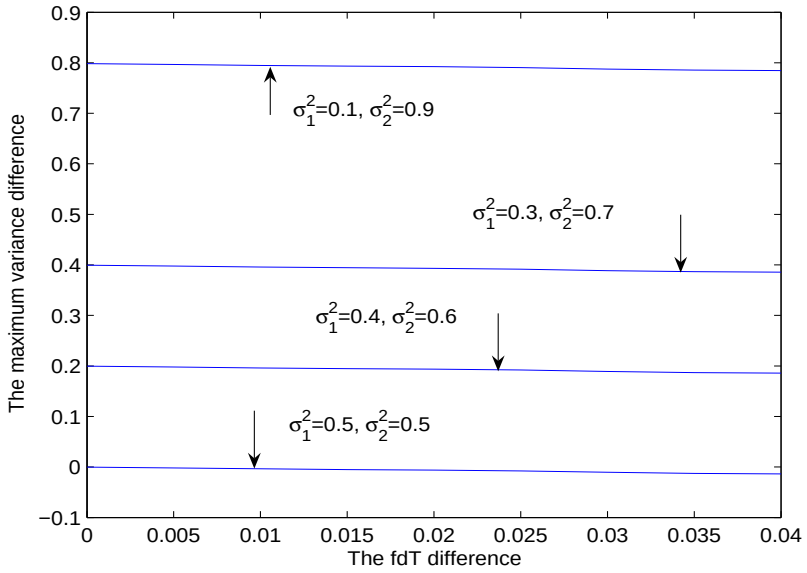


Fig. 5.1: The effects of the differences among the channel fade rates and the variances on the maximum estimation variance difference.

Table 5.1: The proposed 8-state QPSK code generator matrices G^T with two transmit antennas for i.n.i.d. channels with imperfect CSI.

(σ_1^2, σ_2^2)	ICSI
$\sigma_1^2 = 0.1, \sigma_2^2 = 0.9$	[1 3 2 1 1; 2 3 2 1 1]
$\sigma_1^2 = 0.3, \sigma_2^2 = 0.7$	[2 1 1 3 2; 2 3 2 3 1]
$\sigma_1^2 = 0.4, \sigma_2^2 = 0.6$	[2 0 1 2 2; 2 2 2 1 0]

fade rate of $f_dT = 0.05$. In Fig. 5.2, the exact PEP in (5.22), PEP upper bounds in (5.24) and (5.25), and PEP lower bound in (5.26) are plotted. We take the 8-state QPSK code of [1] (abbreviated as TSC code) over the shortest error event path as an example. It can be seen that PUB1 and PLB are very close to the exact PEP, especially for high SNR. Due to the use of the Chernoff bound, PUB2 is a bit loose. However, the advantage of PUB2 is that it nearly has the same trend as the exact PEP.

Fig. 5.3 shows both analytical and simulated BEP performance results for the

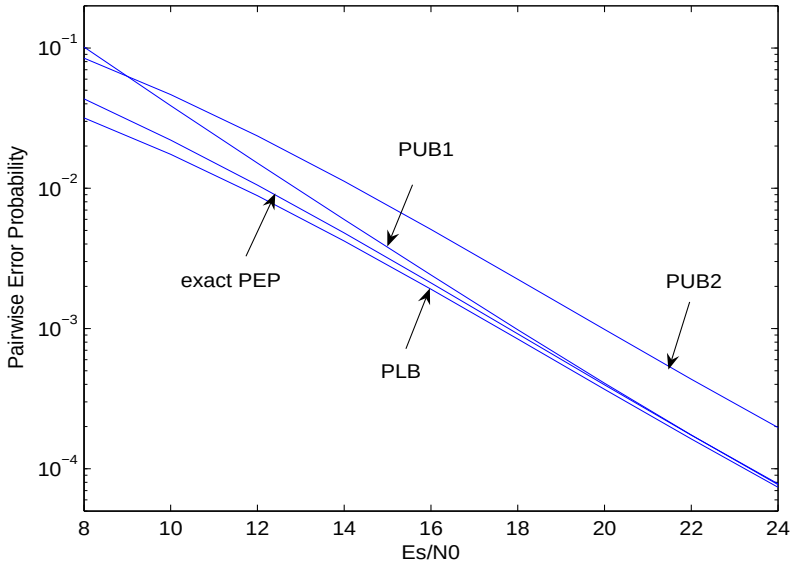


Fig. 5.2: The exact PEP and the PEP bounds for the 8-state QPSK TSC code of [1] over the shortest error event path with imperfect CSI, where $\sigma_1^2 = 0.3$, $\sigma_2^2 = 0.7$, and $f_dT = 0.05$.

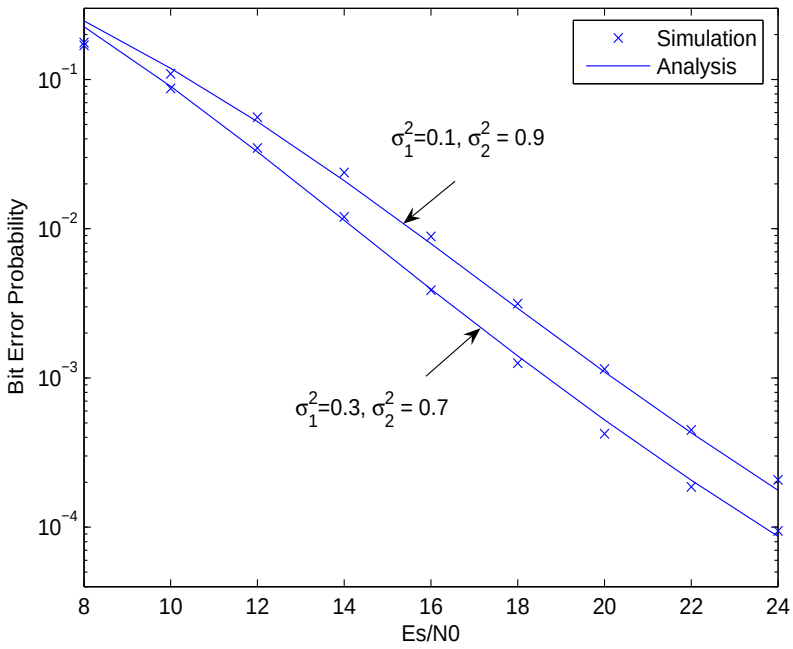


Fig. 5.3: The simulated and analytical BEP results for the 8-state QPSK TSC code of [1] over i.n.i.d. Rayleigh fading channels at $f_dT = 0.05$ with imperfect CSI.

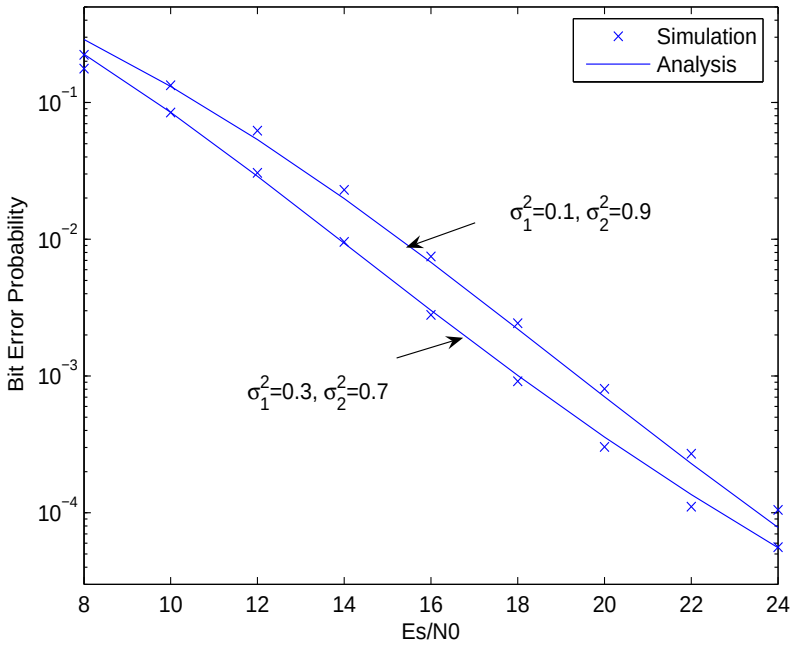


Fig. 5.4: The simulated and analytical BEP results for the 8-state QPSK FVY code of [2] over i.n.i.d. Rayleigh fading channels at $f_dT = 0.05$ with imperfect CSI.

8-state QPSK TSC code over i.n.i.d. fading channels. For all the simulated results, we take the number of frames as $K = 100$. Two channel conditions are considered. One has the variances $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$. The other has $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$. It can be seen that the analytical BEP results obtained using the method of dominant error events are very close to the simulated BEP results. The maximum error event length considered is given by $L_{pmax} = L_{pmin} + \nu - 1$. Similar results can be obtained for the 8-state FVY code of [2], as shown in Fig. 5.4.

Fig. 5.5 demonstrates that the proposed 8-state QPSK ICSI code with $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$ (abbreviated as ICSI-37 code) offers about 3dB performance gain over the TSC code at a BEP of 10^{-4}, when the actual channel variances are $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$. In Fig. 5.6, we compare the proposed 8-state QPSK ICSI code with $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$ (abbreviated as ICSI-19 code) with the TSC code over channels with $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$. It can be seen that the performance gain increases to nearly 4dB at the BEP of 10^{-4}. Thus, the results have shown the superiority of our proposed ICSI codes when only imperfect CSI is available

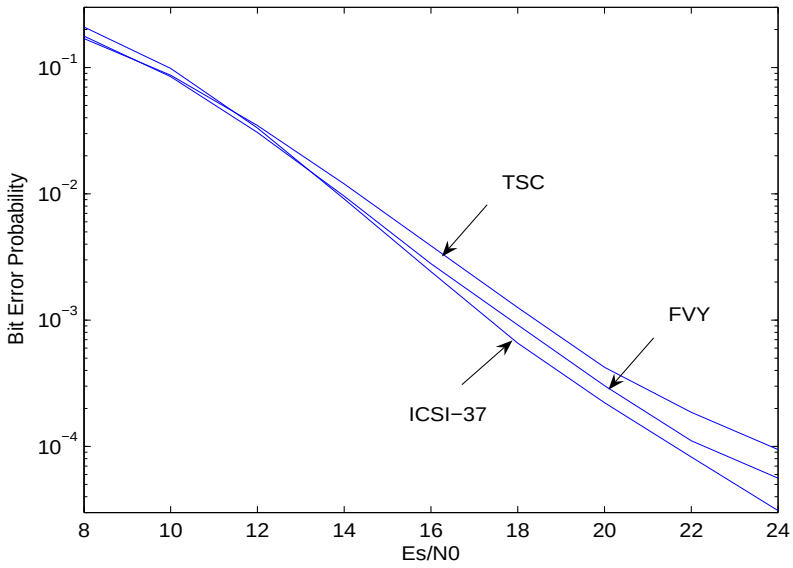


Fig. 5.5: The comparison of the simulated BEP results among the 8-state QPSK ICSI-37 code, the TSC code, and the FVY code over i.n.i.d. fading channels with $\sigma_1^2 = 0.3, \sigma_2^2 = 0.7$.

at the receiver. Furthermore, when the variance differences among channels increase, greater performance gains can be obtained by the proposed codes, and the advantage of using our proposed STTC is more obvious.

In addition to comparing our proposed ICSI codes with the codes known in the literature, Fig. 5.7 and Fig. 5.8 compare the performances of the three proposed ICSI codes under different channel conditions. In Fig. 5.7, the channels have variances $(\sigma_1^2 = 0.4, \sigma_2^2 = 0.6)$. The ICSI code designed with $(\sigma_1^2 = 0.4, \sigma_2^2 = 0.6)$ (abbreviated as ICSI-46 code) performs best, and this is to be expected because the design parameters of this code are matched to the channel parameters. In Fig. 5.8, the channels have variances $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$. Again, it is obvious have that the ICSI-19 code outperforms the other two codes. Thus, we have demonstrated that optimal codes matched to the channel and estimator parameters should be used when these parameters are known at the transmitter, if the degree of imbalance among links is sufficiently high.

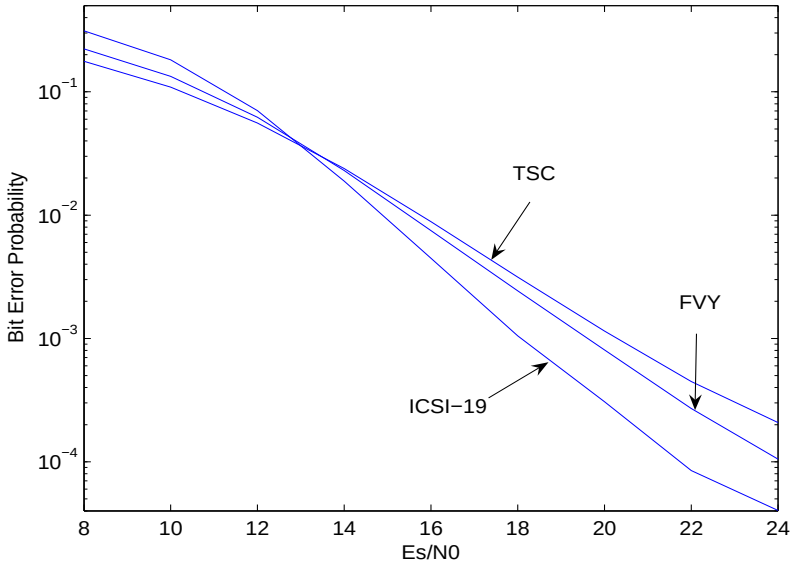


Fig. 5.6: The comparison of the simulated BEP results among the 8-state QPSK ICSI-19 code, the TSC code, and the FVY code over i.n.i.d. fading channels with $\sigma_1^2 = 0.1, \sigma_2^2 = 0.9$.

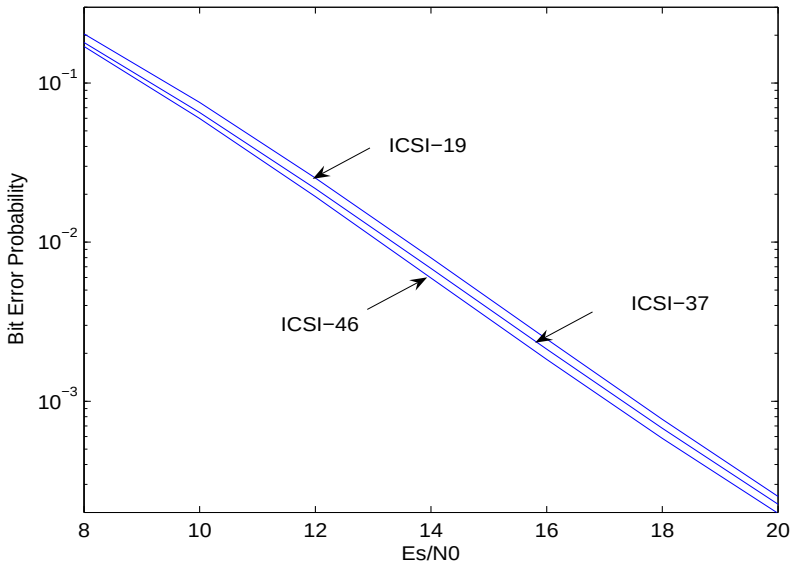


Fig. 5.7: The analytical BEP results of the three proposed 8-state QPSK ICSI codes over i.n.i.d. channels with $N_T = 2$, $N_R = 1$, $f_dT = 0.05$, $\sigma_1^2 = 0.4$ *and* $\sigma_2^2 = 0.6$.

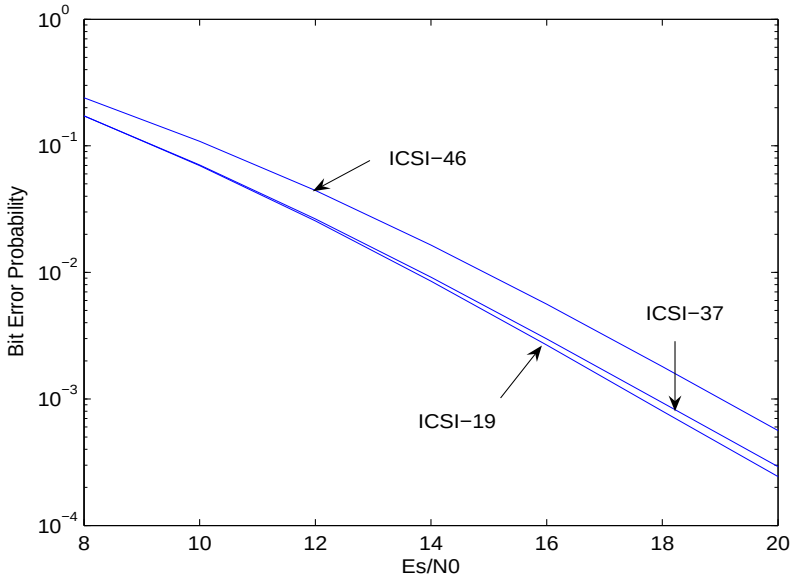


Fig. 5.8: The analytical BEP results of the three proposed 8-state QPSK ICSI codes over i.n.i.d. channels with $N_T = 2$, $N_R = 1$, $f_dT = 0.05$, $\sigma_1^2 = 0.1$ and $\sigma_2^2 = 0.9$, which are estimated with $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p = (E_s)_{PSAM}$.

5.6 Summary

We have extended the work on performance analysis and code design of STTC over rapid, Rayleigh fading channels with imperfect CSI to the i.n.i.d. case, by relaxing the constraint of identical statistical distribution on each link. This will lead to broader applications of STTC in practical MIMO systems. The ML receiver is derived, and shown to use not only the channel estimates but also the second-order statistical information of the estimates. It cannot simply be obtained from the perfect CSI ML receiver by replacing the known channel matrix $\mathbf{H}$ with the imperfectly estimated $\hat{\mathbf{H}}$. Its exact PEP expression, and PEP bounds are derived. Employing the node error event probability as the cost function, a practical code design criterion is presented and new STTC are obtained. When the channels have non-identical distributions, our proposed STTC have better performance than previous codes designed for rapid fading channels. When the variance differences among channel fading processes increase, the performance gains achieved by our proposed STTC are greater.

Chapter 6

Power Allocation with Side Information at the Transmitter

In the previous chapters, we have examined the performance analysis and code design of STTC over rapid Rayleigh fading channels with imperfect CSI at the receiver. The transmitter was assumed to have no knowledge of CSI in the open-loop system. However, the performance of the closed-loop system can be expected to be better than that of the open-loop system, by exploiting side information of channel at the transmitter. Thus, if CSI is available at the receiver, it is desirable to feed back CSI to the transmitter if it is possible. One way to improve performance of closed-loop systems is to use adaptive power allocation schemes based on the side information at the transmitter. Some other schemes, such as adaptive rate or adaptive modulation, may also be adopted. However, these are beyond the scope of this thesis due to the time limitation. We will here focus on the adaptive power allocation schemes.

6.1 Introduction

The study of communication systems with side information at the transmitter can be traced back to Shannon [73]. Then, Goldsmith and Varaiya [27] studied the

case with perfect CSI at the receiver and transmitter, and presented an adaptive coding scheme to achieve capacity. Caire and Shamai [74] investigated the case of imperfect CSI at the transmitter. All these works mentioned concentrate on SISO systems. For MIMO systems, Telatar [6] first examined the capacity with no or perfect CSI at the transmitter, and perfect CSI at the receiver. With no CSI at the transmitter, the optimal transmission strategy to achieve capacity is to allocate equal power in independent orthogonal directions along different transmit antennas. Conversely, with perfect CSI at the transmitter, optimum power allocation to each transmit antenna can be carried out by the well-known water-filling algorithm [75]. The information-theoretic capacity of MIMO channels with imperfect feedback is the subject of recent publications. There are still many open problems that remain to be solved in this area [76]. Narula et al [28] studied how to efficiently use side information in multiple-antenna data transmission over fading channels. The side information is modeled by a random vector, which may consist of noisy estimates of the channel coefficients, or the quantized channel information. Both expected SNR and mutual information are considered as performance measures. Lower performance is obtained as the quality of the side information degrades. For the multiple-input single-output (MISO) case, the optimum SNR can be achieved by beamforming in a direction determined by the eigenstructure of the conditional channel correlation matrix. For the mutual information criterion, unlike for the SNR criterion, beamforming is not always optimum. The condition that determines when beamforming is optimal for two transmit antennas is given. Since computation of the capacity for general channel distribution at the transmitter is a difficult problem, most research focused on some special distributions such as in the mean feedback or the variance feedback cases. For the mean feedback case, the estimate of the channels based on the feedback is known at the transmitter. The variance matrix of the channels is modeled as a matrix that is a scalar multiple of an identity matrix. The scalar may be interpreted as the variance of the estimation error. For the variance feedback case, the transmitter models the channel coefficients as Gaussian random variables with mean zero. The covariance matrix of the

channels is known at the transmitter, and is not necessary to be a scalar multiple of an identity matrix. In [35], the optimum transmission schemes to achieve capacity for MISO systems were presented with either mean feedback or covariance feedback. For mean feedback, when the feedback SNR is larger than a threshold, the optimum scheme is to use beamforming along the mean vector. Otherwise, the optimum scheme is to use transmit diversity, i.e., the principal eigenvector of the optimal input covariance matrix is along the channel mean vector, and the eigenvalues corresponding to the remaining eigenvectors are shown to be equal. For covariance feedback, the optimum solution consists of transmit independent Gaussian inputs along the eigenvectors of the covariance matrix. It is worth noting that all the works mentioned above assumed perfect CSI at the receiver. For more general scenarios, such as the case where only partial or imperfect CSI is known at transmitter/receiver [76, 77], or even no information of the channels is available [78], the capacity analysis is much more difficult. From the results obtained in the previous chapters, it has been known that for the open-loop system, the imperfect CSI at the receiver not only affects the performance and the receiver structure, but also the transmission scheme. Thus, it is important to examine how to exploit the side information at the transmitter for the closed-loop system when only imperfect CSI is available to the receiver. The scenario of imperfect CSI at the receiver has only been examined recently by [56], [79], [80], and [81]. Since the capacity with imperfect CSI at the receiver is difficult to evaluate even for the SISO system, only some bounds are obtained in the aforementioned works. References [56] and [80] considered quasi-static MIMO fading channels, and [79] focused on time-varying SISO systems. The most recently work in [81] extended the results in [80] and [79]. Spatial-temporal power allocation schemes were proposed. However, all the papers which considered the imperfect CSI case assumed both the channel fading gains and the channel estimates are i.i.d. random variables. Therefore, the water-filling power allocation policy [75] can be employed due to the i.i.d. assumption. Our results will show that this water-filling policy cannot work for i.n.i.d. channels. New power allocation schemes need be proposed. Compared with the quasi-static

fading channels, channel estimation for the rapid fading channels is different due to the time-varying nature of the channels. Thus, this results in that the effects of the channel estimation errors on the capacity of the rapid fading channels is different from that of the quasi-static fading channels. These factors motivate us to further explore the capacity of rapid, i.n.i.d. MIMO channels.

We examine here the capacity and the optimal power allocation schemes for STTC over semi-i.n.i.d. rapid, Rayleigh fading channels with imperfect CSI at the receiver. The semi-i.n.i.d. channels assume all the fading links associated with a transmit antenna have the same distribution; but the distribution changes from one transmit antenna to another. This system model for semi-i.n.i.d. channels has been described in Chapter 5. The channel estimates and the corresponding statistics obtained in Chapter 5 are used directly. Conditioned on the pilot channel measurements, a new lower bound on the capacity of rapid, semi-i.n.i.d. MIMO fading channels with imperfect CSI at the receiver is obtained, and the corresponding optimal power allocation scheme is derived. Due to the computational burden of this scheme, we transform the problem into one of designing the optimal transmit weighting matrix by using the analyzed PEP results derived in Chapter 5 as the objective function. When the estimated channel fading matrix is known at the transmitter, and the estimates are sufficiently reliable, a beamforming scheme is proposed to minimize the error performance. On the other hand, if only partial information, such as only the statistical information of channels, is available due to the limited-bandwidth feedback, a simple power allocation scheme is derived.

6.2 Closed-loop TDM System Model

Consider the closed-loop TDM MIMO system model of Fig. 6.1. There are N_T transmit and N_R receive antennas. The data from each user are encoded and multiplexed into frames. The frame structure is same as that of the open-loop TDM system in Chapter 3. Then, frames consisting of data from all users and

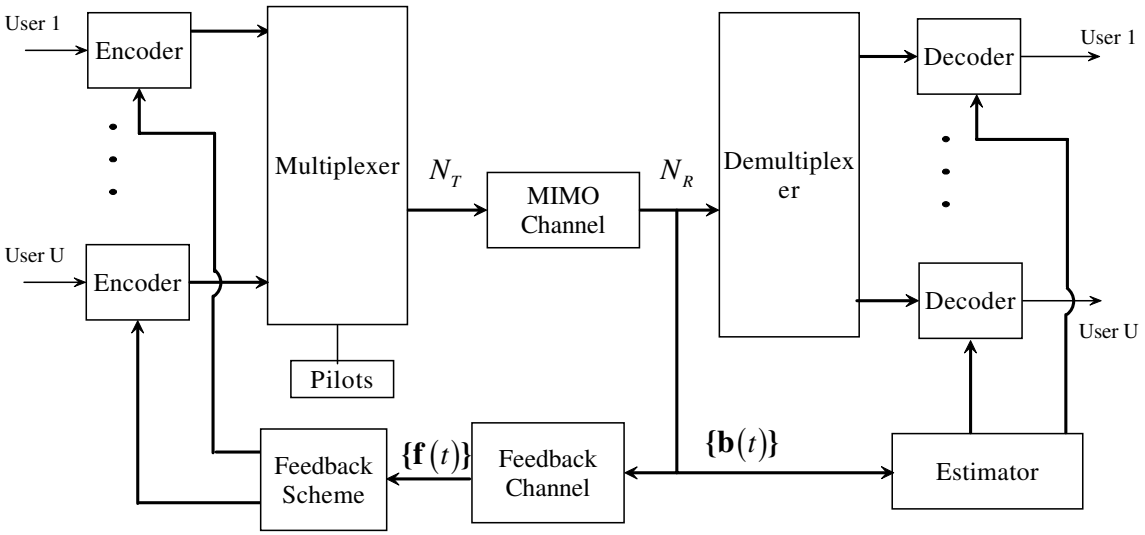


Fig. 6.1: The closed-loop TDM MIMO system model with PSAM.

inserted pilot symbols are transmitted over time-varying MIMO channels. The fading processes between each pair of transmit and receive antennas are assumed to be correlated over time, but independent over space. The channel information at the receiver and transmitter obtained at time t are represented by $\mathbf{b}(t)$ and $\mathbf{f}(t)$, respectively. The side information $\{\mathbf{f}(t)\}$ can be obtained by conveying $\{\mathbf{b}(t)\}$ via a feedback channel. Here we assume the feedback channel is error-free and delay-free. The feedback scheme is described by a function $f(\cdot)$. Thus, we have

$$\mathbf{f}(t) = f(\mathcal{B}) \tag{6.1}$$

where $\mathcal{B} = \{\mathbf{b}(1), \cdots, \mathbf{b}(t), \cdots\}$ is the sequence of channel information at the receiver. It should be emphasized here that the channel information are revealed to the transmitter in a causal fashion, and therefore no predictive encoding is possible [74]. The feedback scheme $f(\cdot)$ may be an estimator, equalizer or encoder, etc., which maps the channel information at the receiver to that at the transmitter. With the side information at the transmitter, the optimal transmission scheme to achieve capacity can be employed by adapting to the channel conditions.

In the TDM system, the data from each user are encoded and detected independently. Without loss of generality, we focus on one particular user in the following, and the signal model for one user is reduced to Fig. 6.2. The encoder of

the user produces a $N_T \times 1$ vector $\mathbf{x}(k)$ in the kth frame. The transmitted signal $\mathbf{x}(k)$ is a function of the input signals $\{u(k)\}$ and the side information $\mathbf{f}(k)$. The corresponding received $N_R \times 1$ vector, $\mathbf{r}(k)$, can be expressed as

$$\mathbf{r}(k) = \mathbf{H}(k)\mathbf{x}(k) + \mathbf{n}(k) \, . \tag{6.2}$$

The noise vector $\mathbf{n}(k)$ is $\mathcal{CN}(\mathbf{0}, N_0\mathbf{I}_{N_R})$ distributed, where $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the complex Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$. The channel fading matrix is $\mathbf{H}(k) = [h_{ij}(k)]$, where for a given k, $h_{ij}(k)$'s are all mutually independent, $\mathcal{CN}(0, 2\sigma_j^2)$ distributed. Here we considier the semi-i.n.i.d. fading channels. Thus, the variances of the fading processes are assumed to depend only on the transmit antenna index j. With the perfect interleaving/de-interleaving of the TDM system, it can be easily shown that the stationary fading process $\{h_{ij}(k)\}_{k=1}^{K}$ experienced by the user is i.i.d. over time. The fading and noise processes are independent of one another. In practice, the exact channel realization $\mathbf{H}(k)$ is not available. We use the PSAM scheme proposed in Chapter 3 to estimate it, and the estimated channel matrix is denoted as $\hat{\mathbf{H}}(k)$. In the PSAM scheme, pilot symbols are transmitted, and the normalized received pilot symbols are used as the channel measurements. The channel measurements from all pilot symbols during the entire transmission are denoted as $\mathcal{I}$. With channel estimation at the receiver, either the channel measurements $\mathcal{I}$ or the estimated channel matrix $\hat{\mathbf{H}}(k)$ obtained can be fed to the feedback channel. However, using the channel measurements during the pilot phase as the input of the feedback channel is more spectrum efficient than using the estimated channel matrix during the data phase as the input. This is due to the fact that the duration of the pilot phase is always shorter than that of the data phase, and, therefore, less data are needed to be transmitted.

The channel information at the receiver, i.e., the channel measurements, is used in two ways for the closed-loop system, as shown in Fig. 6.1. On the one hand, the channel estimator uses the channel measurements to obtain the MMSE channel estimates. On the other hand, the channel measurements are fed to the feedback channel. The usage of the channel measurements $\mathcal{I}$ to estimate the channel fading

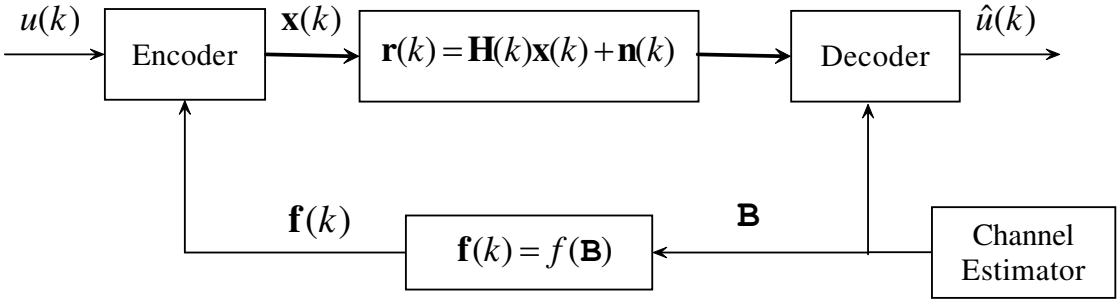


Fig. 6.2: The closed-loop TDM MIMO system model with PSAM for one user.

matrices at the receiver is well-known. We have examined the statistical information of the channel estimates in the previous chapters. For the semi-i.n.i.d. case, the elements of $\hat{\mathbf{H}}(k)$ are independent, $\mathcal{CN}(0, 2\hat{\sigma}_j^2)$ distributed. The elements in the estimation error matrix, $\mathbf{E}(k) = \mathbf{H}(k) - \hat{\mathbf{H}}(k)$, are also independent, $\mathcal{CN}(0, 2\bar{\sigma}_j^2)$ distributed. Since $\mathbf{H}(k)$ are i.i.d. over time due to complete interleaving, the corresponding estimate $\hat{\mathbf{H}}(k)$ and estimation error matrix $\mathbf{E}(k)$ are also i.i.d. over time. Compared to the advanced state of knowledge of using channel information at the receiver, the study of using the channel information at the transmitter for MIMO systems is still at the beginning stage [76]. Here, we assume the feedback scheme is the MMSE estimator. Thus, based on the output of the noisy-free and delay-free feedback channel, the same estimated channel matrix as that at the receiver can be obtained at the transmitter. The channel information at the receiver and transmitter can be expressed as

$$\mathcal{B} = \mathcal{I} \tag{6.3}$$

$$\mathbf{f}(k) = \hat{\mathbf{H}}(k)\,. \tag{6.4}$$

6.3 Capacity of MIMO Channels with Imperfect CSI at the Transmitter and Receiver

The capacity of MIMO channels with perfect CSI at the transmitter and receiver is well known. For a given channel realization, the capacity is given by

[6]

$$C(\mathbf{H}(k)) = \max_{\mathbf{Q}(k):Tr(\mathbf{Q}(k))=P} \log_2 \left| \mathbf{I}_{N_R} + \frac{1}{N_0} \mathbf{H}(k)\mathbf{Q}(k)\mathbf{H}^H(k) \right| \tag{6.5}$$

where $\mathbf{Q}(k) = \mathrm{E}[\mathbf{x}(k)\mathbf{x}^H(k)]$ is the input covariance matrix. The maximum $C(\mathbf{H}(k))$ is achieved by transmitting the signals along the eigenvectors of $\mathbf{H}^H(k)\mathbf{H}(k)$. With the singular value decomposition of the channel matrix, the MIMO channel is converted into parallel SISO channels. The power allocated to these parallel channels can be found by water-filling, i.e.,

$$P_i = \left(\varrho - \frac{N_0}{\lambda_i^2} \right)^+, \ 1 \le i \le \min(N_T, N_R) \tag{6.6}$$

where $\{\lambda_i\}$'s are the singular values of $\mathbf{H}(k)$, and ϱ is the waterfill level which satisfies $\sum_{i=1}^{\min(N_T,N_R)} P_i = P$. P_i is the power in the ith eigenmode of the channel and $(x)^+$ is defined as $\max(x, 0)$. Thus, the channel capacity given $\mathbf{H}(k)$ is

$$C\left(\mathbf{H}(k)\right) = \sum_{i=1}^{\min(N_T,N_R)} \log_2 \left(\frac{\varrho \lambda_i^2}{N_0} \right) . \tag{6.7}$$

By averaging $C(\mathbf{H}(k))$ over all channel realizations, we obtain the ergodic capacity

$$C = \mathrm{E}_{\mathbf{H}(k)} \left[C(\mathbf{H}(k)) \right] . \tag{6.8}$$

However, the capacity of MIMO channels with only imperfect CSI at the transmitter and receiver has not been fully examined because of the difficulty to determine the optimum distribution of the input signals [82], and the problem of computing the differential entropy of the received signal given the imperfect CSI. Some bounds are obtained for i.i.d. fading channels in [56], [81]. No research has been done on the capacity of i.n.i.d. MIMO channels. We next will study the capacity bound for the semi-i.n.i.d. rapid fading channels with imperfect CSI at the transmitter and receiver.

With the rapid fading model considered in Section 6.2, for each use of the channels an independent realization of $\mathbf{H}(k)$ is drawn. In this case, the capacity can be computed as the maximum mutual information [6]. Thus, we start by computing the mutual information $I(\mathbf{x}(k); \mathbf{r}(k)|\mathcal{I})$ between the input and the output of the

channels given the channel measurements at both the transmitter and receiver. Because of the Gaussian statistics, the estimated channel fading matrix $\hat{\mathbf{H}}(k)$ and the associated estimation error covariance matrix summarize all the information concerning $\mathbf{H}(k)$ given in the channel measurements $\mathcal{I}$. The mutual information $I(\mathbf{x}(k);\mathbf{r}(k)|\mathcal{I})$ given $\mathcal{I}$ is

$$I(\mathbf{x}(k);\mathbf{r}(k)|\mathcal{I}) = h(\mathbf{x}(k)|\mathcal{I}) - h(\mathbf{x}(k)|\mathbf{r}(k),\mathcal{I}) \tag{6.9}$$

where $h(\cdot)$ denotes the differential entropy. For the noise-free feedback channel, the channel information at the transmitter is a deterministic function of that at the receiver. It allows optimal codes to be constructed directly over the input alphabet, and be independent of the channel measurements [74].

Since the entropy of a random variable with given variance is upper-bounded by the entropy of a Gaussian random variable with the same variance, we have [79]

$$\begin{aligned} h(\mathbf{x}(k)|\mathbf{r}(k),\mathcal{I}) &= h(\mathbf{x}(k) - \mathbf{c}(k)\mathbf{r}(k)|\mathbf{r}(k),\mathcal{I}) \\ &\leq h(\boldsymbol{\varepsilon}(k)|\mathcal{I}) \\ &\leq \log_2\left[(2\pi e)^{N_T}\left|\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}(k)|\mathcal{I}}\right|\right] \end{aligned} \tag{6.10}$$

where $\mathbf{c}(k)$ is any $N_T \times N_R$ real matrix, $\boldsymbol{\varepsilon}(k)$ is given by $\mathbf{x}(k) - \mathbf{c}(k)\mathbf{r}(k)$, and $\boldsymbol{\Sigma}_{\mathbf{x}} = \mathrm{E}[\mathbf{x}\mathbf{x}^H]$ denotes the autocorrelation function of random vector $\mathbf{x}$, $\boldsymbol{\Sigma}_{(\mathbf{x},\,\mathbf{y})} = \mathrm{E}[\mathbf{x}\mathbf{y}^H]$ denotes the crosscorrelation function of $\mathbf{x}$ and $\mathbf{y}$, and $\boldsymbol{\Sigma}_{(\mathbf{x},\,\mathbf{y}|\mathbf{z})} = \mathrm{E}[\mathbf{x}\mathbf{y}^H|\mathbf{z}]$ is the conditional crosscorrelation function of $\mathbf{x}$ and $\mathbf{y}$ given $\mathbf{z}$. Combining (6.9) and (6.10), the mutual information $I(\mathbf{x}(k);\mathbf{r}(k)|\mathcal{I})$ can be lower-bounded by

$$I(\mathbf{x}(k);\mathbf{r}(k)|\mathcal{I}) \geq \log_2\left|\boldsymbol{\Sigma}_{\mathbf{x}(k)}\boldsymbol{\Sigma}^{-1}_{\boldsymbol{\varepsilon}(k)|\mathcal{I}}\right|. \tag{6.11}$$

The equality holds when $\mathbf{c}(k)\mathbf{r}(k)$ is the linear minimum mean-square error (LMMSE) estimate of $\mathbf{x}(k)$ in terms of $\mathbf{r}(k)$ [79]. The optimal $N_T \times N_R$ coefficient matrix $\mathbf{c}(k)$ is given by

$$\mathbf{c}(k) = \boldsymbol{\Sigma}_{(\mathbf{x}(k),\,\mathbf{r}(k)|\mathcal{I})}\boldsymbol{\Sigma}^{-1}_{\mathbf{r}(k)|\mathcal{I}} \tag{6.12}$$

and the minimum variance of the estimation error $\boldsymbol{\varepsilon}(k)$ is [55]

$$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}(k)|\mathcal{I}} = \boldsymbol{\Sigma}_{\mathbf{x}(k)} - \boldsymbol{\Sigma}_{(\mathbf{x}(k),\,\mathbf{r}(k)|\mathcal{I})}\boldsymbol{\Sigma}^{-1}_{\mathbf{r}(k)|\mathcal{I}}\boldsymbol{\Sigma}^{H}_{(\mathbf{x}(k),\,\mathbf{r}(k)|\mathcal{I})}. \tag{6.13}$$

With the signal model (6.2), it is easy to know that

$$\mathbf{\Sigma}_{(\mathbf{x}(k),\,\mathbf{r}(k)|\mathcal{I})} = \mathbf{\Sigma}_{\mathbf{x}(k)}\hat{\mathbf{H}}^H(k) \tag{6.14}$$

and due to the independence among $\mathbf{x}(k)$, $\mathbf{E}(k)$ and $\mathbf{n}(k)$, we obtain

$$\begin{aligned}\mathbf{\Sigma}_{\mathbf{r}(k)|\mathcal{I}} &= \mathbf{\Sigma}_{\hat{\mathbf{H}}(k)\mathbf{x}(k)} + \mathbf{\Sigma}_{\mathbf{E}(k)\mathbf{x}(k)} + \mathbf{\Sigma}_{\mathbf{n}(k)} \\ &= \hat{\mathbf{H}}(k)\mathbf{\Sigma}_{\mathbf{x}(k)}\hat{\mathbf{H}}^H(k) + \mathbf{\Sigma}_{\bar{\mathbf{n}}(k)}\end{aligned} \tag{6.15}$$

where $\mathbf{\Sigma}_{\bar{\mathbf{n}}(k)} = \mathbf{\Sigma}_{\mathbf{E}(k)\mathbf{x}(k)} + \mathbf{\Sigma}_{\mathbf{n}(k)}$ is the variance matrix of the effective noises. The part of $\mathbf{E}(k)\mathbf{x}(k)$ can be viewed as additional noises caused by the channel estimation errors. Recall that the elements of $\mathbf{E}(k)$ are independent, $\mathcal{CN}(0, 2\bar{\sigma}_j^2)$ distributed. Thus, $\mathbf{\Sigma}_{\bar{\mathbf{n}}(k)}$ can be easily shown to be

$$\mathbf{\Sigma}_{\bar{\mathbf{n}}(k)} = \overline{N}_0\mathbf{I}_{N_R} \tag{6.16}$$

where $\overline{N}_0 = 2\sum_{j=1}^{N_T} P_j\bar{\sigma}_j^2 + N_0$, and $P_j = \mathrm{E}[x_j(k)x_j^*(k)]$. The power of the effective noise is $\overline{N}_0$, which is increased by the channel estimation errors. Replacing $\mathbf{\Sigma}_{(\mathbf{x}(k),\,\mathbf{r}(k)|\mathcal{I})}$ and $\mathbf{\Sigma}_{\mathbf{r}(k)|\mathcal{I}}$ by their expressions in (6.14) and (6.15), respectively, the variance $\mathbf{\Sigma}_{\varepsilon(k)|\mathcal{I}}$ is given by

$$\mathbf{\Sigma}_{\boldsymbol{\varepsilon}(k)|\mathcal{I}} = \mathbf{\Sigma}_{\mathbf{x}(k)} - \mathbf{\Sigma}_{\mathbf{x}(k)}\hat{\mathbf{H}}^H(k)\left[\hat{\mathbf{H}}(k)\mathbf{\Sigma}_{\mathbf{x}(k)}\hat{\mathbf{H}}^H(k) + \mathbf{\Sigma}_{\bar{\mathbf{n}}(k)}\right]^{-1}\hat{\mathbf{H}}(k)\mathbf{\Sigma}_{\mathbf{x}(k)}\,. \tag{6.17}$$

Using the Woodbury's equation [83, sec. A.1], i.e.,

$$(\mathbf{A}+\mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1}+\mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}$$

the inverse of the variance $\mathbf{\Sigma}_{\varepsilon(k)|\mathcal{I}}$ can be expressed as

$$\mathbf{\Sigma}_{\boldsymbol{\varepsilon}(k)|\mathcal{I}}^{-1} = \mathbf{\Sigma}_{\mathbf{x}(k)}^{-1} + \hat{\mathbf{H}}^H(k)\mathbf{\Sigma}_{\bar{\mathbf{n}}(k)}^{-1}\hat{\mathbf{H}}(k)\,. \tag{6.18}$$

Thus, the lower-bound on (6.11) can be shown to be

$$I(\mathbf{x}(k);\mathbf{r}(k)|\mathcal{I}) \geq \log_2\left|\mathbf{I} + \hat{\mathbf{H}}^H(k)\mathbf{\Sigma}_{\bar{\mathbf{n}}(k)}^{-1}\hat{\mathbf{H}}(k)\mathbf{\Sigma}_{\mathbf{x}(k)}\right|. \tag{6.19}$$

Denote the singular value decomposition of $\hat{\mathbf{H}}(k)$ as $\hat{\mathbf{H}}(k) = \hat{\mathbf{U}}^H\hat{\mathbf{D}}\hat{\mathbf{V}}$, where $\hat{U}_{N_R\times N_R}$ and $\hat{V}_{N_T\times N_T}$ are unitary matrices, and $\hat{\mathbf{D}}_{N_R\times N_T}$ is a real diagonal matrix. Define

$\hat{\mathbf{D}} = \text{diag}[\hat{\lambda}_1, \cdots, \hat{\lambda}_r, 0, \cdots, 0]$, where $\hat{\lambda}_i$'s are the singular values and r is the rank of $\hat{\mathbf{H}}(k)$. Then, with the aid of (6.16), $\hat{\mathbf{H}}(k)^H \boldsymbol{\Sigma}_{\tilde{\mathbf{n}}(k)}^{-1} \hat{\mathbf{H}}(\mathbf{k})$ can be written as $\frac{1}{\overline{N}_0} \hat{\mathbf{V}}^H \boldsymbol{\Lambda} \hat{\mathbf{V}}$, where $\boldsymbol{\Lambda} = \hat{\mathbf{D}}^H \hat{\mathbf{D}}$. The lower bound on the mutual information on the right hand side of (6.19) can then be computed as

$$I_{lower}(\hat{\mathbf{H}}(k)) = \log_2 \left| \mathbf{I} + \frac{1}{\overline{N}_0} \boldsymbol{\Lambda} \tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)} \right| \tag{6.20}$$

where $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)} = \hat{\mathbf{V}} \boldsymbol{\Sigma}_{\mathbf{x}(k)} \hat{\mathbf{V}}^H$. This is our new lower bound on the mutual information with imperfect CSI at both the transmitter and receiver. The key difference of our result from others in [80, 81] is that the effective noise $\overline{N}_0$ depends on the unequal variances of the estimates at each transmit antenna due to the nonidentical channel assumption. This leads to that $\overline{N}_0$ is coupled with the transmitted symbol powers at the transmit antennas. Based on $I_{lower}(\hat{\mathbf{H}}(k))$, the lower bound $C_{lower}(\hat{\mathbf{H}}(k))$ on the capacity of the MIMO channels in (6.2), conditioned on a realization of the estimated channel matrix $\hat{\mathbf{H}}(k)$ available at both transmitter and receiver, is given by

$$C_{lower}(\hat{\mathbf{H}}(k)) = \max_{p\left(\mathbf{x}(k)|\hat{\mathbf{H}}(k)\right)} \log_2 \left| \mathbf{I} + \frac{1}{\overline{N}_0} \boldsymbol{\Lambda} \tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)} \right| \tag{6.21}$$

and the lower bound C_{lower} on the ergodic capacity with the estimated channel matrix $\hat{\mathbf{H}}(k)$ known at both transmitter and receiver is given by

$$C_{lower} = \mathrm{E}_{\hat{\mathbf{H}}(k)} \left[C_{lower}(\hat{\mathbf{H}}(k)) \right] = \mathrm{E}_{\hat{\mathbf{H}}(k)} \left[\max_{p\left(\mathbf{x}(k)|\hat{\mathbf{H}}(k)\right)} \log_2 \left| \mathbf{I} + \frac{1}{\overline{N}_0} \boldsymbol{\Lambda} \tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)} \right| \right]. \tag{6.22}$$

With the average power constraint of $\text{Tr}(\tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)}) \leq P$, the right side of (6.21) is maximized with $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)}$ a diagonal matrix, i.e., $\tilde{\boldsymbol{\Sigma}}_{\mathbf{x}(k)} = \text{diag}[P_1 \cdots P_{N_T}]$. Thus, the lower bound on the capacity in (6.21) can be calculated as

$$\begin{aligned} C_{lower}(\hat{\mathbf{H}}(k)) &= \max_{P_j} \sum_{j=1}^{r} \log_2 \left(1 + \frac{\hat{\lambda}_j^2 P_j}{\overline{N}_0} \right) \\ \text{subject to} \quad & \sum_{j=1}^{N_T} P_j \leq P, \quad P_j \geq 0 \end{aligned} \tag{6.23}$$

Note that $\overline{N}_0$ is a function of P_j due to the unequal variances of the channel estimates associated with different transmit antennas. The tight lower bound on

the capacity can be achieved by optimally allocating the power on eigenmodes of the estimated channel matrix. In the next section, we will examine the power allocation schemes.

6.4 Transmit Power Allocation Schemes

In principle, information throughput of MIMO systems can be further increased via an optimized allocation of the available power over transmit antennas [28]. The adaptive transmit power allocation has to be performed based on the channel information at the transmitter. Contrary to the perfect CSI case, the optimal transmit power allocation strategy for MIMO systems when both the transmitter and receiver have only imperfect CSI is still an open problem [84, 85]. Power allocation schemes were designed based on the SNR and the expected mutual information in [28]. However, the expression of mutual information in [28, eq. (26)] considered only imperfect CSI at the transmitter, and perfect CSI at the receiver. The same situation was also examined by [36], but the design of the weighting matrix in [36] is based on a PEP lowerbound. In [36], the authors presented a transmission scheme which adapts a predetermined space-time code to available channel knowledge by a linear transformation. The proposed optimal transmit weighting matrix can be obtained by computer search. Only for some simplified scenarios can the analytical expression for the transmit weighting matrix be obtained. The power allocation for multiple-antenna systems with imperfect channel estimation was examined in [84]. The proposed water-filling policy based on the conditional capacity given the estimated channel information is suboptimal, because they simply replaced the actual channel matrix by its estimate. In this section, we will first design the optimal transmit power allocation scheme using the capacity bound derived in (6.21) for the i.n.i.d. channels with imperfect CSI at both the transmitter and receiver. It is worth noting that the power allocation for the i.n.i.d. channels is quite different from that of i.i.d. channels when there are channel estimation errors. This is due to the fact that the effective noise $\overline{N}_0$ at the receiver is dependent on the power

allocated at each transmit antenna.

6.4.1 Design Based on the Capacity Lower Bound

Given the lower bound on the capacity of MIMO channels with imperfect CSI at the receiver obtained above, one way to design the transmit power allocation scheme is to maximize the right hand side of (6.23). Unlike the optimal power allocation for the perfect CSI case, whether the power allocation based on the capacity lowerbound for the imperfect CSI case is a convex optimization problem is not obvious, because the effective noise $\overline{N}_0$ in the denominator of the capacity lowerbound (6.23) is also a function of the transmission powers P_j, as mentioned. Based on the expression of $C_{lower}(\hat{\mathbf{H}}(k))$ in (6.23), the second derivation of the function

$$f(P_j) = \sum_{j=1}^{r} \log_2 \left(1 + \frac{\hat{\lambda}_j^2 P_j}{\overline{N}_0} \right) \tag{6.24}$$

is given by

$$\frac{\partial^2 f(P_j)}{\partial^2 P_j} = -\frac{\lambda_j^2(\overline{N}_0 - 2P_j\bar{\sigma}_j^2)(4\overline{N}_0\bar{\sigma}_j^2 + \lambda_j^2\overline{N}_0 + 2\lambda_j^2 P_j\bar{\sigma}_j^2)}{(\overline{N}_0^2 + P_j\lambda_j^2\overline{N}_0)^2}\,. \tag{6.25}$$

Since $\overline{N}_0 = 2\sum_{j=1}^{N_T} P_j\bar{\sigma}_j^2 + N_0$, we have $\overline{N}_0 - 2P_j\bar{\sigma}_j^2 > 0$. Therefore, it is easy to show that the second derivation of $f(P_j)$ is negative, i.e.,

$$\frac{\partial^2 f(P_j)}{\partial^2 P_j} < 0 \tag{6.26}$$

Thus, the function $f(P_j)$ is concave, and the optimization problem belongs to the class of convex optimization problems. Maximizing the capacity lower bound is equivalent to minimizing its negative. Since the capacity lower bound is a concave function of power, its negative is convex. The constraints are linear, so they are also convex. Thus, the optimum set of powers P_j's satisfies the Karush-Kuhn-Tucker

(KKT) conditions

$$\sum_{j=1}^{N_T} P_j = P, \quad P_j \geq 0 \tag{6.27}$$

$$-\frac{1}{\ln 2}\frac{\overline{N}_0\hat{\lambda}_j^2 - 2\hat{\lambda}_j^2 P_j\bar{\sigma}_j^2}{\overline{N}_0^2 + \hat{\lambda}_j^2 P_j\overline{N}_0} + \mu - \nu_j = 0 \tag{6.28}$$

$$\mu \geq 0, \quad \nu_j \geq 0, \quad \nu_j P_j = 0 \tag{6.29}$$

where $j = 1, \cdots, N_T$, μ and ν_j are the positive Lagrange multipliers for the power constraint and the positivity constraints on P_j, respectively. Finally, it can be shown that the optimal P_j is

$$P_j = \left(\frac{\overline{N}_0(\hat{\lambda}_j^2 - \mu\overline{N}_0\ln 2)}{\hat{\lambda}_j^2(2\bar{\sigma}_j^2 + \mu\overline{N}_0\ln 2)}\right)^+ \tag{6.30}$$

where μ is chosen so that $\sum_{j=1}^{N_T} P_j = P$. The optimization tool $fmincon$ in Matlab can be used to compute the optimal power allocation at each transmit antenna. From the equation of P_j in (6.30), we have $P_j = 0$ if

$$\frac{\hat{\lambda}_j^2}{\overline{N}_0} \leq \mu\ln 2 \tag{6.31}$$

It indicates that no power is assigned to a channel which is in a bad condition. The channel condition is affected by the effective channel noise $\overline{N}_0$, which, in turn, depends on the channel estimation accuracy and the average SNR. When one channel has much better condition than the others, the beamforming scheme can be used to achieve the capacity. Although this power allocation scheme can approach the channel capacity, it entails high complexity due to the computation of singular values $\hat{\lambda}_i$ and the adaptive optimization procedure. The approximating constant-power allocation scheme may be used to simplify the transmitter design [86]. However, it is only suitable for high SNR because of the saturation property of the logarithmic function at high SNR. To reduce the computational complexity, we may resort to some other possible cost functions.

6.4.2 Design Based on the PEP Lower Bounds

To avoid the computation of the eigenvalues and the numerical search for the optimal transmit powers, we alternatively use the PEP as the criterion to design the optimal power allocation schemes. Some PEP results with imperfect CSI at the receiver have been obtained in the previous chapters, which can be easily computed.

6.4.2.1 The Signal Model with Transmit Weights

By focusing on adapting the transmit power to the channel conditions, we convert the problem into one of designing the optimal transmit weighting matrix at the transmitter. Thus, the signal model in (6.2) can be rewritten as

$$\mathbf{r}(k) = \sqrt{E_s}\mathbf{H}(k)\mathbf{W}(k)\mathbf{v}(k) + \mathbf{n}(k) \tag{6.32}$$

where E_s is the average symbol energy, and each element in the transmit codevector $\mathbf{v}(k)$ has unit energy. $\mathbf{W}(k) = diag[w_1 \cdots w_{N_T}]$ is the diagonal transmit weighting matrix with Frobenius norm $\|\mathbf{W}(k)\|^2 = 1$. Suppose the entire sequence of transmitted and received signals for the user concerned are $\mathbf{v} = [\mathbf{v}^T(1) \cdots \mathbf{v}^T(K)]^T$ and $\mathbf{r} = [\mathbf{r}^T(1) \cdots \mathbf{r}^T(K)]^T$, respectively. The combined weighting matrix and channel fading matrix are $\mathbf{W} = diag[\mathbf{W}(1) \cdots \mathbf{W}(K)]$ and $\mathbf{H} = diag[\mathbf{H}(1) \cdots \mathbf{H}(K)]$, respectively. The estimated channel fading matrix $\hat{\mathbf{H}} = diag[\hat{\mathbf{H}}(1) \cdots \hat{\mathbf{H}}(K)]$ is obtained based on the channel measurements $\mathcal{I}$. The elements in $\hat{\mathbf{H}}(k) = [\hat{h}_{ij}(k)]$ are independent, $\mathcal{CN}(0, 2\hat{\sigma}_j^2)$ distributed. For this semi-i.n.i.d. case, the ML receiver, as shown in (5.28), detects the transmitted signal sequence $\mathbf{v}$ by computing the decision metric $m(\mathbf{r}, \mathbf{v}) = \|\mathbf{r} - \sqrt{E_s}\hat{\mathbf{H}}\mathbf{W}\mathbf{v}\|^2$ for each possible value $\mathbf{v}_l$ of $\mathbf{v}$, and deciding that $\mathbf{v} = \mathbf{v}_j$ if $m(\mathbf{r}, \mathbf{v}_j) = \min_l m(\mathbf{r}, \mathbf{v}_l)$.

If there is no bandwidth limit on the feedback channel, the channel measurements can be conveyed to the transmitter, and the estimated channel matrix is calculated at the transmitter. Otherwise, for the bandwidth-limited feedback channel,

the estimated channel matrix may not be available at the transmitter. The practical assumption is that only partial or quantized channel state information is known at the transmitter. Recalling that the statistical information of the channel estimates depends only on the channel fading model and the channel estimator used, we thus can use a simple feedback scheme, which conveys the design parameters of the channel estimator, namely the pilot spacing and the Wiener filter length, to the transmitter. As for the channel fading model, it is a long-term statistical information of the channel, which can be measured. It is reasonable to assume that the fading model is known at both the transmitter and receiver as a priori information. This simple feedback scheme requires small bandwidth, since only the design parameters of the channel estimator are fed back. Using this feedback scheme, the variances of the channel estimates can be easily computed via (5.3). Based on the information of the channel estimates at the transmitter, we next propose two simple power allocation schemes.

6.4.2.2 Estimated Channel Matrix Known at Transmitter

We assume here that the knowledge of the estimated channel fading matrix is available at the transmitter. Thus, the conditional PEP $P(\mathbf{v}_c \rightarrow \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c)$ given the channel measurements $\mathcal{I}$, is used as the objective function. This conditional PEP that the receiver decides in favor of an erroneous codeword $\mathbf{v}_e = [\mathbf{v}_e^T(1) \ \cdots \ \mathbf{v}_e^T(K)]^T$ when the actual codeword sent is $\mathbf{v}_c = [\mathbf{v}_c^T(1) \ \cdots \ \mathbf{v}_c^T(K)]^T$ and there are no other alternative codewords, is given by (5.18) in Chapter 5,

$$P(\mathbf{v}_c \rightarrow \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) = Q\left(\sqrt{\sum_{k\in\kappa} \frac{E_s\|\hat{\mathbf{H}}(k)\mathbf{W}(k)\mathbf{v}_{ce}(k)\|^2}{2\tilde{N}_0(k)}}\right) \tag{6.33}$$

where $\mathbf{v}_{ce}(k) = \mathbf{v}_c(k) - \mathbf{v}_e(k)$ is the code difference vector, and κ is the set of frame indices k where $\mathbf{v}_c(k) \neq \mathbf{v}_e(k)$. Here, $\mathbf{W}(k)\mathbf{v}_{ce}$ can be viewed as the effective code difference vector. The power of the effective noise at each receive antenna is given by

$$\tilde{N}_0(k) = N_0 + 2E_s \sum_{j=1}^{N_T} \bar{\sigma}_j^2 |w_j(k)|^2 \tag{6.34}$$

If the same power is allocated to each transmit antenna, i.e., $w_j(k) = 1$ for all $j = 1, \cdots, N_T$, equations (6.33) and (6.34) will reduce to the same as those obtained for the open-loop system in (3.26) and (3.24). With the estimated channel matrix known at the transmitter, the optimal transmit weighting matrix can be obtained by minimizing the conditional PEP in (6.33). Define

$$\mathbf{w} = [w_1 \cdots w_{N_T}]^T \tag{6.35a}$$

$$\bar{\Phi} = diag[\bar{\sigma}_1^2, \bar{\sigma}_2^2, \cdots, \bar{\sigma}_{N_T}^2] \tag{6.35b}$$

$$\Psi = \frac{N_0}{2E_s}\mathbf{I}_{N_T} + \bar{\Phi} \tag{6.35c}$$

where $\mathbf{w}$ is the equivalent transmit weighting vector. Then, the matrix form of $\tilde{N}_0(k)/(2E_s)$ is given by

$$\tilde{N}_0(k)/(2E_s) = \mathbf{w}^H(k)\Psi\mathbf{w}(k) \tag{6.36}$$

Due to the monotonically decreasing property of the Q-function, the problem of interest can be formulated into the following constrained optimization problem:

$$\begin{aligned} \text{Find } \arg\max_{\mathbf{w}(k)} f(\mathbf{w}(k)) &= \frac{\mathbf{v}_{ce}^H(k)\mathbf{W}^H(k)\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)\mathbf{W}(k)\mathbf{v}_{ce}(k)}{\mathbf{w}^H(k)\Psi\mathbf{w}(k)} \\ \text{subject to} \quad & \sum_{j=1}^{N_T} |w_j(k)|^2 = 1 \end{aligned} \tag{6.37}$$

Note that the optimal weighting matrix obtained by using this objective function is dependent on the code difference vector. This is impractical to implement. Thus, we use an lower bound on the conditional PEP to obtain the optimal weighting matrix. Replacing the symbol-wise Euclidean distance $|v_{cej}(k)|$, which is the jth element of $\mathbf{v}_{ce}(k)$, by its maximum value d_{max} on the right hand side of (6.33), we have

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathcal{I}, \mathbf{v} = \mathbf{v}_c) \geq Q\left(\sqrt{\sum_{k\in\kappa} \frac{E_s d_{max}^2 \|\hat{\mathbf{H}}(k)\mathbf{w}(k)\|^2}{2\tilde{N}_0(k)}}\right) \tag{6.38}$$

where $d_{max} = 2$ for MPSK schemes with unit energy. Using the PEP lowerbound in (6.38) and dropping the constants, the optimization problem can be reduced to the following:

$$\text{Find } \arg\max_{\mathbf{w}(k)} f(\mathbf{w}(k)) = \frac{\mathbf{w}^H(k)\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)\mathbf{w}(k)}{\mathbf{w}^H(k)\Psi\mathbf{w}(k)} \tag{6.39}$$

Since Ψ is a positive definite matrix, it is known from [87, eq. (2.4.9)] that

$$\lambda'_{N_T} \leq \frac{\mathbf{w}^H(k)\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)\mathbf{w}(k)}{\mathbf{w}^H(k)\Psi\mathbf{w}(k)} \leq \lambda'_1 \tag{6.40}$$

where λ'_1 and λ'_{N_T} are the largest and smallest eigenvalues of $\Psi^{-1}\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)$ with normalized eigenvectors $\mathbf{q}_1$ and $\mathbf{q}_{N_T}$, respectively. The maximum $f(\mathbf{w})$ can thus be obtained at

$$\mathbf{w}_o|_{\hat{\mathbf{H}}(k)} = \mathbf{q}_1 \, . \tag{6.41}$$

This is the optimal power allocation scheme based on the PEP lower bound with the estimated channel fading matrix known at the transmitter. With imperfect CSI at the receiver, both the estimated channel fading matrix $\hat{\mathbf{H}}(k)$, and the matrix Ψ in (6.35c), which depends on the variances of the channel estimation errors and the average SNR, should be used *jointly* to minimize the error performance. When the channel estimates are sufficiently reliable, beamforming is the optimal transmission scheme. Then, the transmitter should transmit signals along the direction of the eigenvector corresponding to the largest eigenvalue of $\Psi^{-1}\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)$, where the estimated channel fading matrix is modified by its corresponding estimation accuracy. Most results on the beamforming scheme assumed that perfect CSI is available at the receiver [28, 36]. No results have been published on the optimal beamforming scheme with imperfect CSI at both the transmitter and receiver. Here, our result clearly shows that with imperfect CSI at the both transmitter and receiver, we cannot simply replace the channel matrix $\mathbf{H}(k)$ by its estimate $\hat{\mathbf{H}}(k)$. The variances of the channel estimation errors should be taken into account.

This power allocation scheme has lower complexity than that based on the capacity lowerbound in the previous section. However, it requires the knowledge of the estimated channel matrix be known at the transmitter. In bandwidth limited feedback systems, the above two schemes may not be suitable. Thus, in the next subsection, we will derive a power allocation scheme based only on the variances of the channel estimates for the bandwidth limited feedback systems, which is simpler to implement.

6.4.2.3 Power Allocation Based on Variances of Channel Estimates

In this case, the estimated channel fading matrix is unknown at the transmitter, and only the variances of the channel estimates are available. The matrix which consists of these variances is defined as

$$\hat{\mathbf{\Phi}} = diag[\hat{\sigma}_1^2, \cdots, \hat{\sigma}_{N_T}^2]. \tag{6.42}$$

The average PEP is used here as the metric. It can be obtained by averaging the conditional PEP in (6.33) over all realizations of the channel fading matrix, and is given by

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{k \in \kappa} \left(\frac{\sin^2 \theta}{\sin^2 \theta + \frac{E_s}{2\tilde{N}_0(k)} \sum_{j=1}^{Nt} \hat{\sigma}_j^2 w_j^2 |v_{cej}|^2} \right)^{N_R} d\theta \,. \tag{6.43}$$

Using the lower bound on the PEP derived in (5.26), and replacing $|v_{cej}|$ by its maximum value d_{max}, $\tilde{N}_0(k)$ by its expression in (6.36), the lower bound on (6.43) can be shown to be

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) \;\; > \;\; \frac{1}{2} \frac{(2q-1)!!}{2q!!} \prod_{k \in \kappa} \left(1 + \frac{E_s d_{max}^2 \mathbf{w}^H(k) \hat{\mathbf{\Phi}} \mathbf{w}(k)}{2\tilde{N}_0(k)} \right)^{-N_R} \tag{6.44}$$

where $q = N_R \times |\kappa|$ is the total diversity order achieved over rapid fading channels, $(2q)!! = 2q(2q-2)\cdots 2$, and $(2q-1)!! = (2q-1)(2q-3)\cdots 1$, as defined in Chapter 5. By minimizing the lowerbound on the average PEP, we can obtain the optimal transmit weighting matrix. Thus, we have the following optimization problem:

$$\arg\max_{\mathbf{w}} f(\mathbf{w}) = \frac{\mathbf{w}^H(k) \hat{\mathbf{\Phi}} \mathbf{w}(k)}{\mathbf{w}^H(k) \mathbf{\Psi} \mathbf{w}(k)} \tag{6.45}$$

$$\text{subject to} \quad \sum_{j=1}^{N_T} |w_j(k)|^2 = 1 \tag{6.46}$$

We can easily obtain the optimal $\mathbf{w}_o$ given $\hat{\mathbf{\Phi}}$ as

$$\mathbf{w}_o|_{\hat{\mathbf{\Phi}}} = \mathbf{q}_1' \tag{6.47}$$

where $\mathbf{q}_1'$ is the normalized eigenvector corresponding to the largest eigenvalue of $\mathbf{\Psi}^{-1}\hat{\mathbf{\Phi}}$. With the definition of $\mathbf{\Psi}$ and $\hat{\mathbf{\Phi}}$ given in (6.35c) and (6.42), we have

$$\mathbf{\Psi}^{-1}\hat{\mathbf{\Phi}} = diag[\zeta_1, \cdots, \zeta_{N_T}] \tag{6.48}$$

where

$$\zeta_j = \frac{\hat{\sigma}_j^2}{\bar{\sigma}_j^2 + N_0/2E_s} \tag{6.49}$$

for $j = 1, \cdots, N_T$. Compared with the optimal weighting vector $\mathbf{w}_o|_{\hat{\mathbf{H}}}(k)$ in (6.39), the matrix $\hat{\mathbf{H}}^H(k)\hat{\mathbf{H}}(k)$ is replaced by $\hat{\mathbf{\Phi}}$ when only the variances of the channel estimates are available at the transmitter. Due to the fact that the matrix $\mathbf{\Psi}^{-1}\hat{\mathbf{\Phi}}$ in (6.48) is diagonal, the eigenvalues of $\mathbf{\Psi}^{-1}\hat{\mathbf{\Phi}}$ are equal to the corresponding diagonal values. With the beamforming scheme, the transmit antenna which has the largest ζ_j is assigned the total power. From the expression of ζ_j in (6.49), it is seen that the choice of the transmit antenna with the best channel condition depends not only on the corresponding variance of the channel estimate $\hat{\sigma}_j^2$, but also on the effective noise caused by the estimation error from that transmit antenna, i.e. $N_0 + 2E_s\bar{\sigma}_j^2$. Our result is the first one to show explicitly the effect of both channel estimation accuracy and the effective noise on the power allocation. Another advantage of this scheme is that the matrix $\mathbf{\Psi}^{-1}\hat{\mathbf{\Phi}}$ is only dependent on the variances of the channel estimates and those of the estimation errors. Thus, the transmit weighting vector can be precomputed once the design parameters, namely, the pilot spacing and the Wiener filter length, are known at the transmitter.

6.5 Pilot Power Allocation Schemes

After examining the transmit power allocation scheme, we will investigate the design of the pilot power allocation schemes. Given a fixed average transmission power, the problem of how to allocate the power between the data symbols and the pilot symbols is an important research topic. The power allocated to the pilot symbols directly affects the channel estimation accuracy and the error performance. A good pilot power allocation scheme can enhance performance at no additional cost of bandwidth and power. Due to the power limitation in wireless communications, pilot power allocation to improve power efficiency has practical importance. In this chapter, we model the relationship between the power of the data and the pilot

symbols by using one parameter. The goal of the pilot power allocation scheme is to find the optimal parameter by minimizing the error performance.

In the previous section, we have discussed several transmit power allocation schemes based on different criteria. Either the estimated channel matrix or the variances of the channel estimates are assumed to be known at the transmitter. The accuracy of the channel estimates is highly related to the power of the pilot symbols. In this section, we will examine how to optimally distribute the power between the data symbols and the pilot symbols. The average BEP performance is used as the design criterion, and no feedback information is needed to determine the pilot power allocation policy. As indicated in the signal models (5.1) and (5.2), we assume that the energy for all pilot symbols is the same, which is denoted by E_p. Similarly, the energy for all data symbols is E_s. However, E_p and E_s could be different. We model the relationship between E_p and E_s by

$$E_p = \alpha E_s \tag{6.50}$$

where α is the pilot-to-data energy ratio. Based on the transmission frame structure, the average energy per symbol is

$$\begin{aligned} E &= \frac{N_T E_p + (L - N_T)E_S}{L} = \frac{(L + (\alpha - 1)N_T)E_S}{L} \\ &= [1 + \rho(\alpha - 1)]E_s \end{aligned} \tag{6.51}$$

where $\rho = N_T/L$. Given the average energy per symbol E, we can easily obtain the pilot symbol energy E_p and the data symbol energy E_s as

$$E_s = \frac{1}{[1 + \rho(\alpha - 1)]}E \tag{6.52}$$

$$E_p = \frac{\alpha}{[1 + \rho(\alpha - 1)]}E\,. \tag{6.53}$$

Our pilot power allocation scheme is to find the optimal α to achieve the minimum error probability, subject to the constraint of a given average SNR per symbol per receive antenna E/N_0. Here, we uses the BEP upperbound obtained by using the dominant error events as the objective function to determine the optimum value of α. Thus, the optimization problem is outlined as

Minimize the BEP upperbound $f(\alpha)$, i.e.,

$$f(\alpha) = \frac{1}{n} \sum_{\mathbf{v}_c} \sum_{\mathbf{v}_e} a(\mathbf{v}_c \to \mathbf{v}_e) P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) P(\mathbf{v}_c) \tag{6.54}$$

where the probability of each codeword $P(\mathbf{v}_c)$ is assumed to be the same, and the PEP $P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v}_c)$ is upperbounded by (5.35)

$$P(\mathbf{v}_c \to \mathbf{v}_e | \mathbf{v} = \mathbf{v}_c) \leq \frac{1}{2} \prod_{k \in \kappa} \prod_{i=1}^{N_R} \left(1 + \frac{\sum_{j=1}^{N_T} \hat{\sigma}_j^2 d_j^2(k)}{2 \left(\frac{N_0}{E} [1 + \rho(\alpha - 1)] + 2 \sum_{j=1}^{N_T} \bar{\sigma}_j^2 \right)} \right)^{-1} \tag{6.55}$$

Subject to

$$E/N_0 = \bar{\gamma}$$

where $\bar{\gamma}$ is the given average SNR per symbol per receive antenna.

We can use the optimization toolbox in Matlab to obtain the optimal value of α, which provides the minimum BEP performance for a given average SNR $\bar{\gamma}$. Given the error events, $d_j^2(k)$'s in (6.55) are easy to obtain. The channel estimate variance $\hat{\sigma}_j^2$ and the channel estimation error variance $\bar{\sigma}_j^2$ can be computed based on the channel fading model and the estimator employed. Thus, the objective function $f(\alpha)$ only depends on α. The optimal value for α, α_o, varies with $\bar{\gamma}$. However, our computations show that α_o is actually not very sensitive to $\bar{\gamma}$. Thus, we use the value of α_o obtained at $\bar{\gamma} = 15\ dB$ as the reference. In Tab. 6.1, we listed the optimum α_o for two 8-state QPSK STTC over rapid fading channels. Our numerical results in the next section show that about 1 dB performance gain can be achieved by using the optimum pilot power allocation scheme.

6.6 Numerical Results and Discussion

In the numerical results, we assume that the channel fading process on each link has a Jake's power spectrum so that $R(\tau) = J_0(2\pi f_d \tau)$. The variance of the fading

Table 6.1: The optimal α_o for the QPSK 8-state TSC code of [1] and FVY code of [2] over rapid fading channels.

α_o	TSC	FVY
$\sigma_1^2 = 0.5, \sigma_2^2 = 0.5$	2.4391	2.4390
$\sigma_1^2 = 0.3, \sigma_2^2 = 0.7$	2.434	2.4349
$\sigma_1^2 = 0.1, \sigma_2^2 = 0.9$	2.4235	2.4226

processes associated with the jth antenna is denoted by σ_j^2. The PSAM scheme is used to estimate the channels with pilot spacing $L = 8$ and the Wiener filter length $N = 6$, for all channel fade rates that are lower than $f_dT = 0.05$. Without loss of generality, we consider the user whose time slot is located immediately after the pilot phase in the first block.

We first compare the capacity lower bound for the case of imperfect CSI at the receiver with that of perfect CSI. Either the optimal transmit power allocation scheme based on the capacity lower bound, or the constant power scheme that assigns equal power to each transmit antenna, is used. Fig. 6.3 and 6.4 present the capacity lower bounds with *fixed* E_p for semi-i.n.i.d. channels. All the results show that there is capacity loss due to the imperfect CSI at the receiver. Capacity gain can be achieved with the optimal transmit power allocation, especially at low SNR. At high SNR, the capacity lower bound with constant power allocation is shown to be very close to that with optimal power allocation. Since the power of pilot symbols E_p is fixed, the variances of the channel estimates are determined. The increase in average SNR of data symbols P/N_0 leads to a greater effective noise $\overline{N}_0 = 2\sum_{j=1}^{N_T} P_j\bar{\sigma}_j^2 + N_0$. Thus, the capacity bound for the imperfect CSI case does not increase linearly with the average SNR of data symbols after a certain threshold. It is also observed that the capacity lower bound with imperfect CSI for the channel with $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$ is less than that with $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$ for the same fade rate of $f_dT = 0.05$. Therefore, the capacity loss is increased with the difference among the variances of the channel fading coefficients associated with different transmit antennas. In Fig. 6.5 and 6.6, the capacity lower bounds with

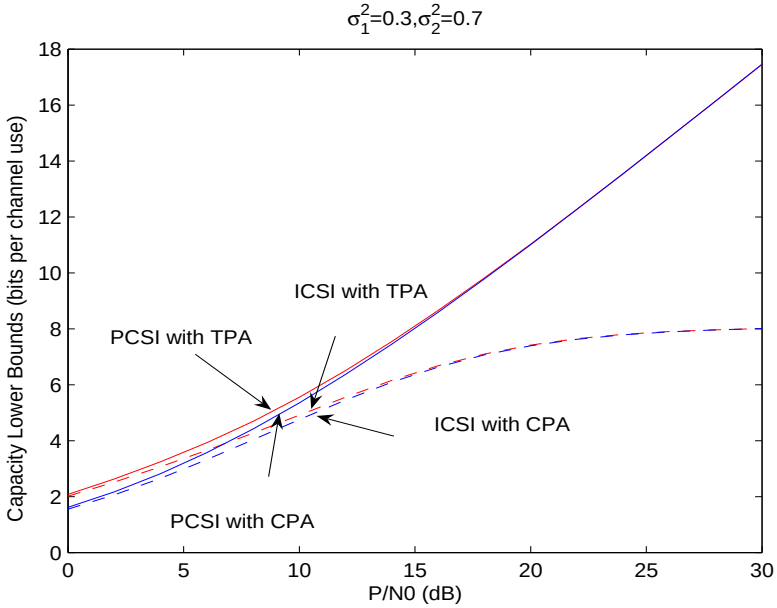


Fig. 6.3: The capacity lower bound with optimal transmit power allocation (TPA) in (6.30) or constant power allocation (CPA) for both the perfect CSI and imperfect CSI cases. The channels with two transmit and two receive antennas have $\sigma_1^2 = 0.3, \sigma_2^2 = 0.7$ *and* $f_d T = 0.05$*, which are estimated with* $L = 8$*,* $N = 6$ *at* $E_p/N_0 = 15$ *dB.*

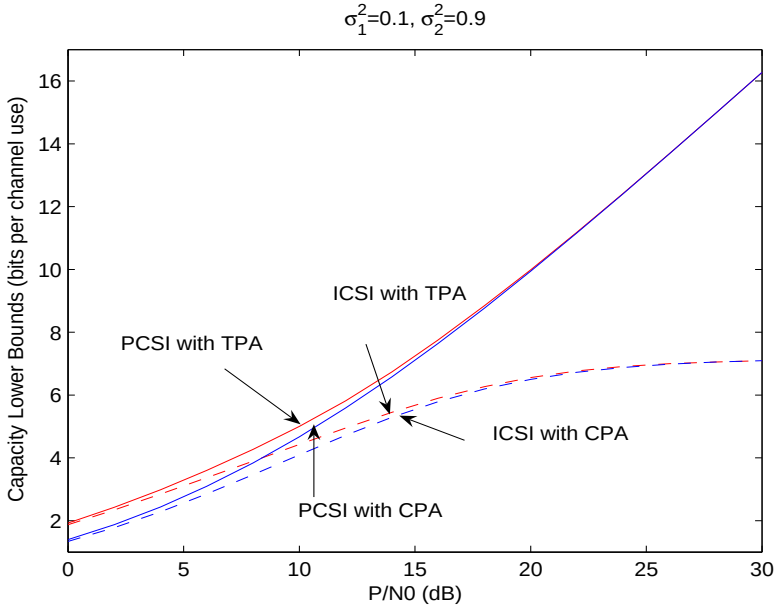


Fig. 6.4: The capacity lower bound with optimal transmit power allocation (TPA) in (6.30) or constant power allocation (CPA) for both the perfect CSI and imperfect CSI cases. The channels with two transmit and two receive antennas have $\sigma_1^2 = 0.1, \sigma_2^2 = 0.9$ *and* $f_d T = 0.05$*, which are estimated with* $L = 8$*,* $N = 6$ *at* $E_p/N_0 = 15$ *dB.*

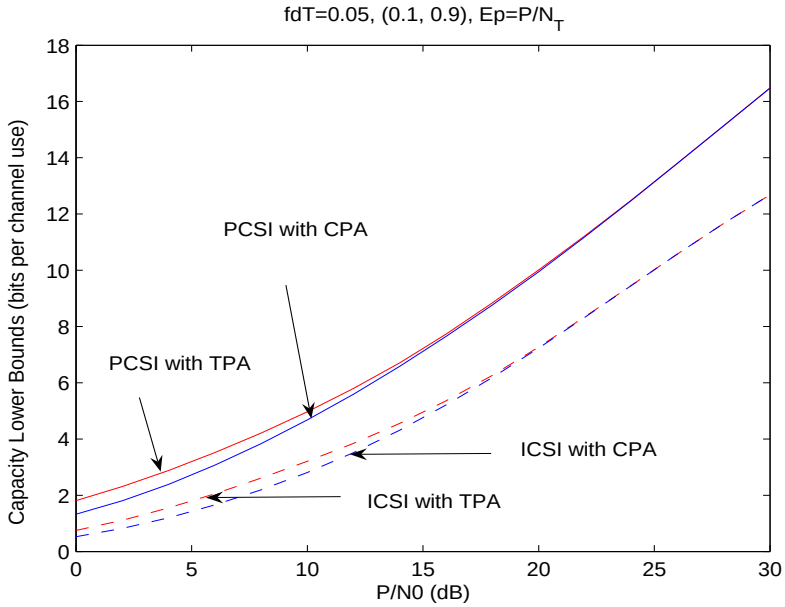


Fig. 6.5: The capacity lower bound with optimal transmit power allocation (TPA) in (6.30) or constant power allocation (CPA) for both the perfect CSI and imperfect CSI cases. The channels with two transmit and two receive antennas have $\sigma_1^2 = 0.1, \sigma_2^2 = 0.9$ *and* $f_dT = 0.05$*, which are estimated with* $L = 8$*,* $N = 6$ *at* $E_p = P/N_T$*.*

$E_P = P/N_T$ are presented. It can be seen that the capacity loss is larger for the channel with fade rate of 0.05 than for the channel with fade rate of 0.01. The larger capacity loss, either for the case of higher channel fade rates, or for the case of larger variance differences, can be explained by the accuracy of channel estimation. It has been shown by Fig. 5.1 in Chapter 5 that the variances of the channel estimation errors increase with the channel fade rate and the variance differences of the channel fading coefficients associated with the different transmit antennas.

We use the power allocation gain (PAG) to show the effectiveness of the proposed allocation schemes. PAG is computed as

$$PAG = \frac{f^*(P1 \cdots P_{N_T})}{f^{(0)}(P1 \cdots P_{N_T})}$$

where $f(P1 \cdots P_{N_T})$ can be the capacity lowerbound (6.23), or the functions inversely proportional to the conditional PEP lowerbound (6.39) and the average PEP lowerbound (6.45), respectively. Thus, the larger the value of $f(P1 \cdots P_{N_T})$,

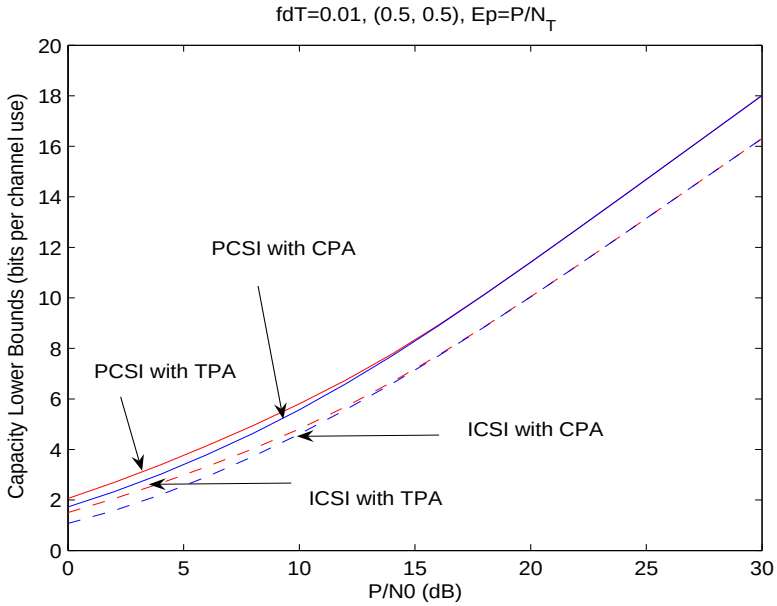


Fig. 6.6: The capacity lower bound with optimal transmit power allocation (TPA) in (6.30) or constant power allocation (CPA) for both the perfect CSI and imperfect CSI cases. The channels with two transmit and two receive antennas have $\sigma_1^2 = 0.5, \sigma_2^2 = 0.5$ and $f_dT = 0.01$, which are estimated with $L = 8$, $N = 6$ at $E_p = P/N_T$.

the better the performance that can be achieved for all the three criteria. The PAG of each scheme is obtained by using the corresponding expression of $f(P1 \cdots P_{N_T})$. Note that $f^*(P1 \cdots P_{N_T})$ is attained by distributing the total available power over all transmit antennas using the optimal allocation scheme, and $f^{(0)}(P1 \cdots P_{N_T})$ by *evenly* distributing the total power over all transmit antennas. Since the function $f(P1 \cdots P_{N_T})$ used is related to the lower bound on the capacity or the PEP, the PAG achieved by the optimal allocation scheme derived by optimizing $f(P1 \cdots P_{N_T})$, thus, depends on the tightness of the bound. Furthermore, PAG also depends on operating SNR and the variances of the channel estimates [80]. When the variances of the channel estimates or the operating SNR is high, the PAG obtained approaches unity. The PAG's shown in our numerical results are obtained by averaging over a large number of realizations of $\hat{\mathbf{H}}(k)$. When $PAG > 1$, it means exploiting the feedback at the transmitter is useful. When PAG approaches unity, the feedback information provides no gains.

Fig. 6.7 presents the PAG attained by the optimal power allocation scheme

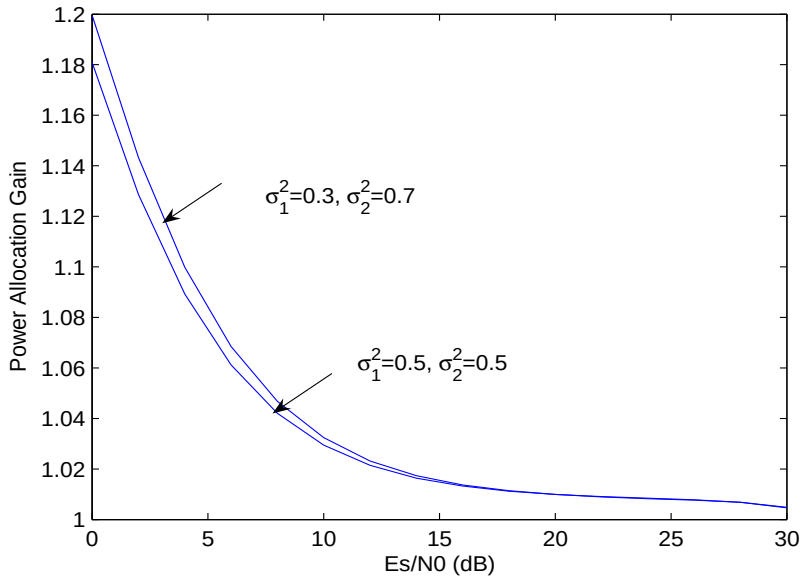


Fig. 6.7: The power allocation gain achieved by power allocation scheme based on the capacity lower bound in (6.30). The channels have a fade rate of $f_dT = 0.05$, which are estimated with $L = 8$, $N = 6$ at $E_p/N_0 = 15$ dB.

in (6.30) based on the capacity lower bound. The total available power is P. Both i.i.d. and i.n.i.d. channels are examined. It can be seen that PAG is high for low SNR and approaches unity with the increase of SNR. The reason is that capacity is a logarithmic function of power, and is insensitive to the exact power allocation at high SNR. Thus, the constant-power allocation concept, i.e., the active transmitter antennas are assigned the same power, can also be adopted for the channel estimation scenario. It is also found that the PAG attained over i.n.i.d. channels is greater than that over i.i.d. channels at low SNR.

Fig. 6.8 compared the two power allocation schemes based on PEP bounds, i.e., (6.41) and (6.47), respectively. The results show that $PAG > 1$ can be obtained, and the PAG obtained by exploiting the estimated channel fading matrix is greater than that when only the variances of the channel estimates are known at the transmitter. Obviously, there is a tradeoff between the performance gain and the complexity of the feedback implementation.

Fig. 6.9 shows both analytical and simulated BEP performance results for

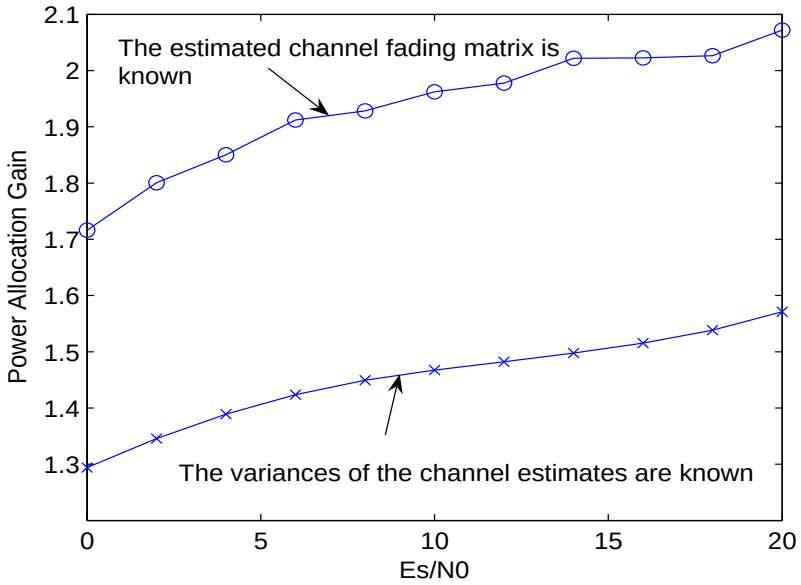


Fig. 6.8: The comparison of the power allocation gains achieved by schemes based on PEP given the knowledge of CSIT in (6.41) and (6.47), respectively. The channels are i.i.d. with $\sigma_1^2 = 0.5$, $\sigma_2^2 = 0.5$, *and* $f_dT = 0.05$, *which are estimated with* $L = 8$, $N = 6$ *at* $E_p/N_0 = 15$ *dB.*

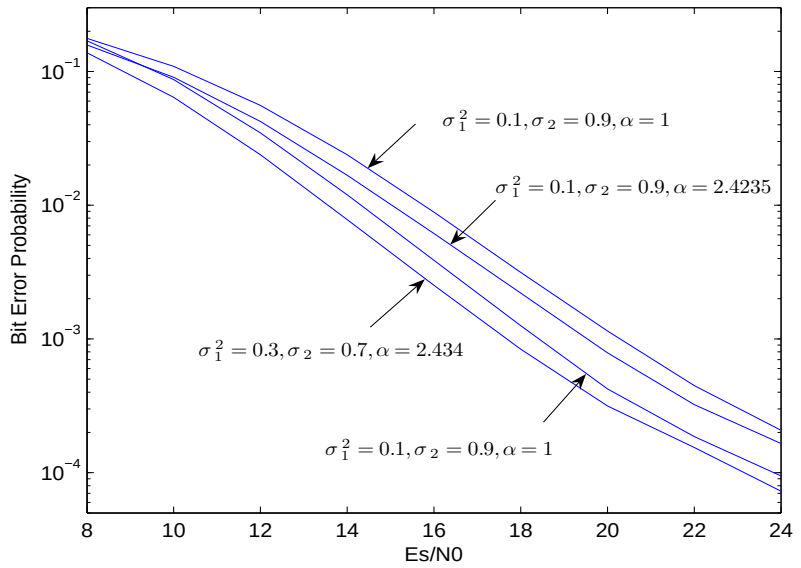


Fig. 6.9: The simulated BEP results for the 8-state QPSK code of [1] over i.n.i.d. Rayleigh fading channels at $f_dT = 0.05$, *using two transmit and one receive antenna. The channel is estimated with* $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, *and* $E_p = \alpha(E_s)_{PSAM}$.

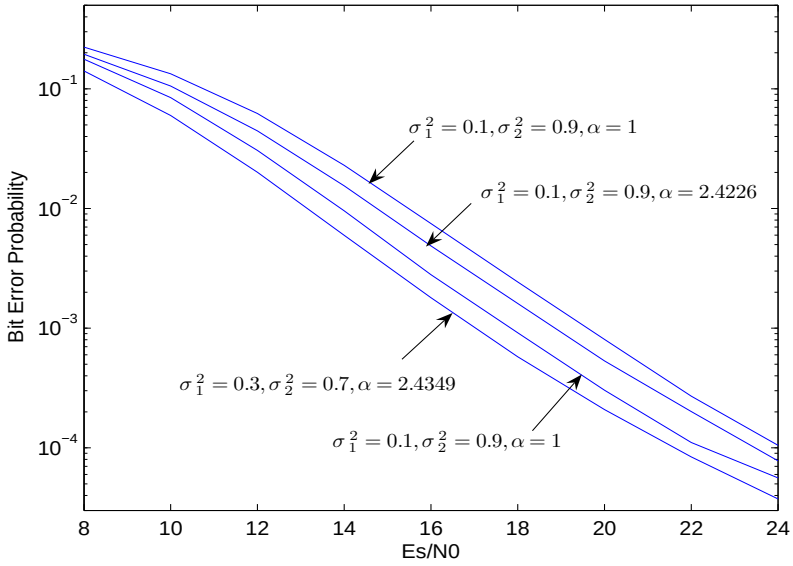


Fig. 6.10: The simulated BEP results for the 8-state QPSK code of [2] over i.n.i.d. Rayleigh fading channels at $f_dT = 0.05$, using two transmit and one receive antenna. The channel is estimated with $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p = \alpha(E_s)_{PSAM}$.

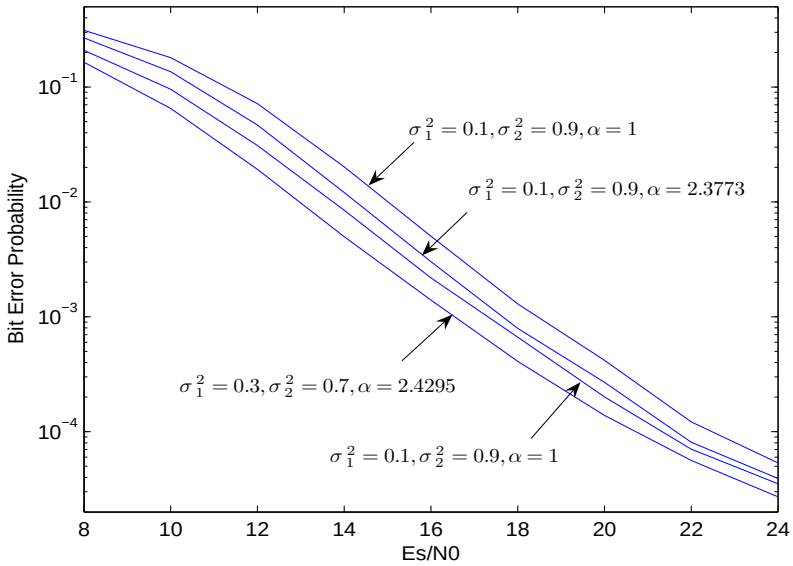


Fig. 6.11: The simulated BEP results for the 8-state ICSI code over i.n.i.d. Rayleigh fading channels at $f_dT = 0.05$, using two transmit and one receive antenna. The channel is estimated with $L = 8$, $N = 6$, $(E_s)_{PSAM} = E_s(L - N_T)/L$, and $E_p = \alpha(E_s)_{PSAM}$.

the 8-state TSC code of [1] over i.n.i.d. fading channels. Two channel situations are considered. One has the variances $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$. The other case has $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$. For both cases, two pilot power schemes are examined. Without using the optimal pilot power allocation scheme, the pilot symbol energy is simply assumed to be equal to the data symbol energy, i.e., $\alpha = 1$. Otherwise, the optimal pilot power allocation is employed by using α_o obtained in Tab. 6.1. There is a performance gain of about 1 dB that can be achieved by the proposed pilot allocation scheme at no extra cost of bandwidth and power. Thus, it is efficient to use the pilot power allocation scheme, especially at low SNR. Similar results can be obtained for the 8-state FVY code of [2], as shown in Fig. 6.10. Besides using the known codes in the literature, we also examine the BEP performance of the proposed codes with imperfect CSI at the receiver for the i.n.i.d. channels. Take the ICSI code designed for the $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$ case, for example, which is given in Tab. 5.1. It is shown in Fig. 6.11 that the performance gain achieved by the ICSI code is similar to the TSC and FVY codes. However, the performance gap between the two channel conditions, i.e., $(\sigma_1^2 = 0.3, \sigma_2^2 = 0.7)$ and $(\sigma_1^2 = 0.1, \sigma_2^2 = 0.9)$, are smaller than that of the TSC and the FVY codes. This shows that our proposed ICSI codes are more suitable for the i.n.i.d. channels, and are robust to the channel variations.

6.7 Summary

In this chapter, we examine power allocation schemes for closed-loop TDM MIMO systems with imperfect CSI at the receiver. First, the closed-loop system model is built up. The estimated channel matrix can be obtained at the transmitter as the side information. The capacity of a MIMO system with side information at the transmitter is studied. A new lower bound on the channel capacity with imperfect CSI at both the transmitter and receiver for semi-i.n.i.d. channels is derived. There is capacity loss due to the channel estimation errors. Based on the

capacity lower bound obtained, an optimal transmit power allocation scheme is designed to approach the capacity. To reduce the complexity of the power allocation scheme, the PEP bounds are alternatively used as the criterion, and the problem is transformed into one of designing the optimal transmit weighting matrix. Based on the side information at the transmitter, different optimal transmit weighting matrices are derived. Our results show that all the proposed transmit power allocation schemes can achieve performance gains, compared to the case with no feedback information at the transmitter. However, there is a tradeoff between the complexity of the feedback scheme and the performance gain. In addition to using transmit power allocation schemes, we also investigated how to optimally distribute the power between the data and pilot symbols. With the power allocation schemes, about 1 dB performance gain can be achieved at no additional cost of bandwidth and power.

Chapter 7

Conclusions and Proposals for Future Research

7.1 Conclusions

In this thesis, we addressed two important, practical topics of STTC for both open-loop and closed-loop systems, where only imperfect CSI is available at the receiver. One is error performance analysis and code design. The other is space-time transmission schemes. For the open-loop system, we investigated the performance analysis of STTC over rapid Rayleigh fading channels with imperfect CSI, and analyzed the effects of channel estimation on the performance. Some useful insights were gained to combat the performance loss caused by channel estimation errors. This also provided a powerful design criterion for the imperfect CSI case. New, improved codes were proposed based on our general design criterion. For the closed-loop system, transmit power allocation schemes were examined to achieve performance gains. These schemes adapted the power at each transmit antenna to the current channel state based on the side information at the transmitter. Further, pilot power allocation schemes were developed to improve the achievable performance of STTC at no additional cost of power and bandwidth. In the following, we first summarize the results for open-loop systems.

7.1.1 Performance Analysis Results

Performance analysis is based on receiver structure. First, the ML receiver structure with channel estimation is derived. It is found that when the MIMO channel is i.i.d., the receiver use the estimated channel matrix as if it were the exact channel matrix, for constant-energy modulation schemes. However, when the channel is i.n.i.d., the receiver also requires the second-order statistical information of channel estimates, in addition to the estimates themselves. The channel estimation accuracy plays an important role in determining the weight on the signals received at each receive antenna. Both the first and second-order statistical information of the channel estimates can be easily computed, once the channel fading model and the channel estimator are given. The receiver for the i.n.i.d. channels is more general and practical, which can be reduced to that of the i.i.d. channels.

After obtaining the ML receiver, we proceed to derive the performance results. Based on the ML receiver, both an exact expression of PEP, explicit PEP upperbounds and a tight BEP upperbound are derived for the general i.n.i.d. fading channels. The corresponding results for i.i.d. channels can be easily obtained from those of the i.n.i.d. channels. Our explicit PEP upperbounds clearly show that the effects of channel estimation can be completely summarized in the variances of the channel estimates and those of the estimation errors. For i.n.i.d. channels, these variances depend on both the transmit and receive antenna index. For i.i.d. channels, they only depend on the transmit antenna index. Under imperfect channel estimation conditions, channel estimation errors increase the power of the effective noise at each receiver antenna, and result in performance degradation. The channel estimation accuracy is determined by the channel fade rate, pilot SNR, and the design parameters of the PSAM scheme at the receiver. The optimal values of the design parameters, namely, the pilot spacing and the Wiener filter length, can be found by the general systematic procedure that we proposed. The estimation accuracy is increased with the increase in pilot SNR, but decreased with the increase in the channel fade rate. The performance loss of STTC caused by channel esti-

mation errors is directly related to the channel estimation accuracy. In contrast to the perfect CSI case, increasing the number of transmit antennas in the imperfect CSI case may lead to a reduced or even negative transmit diversity advantage if the fade rate is high, because of the increased total estimation error variance. Another important finding from our explicit PEP results is that the symbol-wise Euclidean distances between any two codewords are weighted by the unequal variances of the channel estimates among different transmit antennas. Therefore, to minimize the error performance, the symbol-wise Euclidean distances should be optimally distributed based on the statistical information of the channel estimates. This makes the code design of STTC with channel estimation different from previous code designs, which assume that perfect CSI is available at the receiver and each channel has identical distribution.

7.1.2 Code Design of STTC with Channel Estimation

Based on the performance analysis results, a new, practical code design criterion with imperfect CSI is provided. We employ the node error probability as the performance measure, instead of using the traditional PEP as the objective function. More importantly, our new design criterion incorporates statistical information of channel estimates to minimize the performance loss caused by channel estimation errors. It is found that the effect of channel estimation errors on code design can be measured by the maximum estimation variance difference. For i.i.d. channels, this maximum estimation variance difference increases with the channel fade rate and the number of transmit antennas. This implies that with an increase in the channel fade rate or the number of transmit antennas, the codes based on our new criterion can achieve more performance gains than those of existing STTC which are designed under the perfect CSI assumption. For i.n.i.d. channels, the maximum estimation variance difference is mainly related to the differences among the variances of the actual channel fading coefficients associated with different transmit antennas. The advantage of our new code design criterion is more

obvious for i.n.i.d. channels, because of the inherent variance differences among transmit and receive antennas. Because of the practical importance of non-identical statistical distributions among different links, and imperfect CSI at the receiver, it is important to use our proposed design criterion in practice. By exploiting the statistical information of channel estimation, improved performance can be achieved by our new codes. The new codes are obtained by our proposed iterative code search algorithm. In the iterative algorithm, the codewords are grouped based on their error event lengths. At each iterative step, the codewords with the same error event length are considered. At the first step, the error event length is set to its minimum value. The codewords are compared to choose those that have the optimal performance over this minimum error event length. Then, the error event length is increased by one from the previous step, and the same search process is performed. The search keeps running until the set of the optimal code generators obtained does not change for two consecutive steps. At each step, many codes which do not satisfy the design criterion can be discarded. Thus, a smaller set needs to be searched at each step, and the search complexity is reduced.

7.1.3 Power Allocation Schemes

All the above results are for open-loop systems, where no side information is available at the transmitter. In closed-loop systems, side information can be obtained by using feedback schemes. Due to the power limitation in wireless communication, it is very important to use the available power efficiently. In this thesis, we focused on investigating the optimal power adaptation to the current channel state based on the feedback information in closed-loop systems. Three optimal transmit power allocation schemes are proposed based on different criteria, for the case of imperfect CSI at both the transmitter and the receiver. One widely used criterion is the channel capacity. Due to the difficulty of the capacity analysis with only imperfect CSI at the transmitter and receiver, a new capacity lower bound is derived for this purpose. Using the capacity lower bound, an optimal transmit

power allocation scheme is obtained. It is found that when one channel has much better condition than other channels, the beamforming scheme is the optimal one to achieve the channel capacity. The estimated channel matrix known at the transmitter, and the effective noise at the receiver convey the channel condition. The effective noise includes both the additive channel noise and the noise due to the channel estimation errors. To reduce the complexity of the power allocation based on the capacity lower bound, another commonly used error performance criterion, namely, the PEP criterion, is adopted to design the optimal transmit weighting matrix. Different optimal transmit weighting matrices are obtained based on the side information at the transmitter. When the estimated channel matrix is known at the transmitter, the optimal transmit weighting matrix can achieve great performance gains, especially at low SNR. However, it requires that the feedback channel has enough bandwidth. For bandwidth-limited feedback channels, we also proposed a simple feedback scheme where only the variances of the channel estimates are fed back to the transmitter. Then, the optimal transmit weighting matrix is easier to compute, but the performance gain is less than that achieved by the previous case due to the loss of some channel information. Therefore, there is a tradeoff between the complexity of the feedback scheme and the performance gain. All these schemes can be used to achieve performance gains over those that distribute the same power to all transmit antennas. Using the optimal transmit allocation schemes at the transmitter, more transmission power is assigned to the channels with better channel condition. In addition to using transmit weight matrix, the power allocation between the pilot and the data symbols is also exploited. By optimally allocating the pilot power based on minimizing the BEP upperbound, we can achieve performance enhancement for MIMO systems at no cost of additional bandwidth and power.

7.2 Proposals for Future Research

As shown in this thesis, it is important to consider the actual imperfect CSI case in practical systems. In addition to the performance analysis and code design, the transmission schemes are also affected by the imperfect channel estimates. Although our analysis is based on the Rayleigh fading model, the principal ideas can be straightforwardly extended to other fading models. Recently, the selection diversity-based MIMO systems have received a great deal of research attention since the antenna selection schemes can reduce complexity and maintain the advantage of full diversity when perfect CSI is available at both the transmitter and receiver [88]. However, when only imperfect CSI is available, the research on transmit/receive antenna selection is not much. Thus, this area deserves further exploration. The performance analysis results obtained in this thesis can be applied to help the study of antenna selection by incorporating imperfect CSI. After extensive investigation of MIMO in physical layer, research interests are now shifting to MIMO in wireless networks. Cross-layer design [89] and cooperative diversity [90] are the two hot research topics. How to develop the corresponding physical layer signaling schemes for wireless networks is very important and interesting. In the following sections, we present some ideas that we feel ought to be addressed in future research.

7.2.1 Other Fading Models

The Rayleigh fading channel, which is one of the widely used channel model for land mobile terrestrial systems, has been examined in this thesis for the imperfect CSI case. For a Rayleigh fading channel, the channel impulse response is modeled as a zero-mean, complex, Gaussian process. When there exists line-of-sight paths, the channel impulse response can no longer be modeled as having zero-mean. In this case, the envelope of the response has a Rice distribution and the channel is said to be a Rician fading channel [91]. The Nakagami-m fading model is a more general distribution, which provides greater flexibility in matching experimental

data collected in a variety of fading environments [91]. Thus, it is straightforward to extend our current work from the Rayleigh to the Rician or Nakagami-m fading model. In fact, the performance analysis of space-time coding with Nakagami-m fading model is only the topic of recent publications, even for the perfect CSI case [92–95]. There are no results yet on the performance of space-time coding over Nakagami fading channels with imperfect CSI. A lot of open problems still remain in this area. Following the procedures in this thesis, the design of the ML receiver, the performance analysis and the computation of the capacity of the rapid fading channel model for the more general statistical fading distributions with imperfect CSI can be addressed.

7.2.2 Transmit Antenna Selection

Using selection diversity-based techniques, a *subset* of the available antennas are selected at transmitter and/or receiver based on some criteria, such as the capacity, average SNR or error probability. These techniques not only cut the implementation cost by reducing the number of RF chains, but also achieve the full diversity as all antennas are used [88]. Since the increase in the number of antennas at the receiver leads to larger combining gains, much research effort has been devoted to transmit antenna selection with maximal ratio combing (MRC) at the receiver [96–98]. With selection combining (SC) scheme, a *single* transmit antenna is selected and MRC is performed at the receiver. By choosing the transmit antenna with the highest channel power, this scheme can maximize the total received signal power at the receiver. If the feedback information is perfect, the SC/MRC scheme is the optimal. However, when the feedback a transmitter receives experiences a time-delay, the current best antenna may have changed at the moment when the transmitter receives the feedback after the delay. Under this nonideal transmit antenna selection situation, more than one transmit antenna should be selected to increase the degree of robustness against feedback errors [98]. Selection combing with space-time coding is one way to solve the above-mentioned problem. Either

SC-STBC or SC-STTC can be used. The selection criterion and performance analysis of transmit antenna selection schemes mostly focus on the slow, i.i.d. fading model, and assume perfect CSI at both the transmitter and the receiver. STBC using transmit antenna selection with imperfect CSI was studied in [99]. The results for STTC using transmit antenna selection with imperfect CSI are few. Based on the receiver structure derived in Chapter 5 for the general i.n.i.d. channels, we can further study the performance of STTC with transmit antenna selection. It is expected that the imperfect CSI at the transmitter and the receiver will affect the selection criterion. With the performance analysis results, the investigation of the optimal number of the active transmit antennas can proceed. Here, the imperfect CSI considered is caused by the channel estimation errors at the receiver. For some cases, the imperfect CSI at the transmitter may be due to the quantized feedback information [100]. If the quantization errors can be modeled properly, a general framework for the performance analysis with imperfect CSI can be established in future works.

7.2.3 MIMO Wireless Networks

As is well known, the layered-design is not efficient for wireless networks due to the nature of mobile channels [89, 101]. For wireless communications, channel quality changes dynamically, which leads to a tight interdependence among layers. Thus, cross-layer design is often considered [102–104], in which the overall system performance can be improved by taking advantage of the available information across different layers. If the instantaneous CSI is known, the channel-aware scheduling at the link layer can be used to improve the system throughput. However, again, the effect of imperfect CSI on the channel-aware scheduling is still unclear. Therefore, the performance analysis results obtained in this thesis can be used to facilitate the cross-layer design for wireless networks by providing the actual physical layer information.

MIMO channels promise an increase in capacity only if independent signals

are present at the antenna elements. Naturally, physical limitations within the mobile terminal will lead to mutual correlation among the elements, which results in reduced MIMO capacity [105]. A novel solution to overcome this problem is user cooperative diversity [90]. Through cooperation of in-cell users, a new form of spatial diversity can be achieved by the use of other users' antennas. The antennas of the cooperative users create virtual MIMO channels [106]. For the downlink transmission, a base station array consisting of several antenna elements transmits a space-time encoded data stream to the associated mobile terminals. Each mobile terminal within a group receives the entire data stream, extracts its own information and concurrently relays the information of other users to their mobile terminals. It then receives more of its own information from the surrounding mobile terminals and, finally, processes the entire data stream. In this distributed MIMO system, the statistics of the fading coefficients on different virtual links are likely to be quite different, since the users may be separated far away. Therefore, our performance analysis and code design for i.n.i.d. fading channels with imperfect CSI can be further explored in the distributed MIMO systems. The optimal cooperative coding scheme based on the error performance need be carefully addressed. For the uplink transmission, each user not only transmits his own information, but also the information of his partners. However, this is complicated by the fact that the interuser channel is noisy [90]. How to track the fading coefficients of the interuser channel, and how to decode the information for each user with imperfect knowledge of the fading parameters have not yet been examined. The channel estimation techniques and system designs deserve further investigation in the user cooperative wireless networks.

Bibliography

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

[2] W. Firmanto, B. Vucetic, and J. Yuan, "Space-time tcm with improved performance on fast fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 154–156, Apr. 2001.

[3] J. Y. Z. Chen, B. S. Vucetic and K. L. Lo, "Space-time trellis codes for 4-psk with three and four transmit antennas in quasi-static flat fading channels," *IEEE Commun. Lett.*, vol. 6, pp. 67–69, Feb. 2002.

[4] J. Winters, "On the capacity of radio communication systems with diversity in a rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 871–878, June 1987.

[5] G. J. Foschini, "Layered space-time architecutre for wireless communication in fading environments when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.

[6] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, pp. 585–596, Nov. 1999.

[7] J. Yuan, Z. Chen, B. Vucetic, and W. Firmanto, "Performance and design of space-time coding in fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 1991–1996, Dec. 2003.

[8] M. Tao and R. S. Cheng, "Improved design criteria and new trellis codes for space-time coded modulation in slow flat fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 313–315, July 2001.

[9] J. Zhang, Y. Qiang, J. Wang, and D. Li, "On the design of space-time code for fast fading channels," *in Proc. PIMRC*, vol. 2, pp. 1045–1048, Sept. 2003.

[10] Q. Yan and R. S. Blum, "Improved space-time convolutional codes for quasi-static slow fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 563–571, Oct. 2002.

[11] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Select. Area Comm.*, vol. 16, pp. 1451–1458, Oct. 1998.

[12] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Info. Theory*, vol. 45, pp. 1456–1466, July 1999.

[13] V. Tarokha, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communciations: performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–459, Mar. 1999.

[14] M. Tao and R. S. Cheng, "Diagonal block space-time code design for diversity and coding advatange over flat fading channels," *IEEE Trans. Signal Process.*, vol. 52, pp. 1012–1020, Apr. 2004.

[15] Z. Safar and K. J. Liu, "Space-time trellis code construction for fast fading channel," *in Proc. ICC*, vol. 1, pp. 563–246, May 2002.

[16] G. D. Golden, G. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky, "Detection algorithm and initial laboratory results using the v-blast space-time communication architecture," *Electronics Letters*, vol. 35, pp. 14–15, Jan. 1999.

[17] H. E. Gamal and R. Hammons, "A new approach to layered space-time coding and signal processing," *IEEE Trans. Info. Theory*, vol. 47, pp. 2321–2334, Sept. 2001.

[18] G. Caire and G. Colavolpe, "On low-complexity space-time coding for quasi-static channels," *IEEE Trans. Info. Theory*, vol. 49, pp. 1400–1416, June 2003.

[19] C. B. Papadias and G. J. Foschini, "Capacity-approaching space-time codes for systems employing four transmitter antennas," *IEEE Trans. Info. Theory*, vol. 49, pp. 726–733, Mar. 2003.

[20] H. Bouzekri and S. L. Miller, "Analytical tools for space-time codes over quasi-static fading channels," *in Proc. ICC*, vol. 3, pp. 1377–1381, May 2002.

[21] M. K. Byun, D. Park, and B. G. Lee, "Performance and distance spectrum of space-time codes in fast rayleigh fading channels," *in Proc. WCNC*, vol. 1, pp. 257–261, Mar. 2003.

[22] G. Tarricco and E. Biglieri, "Exact pairwise error probability of space-time codes," *IEEE Trans. Info. Theory*, vol. 48, pp. 510–513, Feb. 2002.

[23] M. Uysal and C. N. Georghiades, "Error performance analysis of space-time codes over rayleigh fading channels," *in Proc. VTC*, vol. 5, pp. 2285–2290, Sept. 2000.

[24] M. K. Byun and B. G. Lee, "New bounds of pairwise error probability for space-time codes in rayleigh fading channels," *in Proc. WCNC*, vol. 1, pp. 89–93, Mar. 2002.

[25] M. Uysal and C. N. Georghiades, "On the error performance analysis of space-time trellis codes," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1118–1123, July 2004.

[26] H. Shin and J. H. Lee, "Upper bound on the error probability for space-time codes in fast fading channels," *in Proc. VTC*, vol. 1, pp. 243–246, Sept. 2002.

[27] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Info. Theory*, vol. 43, pp. 1986–1992, Nov. 1997.

[28] 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 J. Select. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[29] V. K. N. Lau, Y. Liu, and T. Chen, "The role of transmit diversity on wireless communications-reverse link analysis with partial feedback," *IEEE Trans. Commun.*, vol. 50, pp. 2082–2090, Dec. 2002.

[30] T. Lo, "Adaptive space-time transmission with side information," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1496–1501, Sept. 2004.

[31] P. Xia, S. Zhou, and G. B. Giannakis, "Multiantenna adaptive modulation with beamforming based on bandwidth-constrained feedback," *IEEE Trans. Commun.*, vol. 53, pp. 526–536, Mar. 2005.

[32] A. F. Molisch and M. Z. Win, "Mimo systems with antenna selection," *IEEE Microwave Mag.*, pp. 46–56, Mar. 2004.

[33] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1437–1450, Oct. 1998.

[34] A. Santoso, Y. Li, and B. Vucetic, "Weighted space-time trellis codes," *IEE Electronics Lett.*, vol. 40, pp. 54–55, Feb. 2004.

[35] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Info. Theory*, vol. 47, pp. 2632–2639, Sept. 2001.

[36] G. Jongren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Trans. Info. Theory*, vol. 48, pp. 611–627, Mar. 2002.

[37] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Info. Theory*, vol. 46, pp. 1496–1501, Nov. 2000.

[38] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 1496–1501, Dec. 2000.

[39] A. Papoulis, *Probability, random variables, and stochastic process*, 4th ed. McGraw-Hill, Inc., 2002.

[40] J. K. Cavers, "An analysis of pilot symbol assisted modulation for rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 686–693, Nov. 1991.

[41] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," *IEEE Trans. Commun.*, vol. 47, pp. 199–207, Feb. 1999.

[42] ——, "Errata to 'space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths'," *IEEE Trans. Commun.*, vol. 51, p. 2141, Dec. 2003.

[43] P. Garg, R. K. Mallik, and H. M. Gupta, "Performance analysis of space-time coding with imperfect channel estimation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 257–265, Jan. 2005.

[44] Y. Ma and S. Pasupathy, "Performance of generalized selection combining on generalized fading channels," *in Proc. ICC*, vol. 5, pp. 3041–3045, May 2003.

[45] L. B. Milstein, "Guidelines for evaluation of radio transmission techniques for imt-2000," 1997.

[46] JTC(AIR), "Guidelines for system deployment modeling and simulation," 1994.08.01-065R4.

[47] N. Kong, "Average signal-to-interference-plus-noise ratio of a generalized optimum selection combiner for non-identical independent rayleigh fading channels in the presence of co-channel interference," *in Proc. ICC*, vol. 4, pp. 990 – 994, June 2001.

[48] M. Tao and P. Y. Kam, "Optimal differential detection and performance analysis of orthogonal space-time block codes over semi-identical mimo fading channels," *in Proc. VTC*, vol. 4, pp. 2403–2407, Sept. 2005.

[49] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

[50] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 920–928, June 2003.

[51] Z. Ding and B. Ward, "Subspace approach to blind and semi-blind channel estimation for spacectime block codes," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 357–362, 2005.

[52] P. Y. Kam, "Optimal detection of digital data over the nonselective rayleigh fading channel with diversity reception," *IEEE Trans. Commun.*, vol. 39, pp. 214–219, Feb. 1991.

[53] ——, "Adaptive diversity reception over a slow nonselective fading channel," *IEEE Trans. Commun.*, vol. COM-35, pp. 572–574, May 1987.

[54] F. Davarian, "Mobile digital communications via tone-calibration," *IEEE Trans. Veh. Technol.*, vol. 36, p. 55C62, May 1987.

[55] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1979.

[56] D. Samardzija and N. Mandayam, "Pilot assisted estimation of mimo fading channel response and achievalbe data rates," *IEEE Trans. Signal Process.*, vol. 51, pp. 2882–2890, Nov. 2003.

[57] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for mimo systems," *in Proc. ICC*, pp. 2326–2329, May 2003.

[58] J. Gao and H. Liu, "Decision-direction estimation of mimo time-varying rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1412–1417, July 2005.

[59] S. Zhang, P. Y. Kam, and P. Ho, "Performance of pilot-symbol-assisted-modulation with transmit-receive diversity in nonselective rayleigh fading channel," *in Proc. VTC*, vol. 2, pp. 1840–1844, Sept. 2004.

[60] R. Gozali and B. D. Woerner, "Upper bounds on the bit-error probability of space-time trellis codes using generating function techniques," *in Proc. VTC*, vol. 2, pp. 1318–1323, May 2001.

[61] M. P. Fitz, J. Grimm, and S. Siwamogsatham, "A new view of performance analysis techniques in correlated rayleigh fading," *in Proc. WCNC*, vol. 1, pp. 139–144, Sept. 1999.

[62] E. Biglieri and G. Tarrico, "Performance of space-time codes for a large number of antennas," *IEEE Trans. Info. Theory*, vol. 48, pp. 1794–1830, July 2002.

[63] J. K. Cavers and P. Ho, "Analysis of the error performance of trellis coded modulations in rayleigh fading channels," *IEEE Trans. Commun.*, vol. 40, pp. 74–83, Jan. 1992.

[64] G. Taricco and E. Biglieri, "Decoding space-time codes with imperfect channel estimation," *in Proc. ICC*, vol. 5, pp. 2741–2745, July 2004.

[65] J. G. Proakis, *Digital Communications*, 3rd ed. McGraw-Hill, Inc., 2001.

[66] B. Vucetic and J. Yuan, *Space-time Coding*. Hoboken, NJ : Wiley, 2003.

[67] G. Jongren and M. Skoglund, "Design of channel-estimate-dependent space-time block codes," *IEEE Trans. Commun.*, vol. 52, pp. 1191–1203, July 2004.

[68] Y. Li and P. Y. Kam, "Performance analysis of space-time trellis codes over rapid rayleigh fading channels with channel estimation," *in Proc. VTC*, vol. 1, pp. 497–501, Sept. 2005.

[69] ——, "Space-time trellis codes over rapid rayleigh fading channels with channel estimation −part i: Receiver design and performance analysis," *to appear in IEEE Trans. Commun.*

[70] X. T. Lin and R. S. Blum, "Systematic design of space-time codes employing multiple trellis coded modulation," *IEEE Trans. Commun.*, vol. 50, pp. 608–615, Apr. 2002.

[71] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, Inc., 1980.

[72] Y. Li and P. Y. Kam, "Space-time trellis code design over rapid rayleigh fading channels with channel estimation," *in Proc. WCNC*, vol. Physical Layer Track, PHY41, 2006.

[73] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, pp. 289–293, 1958.

[74] G. Caire and S. S.Shamai, "On the capacity of some channels with channel state information," *IEEE Trans. Inform. Theory*, vol. 45, pp. 2007–2019, Sept. 1999.

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

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

[77] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, pp. 139–157, Jan. 1999.

[78] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Uniform power allocation in mimo channes: a game-theoretic approach," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1707– 1727, July 2003.

[79] M. Médard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Info. Theory*, vol. 46, pp. 933–946, May 2000.

[80] E. Baccarelli and M. Biagi, "On the information throughput and optimized power allocation for mimo wireless systems with imperfect channel estimation," *IEEE Trans. Signal Process.*, vol. 53, pp. 2335–2347, July 2005.

[81] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading mimo channels with channel estimation error," *IEEE Trans. Inform. Theory*, vol. 52, pp. 2203–2214, May 2006.

[82] B. Hasibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, pp. 951– 963, Apr. 2003.

[83] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ : Prentice Hall, 1999.

[84] E. Baccarelli and M. Biagi, "Power-allocation policy and optimized design of multiple-antenna systems with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 136–145, Jan. 2004.

[85] J. Luo, R. S. Blum, L. Cimini, L. Greenstein, and A. Haimovich, "Power allocation in a transmit diversity system with mean channel gain information," *IEEE Commun. Lett.*, vol. 9, pp. 616–618, July 2005.

[86] W. Yu and J. M. Cioffi, "Constant-power waterfilling: performance bound and low-complexity implementation," *IEEE Trans. Commun.*, vol. 54, pp. 23–28, Jan. 2006.

[87] A. M. Mathai and S. B. Provost, *Quadratic forms in random variables: theory and applications*. Narcel Dekker, INC, 1992.

[88] Z. Chen, J. Yuan, and B. Vucetic, "Analysis of transmit antenna selection/maximal-ratio combining in rayleigh fading channels," *IEEE Trans. Veh. Technol.*, pp. 1312–1321, July 2005.

[89] S. Shakkottai, S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Commun. Mag.*, pp. 74–80, Oct. 2003.

[90] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity–part i: System description," *IEEE Trans. Commun.*, pp. 1927–1938, Nov. 2003.

[91] J. D. Parson, *The mobile radio propagation channel*. New York: Wiley, 1992.

[92] M. Uysal, "Pairwise error probability of space-time codes in rician-nakagami channels," *IEEE Commun. Lett.*, vol. 8, pp. 132–134, Mar. 2004.

[93] A. Maaref and S. Aissa, "Performance analysis of orthogonal space-time block codes in spatially correlated mimo nakagami fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 807–817, 2006.

[94] G. Femenias, "Ber performance of linear stbc from orthogonal designs over mimo correlated nakagami-m fading channels," *IEEE Trans. Veh. Technol.*, vol. 53, pp. 307–317, Mar. 2004.

[95] X. Zhang and N. C. Beaulieu, "Ser of threshold hybrid selection/maximal-ratio combining in correlated nakagami fading," *IEEE Trans. Commun.*, vol. 53, pp. 1423–1426, Sept. 2005.

[96] D. A. Gore and A. J. Paulraj, "Mimo antenna subset selection with space-time coding," *IEEE Trans. Signal Process.*, pp. 2580–2588, Oct. 2002.

[97] S. Thoen, L. V. d. Perre, B. Gyselinckx, and M. Engels, "Performance analysis of combined transmit-sc/receive-mrc," *IEEE Trans. Commun.*, pp. 5–8, Jan. 2001.

[98] J. Tang, X. Zhang, and Q. Du, "Alamouti scheme with joint antenna selection and power allocation over rayleigh fading channels in wireless networks," *in Proc. GLOBECOM*, pp. 3319–3323, Nov. 2005.

[99] C. Shan, "Performance analysis of space-time block coded systems with channel estimat," Ph.D. dissertation, NATIONAL UNIVERSITY OF SINGAPORE, 2005.

[100] J. C. Roh and B. D. Rao, "Design and analysis of mimo spatial multiplexing systems with quantized feedback," *IEEE Trans. Signal Process.*, vol. 54, pp. 2874–2886, Aug. 2006.

[101] H. Jiang, W. Zhuang, and X. Shen, "Cross-layer design for resource allocation in 3g wireless networks and beyond," *IEEE Commun. Mag.*, pp. 120–126, Dec. 2005.

[102] P. P. Pham, S. Perreau, and A. Jayasuriya, "New cross-layer design approach to ad hoc networks under rayleigh fading," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 28–39, Jan. 2005.

[103] M. A. Haleem and R. Chandramouli, "Adaptive downlink scheduling and rate selection: A cross-layer design," *IEEE J. Sel. Areas Commun.*, vol. 23, pp. 1287–1297, June 2005.

[104] V. K. N. Lau, M. Jiang, and Y. Liu;, "Cross layer design of uplink multi-antenna wireless systems with outdated csi," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 1250 – 1253, June 2006.

[105] Q. T. Zhang, X. W. Cui, and X. M. Li, "Very tight capacity bounds for mimo-correlated rayeleigh-fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 681–689, Mar. 2006.

[106] M. Dohler, "Virtual antenna arrays," Ph.D. dissertation, Kings College London, University of London, 2003.

List of Publications

1. **Yan Li** and Pooi Yuen Kam, "Space-time trellis codes over rapid Rayleigh fading channels with channel estimation −part I: receiver design and performance analysis," *to appear in IEEE Trans. Commun.*.

2. **Yan Li** and Pooi Yuen Kam, "Space-time trellis codes over rapid Rayleigh fading channels with channel estimation −part II: performance analysis and code design for non-identical distributions," *to appear in IEEE Trans. Commun.*.

3. **Yan Li** and Pooi Yuen Kam, "Power Allocation for STTC over Rapid Rayleigh Fading Channels with Channel Estimation," *to be submitted to IEEE Trans. Commun.*.

4. **Yan Li** and Pooi Yuen Kam, "Performance analysis of space-time trellis codes over rapid Rayleigh fading channels with channel estimation," *in Proc. VTC*, vol. 1, pp. 497-501, Sept. 2005.

5. **Yan Li** and Pooi Yuen Kam, "Space-time Trellis Code Design over Rapid Rayleigh Fading Channels with Channel Estimation," *in Proc. WCNC*, vol. 3, pp. 1655-1659, Apr. 2006.

6. **Yan Li** and Pooi Yuen Kam, "Space-time Trellis Codes Over Independent, Non-identically Distributed, Rapid, Rayleigh Fading Channels with Channel Estimation,' *in Proc. GLOBECOM*, sec. CTH12-2, pp. 1-5, Nov. 2006.

Appendix A

Derivation of The Covariance Matrix $\mathbf{\Gamma}$ in (5.11)

In this appendix, we derive the covariance matrix $\mathbf{\Gamma}$ in (5.11) of the receive signal $\mathbf{r}$ for the i.n.i.d. case.

We first define $\boldsymbol{\alpha} = \mathbf{E}\mathbf{v}$. Due to the block-diagonal structure of the estimation error matrix $\mathbf{E}$, we can easily show that $\boldsymbol{\alpha} = \left[\boldsymbol{\alpha}^T(1) \ \cdots \ \boldsymbol{\alpha}^T(K)\right]^T$, where $\boldsymbol{\alpha}(k) = \mathbf{E}(k)\mathbf{v}(k) = [\alpha_1(k) \ \cdots \ \alpha_{N_R}(k)]^T$ for each k. Each element $\alpha_i(k)$ of $\boldsymbol{\alpha}(k)$ is now given by

$$\alpha_i(k) = \sum_{j=1}^{N_T} e_{ij}(k)v_j(k) \, . \tag{A.1}$$

Due to the independence of the elements $\{e_{ij}(k)\}$ of $\mathbf{E}$, it can be shown that

$$\mathrm{E}\left[\alpha_i(k)\alpha_{i'}^H(l)\right] = \mathrm{E}\left[\left|\sum_{j=1}^{N_T} e_{ij}(k)v_j(k)\right|^2\right] \delta(k-l)\delta(i-i') \, . \tag{A.2}$$

With equal-energy MPSK modulation, we have $|v_j(k)|^2 = 1$, for all k and j. Then, the result in (A.2) can be reduced to

$$\mathrm{E}\left[\alpha_i(k)\alpha_{i'}^H(l)\right] = \left(2\sum_{j=1}^{N_T} \bar{\sigma}_{ij}^2\right) \delta(k-l)\delta(i-i') \, . \tag{A.3}$$

Thus, we have

$$\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right] = \mathrm{E}\left[\boldsymbol{\alpha}\boldsymbol{\alpha}^H\right] = \mathbf{I}_K \otimes \mathbf{\Xi} \tag{A.4}$$

where $\mathbf{\Xi}$ is given by,

$$\mathbf{\Xi} = diag.\left[2\sum_{j=1}^{N_T}\bar{\sigma}_{1j}^2 \ \cdots \ 2\sum_{j=1}^{N_T}\bar{\sigma}_{N_Rj}^2\right] . \tag{A.5}$$

By substituting the expression for $\mathrm{E}\left[\mathbf{E}\mathbf{v}\mathbf{v}^H\mathbf{E}^H\right]$ in (A.4) into the first term on the rightmost side of (5.10), we can straightforwardly obtain the covariance matrix $\mathbf{\Gamma}$ of $\mathbf{r}$ in (5.11).

Appendix B

The Statistics of X in (5.15)

To compute the conational PEP given in eq. (5.15) in Chapter 5, we first need derive the statistics of the random variable X. Here X is given by

$$X = \sum_{k=1}^{K} \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} \Big\{ |r_i(k) - \sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}_c(k)|^2 - |r_i(k) - \sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}_e(k)|^2 \Big\} . \tag{B.1}$$

Define

$$q_{ic}(k) = \sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}_c(k) \tag{B.2a}$$

$$q_{ie}(k) = \sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}_e(k) \tag{B.2b}$$

where $q_{ic}(k)$ is the conditional mean of $r_i(k)$, i.e., $q_{ic}(k) = \mathrm{E}\left[r_i(k)|\mathcal{I}, \mathbf{v}\right]$. Thus, X can be rewritten as

$$X = \sum_{i=1}^{N_R} (\overline{N}_{0_i})^{-1} \sum_{k \in \kappa} \left\{ |r_i(k) - q_{ic}(k)|^2 - |r_i(k) - q_{ie}(k)|^2 \right\} .$$

where

$$\begin{aligned} & |r_i(k) - q_{ic}(k)|^2 - |r_i(k) - q_{ie}(k)|^2 \\ = \; & \left(-r_i^*(k)q_{ic}(k) - r_i(k)q_{ic}^*(k) + |q_{ic}(k)|^2 + r_i^*(k)q_{ie}(k) + r_i(k)q_{ie}^*(k) - |q_{ie}(k)|^2 \right) \end{aligned}$$

Given the pilot measurements $\mathcal{I}$ and the transmitted signal sequence $\mathbf{v}$, X is conditionally Gaussian and the conditional mean of X can be calculated as

$$\begin{aligned}
&\mathrm{E}\left[X|\mathcal{I},\mathbf{v}=\mathbf{v}_c\right] \\
=\;& \sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}\sum_{k\in\kappa}\left(-2|q_{ic}(k)|^2+|q_{ic}(k)|^2+q_{ic}^*(k)q_{ie}(k)+q_{ic}(k)q_{ie}^*(k)-|q_{ie}(k)|^2\right) \\
=\;& \sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}\sum_{k\in\kappa}\left(-|q_{ic}(k)-q_{ie}(k)|^2\right) \\
=\;& -E_s\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}\sum_{k\in\kappa}|\hat{\mathbf{h}}_i^T(k)\mathbf{v}_{ce}(k)|^2\,. \qquad \text{(B.3)}
\end{aligned}$$

Letting $q_{ice}(k)=\sqrt{E_s}\hat{\mathbf{h}}_i^T(k)\mathbf{v}_{ce}(k)$, X may be rewritten again as

$$X=\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}\sum_{k\in\kappa}\left(-r_i^*(k)q_{ice}(k)-r_i(k)q_{ice}^*(k)+|q_{ic}(k)|^2-|q_{ie}(k)|^2\right). \qquad \text{(B.4)}$$

Thus, the conditional variance of X can be calculated as

$$\begin{aligned}
\mathrm{Var}\left[X|\mathcal{I},\mathbf{v}=\mathbf{v}_c\right] &= 4\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-2}\sum_{k\in\kappa}\mathrm{Var}\left[\Re(r_i^*(k)q_{ice}(k))|\mathcal{I},\mathbf{v}=\mathbf{v}_c\right] \\
&= 2\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-2}\sum_{k\in\kappa}\overline{N}_{0_i}|q_{ice}(k)|^2 \\
&= 2E_s\sum_{i=1}^{N_R}(\overline{N}_{0_i})^{-1}\sum_{k\in\kappa}|\hat{\mathbf{h}}_i^T(k)\mathbf{v}_{ce}(k)|^2\,. \qquad \text{(B.5)}
\end{aligned}$$

Appendix C

Derivation of The Characteristic Function in (5.20)

It has shown that the PEP in eq. (5.21) in Chapter 5 can be obtained by the method of moment generation function, based on the characteristic function of $\mathcal{D} = \sum_{k\in\kappa} \hat{\mathbf{h}}^H(k)\mathbf{A}(k)\hat{\mathbf{h}}(k)$. Here, we give the derivation of the characteristic function of $\mathcal{D}$.

Define $\hat{\mathbf{h}}(k) = \text{vec}(\hat{\mathbf{H}}^T(k)) = [\hat{\mathbf{h}}_1^T(k)\cdots\hat{\mathbf{h}}_{N_R}^T(k)]^T$, $\mathbf{B}^H(k) = (\overline{\mathbf{N}}_0)^{-\frac{1}{2}} \otimes \mathbf{v}_{ce}^*(k)$, and $\mathbf{A}(k) = \mathbf{B}^H(k)\mathbf{B}(k)$, where vec$(\cdot)$ is the vectorization operator. Then, (5.18) can be rewritten as $P(\mathbf{v}_c \to \mathbf{v}_e|\mathcal{I}, \mathbf{v} = \mathbf{v}_c) = Q\left(\sqrt{\frac{E_s}{2}\mathcal{D}}\right)$, where $\mathcal{D} = \sum_{k\in\kappa}\mathcal{D}_k$. The quantity $\mathcal{D}_k = \hat{\mathbf{h}}^H(k)\mathbf{A}(k)\hat{\mathbf{h}}(k)$ is a quadratic form in the $N_R N_T \times 1$ random vector $\hat{\mathbf{h}}(k)$, and $\hat{\mathbf{h}}(k) \sim \mathcal{CN}(\mathbf{0}, 2\mathbf{\Lambda})$, where $\mathcal{CN}(\mathbf{u}, \mathbf{\Sigma})$ denotes a complex, Gaussian random vector with mean $\mathbf{u}$ and covariance matrix $\mathbf{\Sigma}$. Here $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = diag.[\mathbf{\Lambda}_1 \cdots \mathbf{\Lambda}_{N_R}] \tag{C.1}$$

where

$$\mathbf{\Lambda}_i = \begin{bmatrix} \hat{\sigma}_{i1}^2 & 0 & \cdots & 0 \\ 0 & \hat{\sigma}_{i2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \hat{\sigma}_{iN_T}^2 \end{bmatrix}. \tag{C.2}$$

From [87], the characteristic function of $\mathcal{D}_k$ is

$$\psi_{\mathcal{D}_k}(\jmath\omega) \;=\; \mathrm{E}[e^{\jmath\omega\mathcal{D}_k}] = |\mathbf{I} - 2\jmath\omega\mathbf{\Lambda}^{\frac{1}{2}}\mathbf{A}(k)\mathbf{\Lambda}^{\frac{1}{2}}|^{-1} = \prod_{i=1}^{r}(1-2\jmath\omega\lambda_i)^{-1} \quad \text{(C.3)}$$

where $\lambda_1, \cdots \lambda_r$ are the eigenvalues of $\Phi = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{A}(k)\mathbf{\Lambda}^{\frac{1}{2}}$. Due to the block structure of both $\mathbf{A}(k)$ and $\mathbf{\Lambda}$, Φ can be expressed as $\Phi = diag.[\Phi_1 \cdots \Phi_{N_R}]$, where Φ_i is given by

$$\Phi_i = \frac{\mathbf{\Lambda}_i^{\frac{1}{2}}\mathbf{v}_{ce}^{*}\mathbf{v}_{ce}^{T}\mathbf{\Lambda}_i^{\frac{1}{2}}}{\overline{N}_{0_i}} \,. \quad \text{(C.4)}$$

Since the rank of Φ_i is one, the corresponding eigenvalue is

$$\lambda_i = \frac{\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)}{\overline{N}_{0_i}} \quad \text{(C.5)}$$

where $d_j(k) = |v_{cej}(k)|$, and $v_{cej}(k)$ is the jth element of $\mathbf{v}_{ce}(k)$. Thus, the characteristic function of $\mathcal{D}_k$ is

$$\psi_{\mathcal{D}_k}(\jmath\omega) = \prod_{i=1}^{N_R}\left[1 - 2\jmath\omega(\overline{N}_{0_i})^{-1}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right]^{-1} \,. \quad \text{(C.6)}$$

Due to the independence of the $\mathcal{D}_k$'s, the characteristic function of $\mathcal{D}$ can be expressed as

$$\psi_{\mathcal{D}}(\jmath\omega) = \prod_{k\in\kappa}\prod_{i=1}^{N_R}\left[1 - 2\jmath\omega(\overline{N}_{0_i})^{-1}\sum_{j=1}^{N_T}\hat{\sigma}_{ij}^2 d_j^2(k)\right]^{-1} \,. \quad \text{(C.7)}$$