Signal Processing Techniques for

Wireless Communication Systems

A DISSERTATION SUBMITTED TO

THE DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING

OF THE UNIVERSITY OF ADELAIDE

BY

Van Khanh Nguyen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF DOCTOR OF PHILOSOPHY

Commenced, February 2001

Dissertation Submitted, September 2003

Declaration of Originality

I declare that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university and that, to the best of my knowledge and belief, the thesis contains no material previously published or written by any other person except where due reference is made in the text.

I consent to the thesis being made available for photocopying and loan if accepted for the award of the degree.

Van Khanh Nguyen

Center for Internet Research (CIR)

Department of Electrical and Electronic Engineering

The University of Adelaide

September 2003

Acknowledgments

First and foremost, I would like to express my gratitude to my supervisor Professor Lang White for his guidance, support and encouragement throughout my PhD studies. Working with him has been a true privilege and a great experience for me. I would also like to thank Professor Lang White for his financial support for my travel to ETIS, ENSEA, University of Cergy-Pontoise, France, to conferences and to a training course on Space-time Coding.

I would like to thank Dr Jinho Choi, Dr Stephan Hanly, Professor Reginald Coutts, Mr Justin Hernandez and Mr Adriel Cheng for reading through my thesis and providing comments. I am also grateful to Professor Inbar Fijakow, Assoc. Professor Bruce Davis, Assoc. Professor Cheng Chew Lim, Mr David Bowler, Mr Stephen Guest and Mrs Rose-Marie Descalzi for their help during my research. My special thanks go to Assoc. Professor Bruce Davis and Dr Ken Sarkies for helping me to obtain the Premier's Scholarship in Information and Communication Technology and the Australian Postgraduate Award.

Finally, I would like to thank my family and friends for their continual support and encouragement. In particular, I'll single out Mum, Dad and Thanh Trang for always being there for me.

Abstract

The demands for wireless communication services are growing at a rapid rate. Meeting these demands is challenging since the availability of the radio spectrum at the frequencies of interest is limited. Furthermore, wireless communications also must cope with several other difficulties such as multiple access interference (MAI), channel fading, and limitations on the power and size of the mobile terminals. This thesis investigates the problems of MAI and channel fading in wireless communications, and focuses on developing spectrally efficient coding and signal processing techniques to mitigate the effects of these problems.

The first part of the thesis discusses the use of multiuser detection techniques to overcome the problems of MAI in code-division multiple access (CDMA) systems. The thesis develops two new interference cancellation detection techniques and an adaptive multiuser detector for joint parameter estimation and symbol detection. In addition, the thesis derives a novel framework for analysing the convergence behaviour of an interference cancellation technique which is commonly known as parallel interference cancellation.

In the second part of the thesis, the effects of channel fading on the performance of wireless communication systems are considered. The thesis examines the use of multiple transmit and multiple receive antennas in conjunction with coding for providing diversity to combat channel fading. Particular focus is given to the case when the propagation paths are spatially correlated. The performance of such communication systems is analysed and design criteria for constructing good codes are derived subsequently. The thesis then develops a receiver for joint decoding and channel estimation in time-varying fading channels.

Finally, since there are many different types of diversity which can be exploited in wireless communication systems, the thesis develops a generalised and unified taxonomy for system modelling and signal processing for such systems.

Publications

The following papers have been written based on the materials presented in this thesis.

Journal Papers

1. V. K. Nguyen and L. B. White, "Iterative multiuser detection with parameter estimation", *Digital Signal Processing: A Review Journal*, vol. 12, no. 2/3, pp. 145-158, July 2002. (doi:10.1006/dspr.2002.0436)

2. V. K. Nguyen, L. B. White, E. Jaffrot, M. Soamiadana and I. Fijalkow, "Recursive receivers for diversity channels with correlated fading", *IEEE Journal of Selected Areas in Communication*, Special issue on MIMO systems and applications, vol. 21, no. 5, pp. 754-764, June 2003.

3. V. K. Nguyen and L. B. White, "Joint space-time trellis decoding and channel estimation in correlated fading channels", *IEEE Signal Processing Letters*, Oct. 2003, accepted.

Conference Papers

1. V. K. Nguyen and L. B. White, "Interference cancellation schemes for CDMA systems", *IEEE Information, Decision and Control*, pp. 47-52, Feb. 2002, Adelaide, Australia.

2. V. K. Nguyen and L. B. White, "Recursive receiver structures for general diversity channels with time correlated flat fading", *Proc. 3rd Australian Communications Theory Workshop*, pp. 48-52, Feb. 2002, Canberra, Australia.

3. E. Jaffrot, V. K. Nguyen, M. Soamiadana, L. B White and I. Fijalkow, "Symbol by symbol reduced complexity highly selective OFDM channel estimation", *Proc. XI European Signal Processing Conference*, Sept. 2002, France.

4. V. K. Nguyen and L. B. White, "Space-time trellis decoder for spatially correlated time-varying fading channel", *Proc. 4th Australian Communication Theory Workshop*, Feb. 2003, Melbourne, Australia.

5. V. K. Nguyen and L. B. White, "Space-Time Coding in Spatially Correlated Rayleigh Fading Channels", Proc. 7th *International Symposium on Digital Signal Processing and Communication Systems*, Dec. 2003, Gold Coast, Australia.

Journal Paper in Preparation

1. V. K. Nguyen and L. B. White, "A general framework for analysing the convergence behaviour of parallel interference cancellation".

Contents

List of Figures

List of Tables

List of Acronyms

1G	First Generation
2G	Second Generation
3G	Third Generation
AMPS	Advanced Mobile Phone System
AOA	Angle of Arrival
APP	A Posteriori Probability
AR	Autoregressive
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BS	Base Station
CDMA	Code Division Multiple Access
CSI	Channel State Information
DFT	Discrete Fourier Transform
EM	Expectation-Maximization
FDMA	Frequency Division Multiple Access
FER	Frame Error Rate
GLRT	Generalised Likelihood Ratio Test
GSM	Global System for Mobile Communication
IMT-2000	International Mobile Telecommunications in the year 2000
IS-136	Interim Standard-136
IS-95	Interim Standard-95
JTACS	Japanese Total Access Communication System
KF	Kalman Filter
LMS	Least Mean Square
MAI	Multiple Access Interference

MAP	Maximum a Posteriori
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLSE	Maximum Likelihood Sequence Estimation
MMSE	Minimum Mean Square Error
MRC	Maximum Ratio Combining
MSC	Mobile Switching Center
MU	Mobile Unit
NCML	Noncoherent Maximum Likelihood
NMT	Nordic Mobile Telephone
NTT-800	Nippon Telephone and Telegraph-800
OFDM	Orthogonal Frequency Division Multiplex
PIC	Parallel Interference Cancellation
PDC	Personal Digital Cellular
PSP	Per-Survivor Processing
QPSK	Quadrature Phase Shift Keying
SER	Symbol Error Rate
SIC	Successive Interference Cancellation
SISO	Single Input Single Output
SNR	Signal to Noise Ratio
STD	Switched Transmit Diversity
TACS	Total Access Communication System
TDMA	Time Division Multiple Access
TXAA	Transmit Adaptive Array
UMTS	Universal Mobile Telecommunication System

Notation and Symbols

Below are the common mathematic notation and symbols used in this dissertations.

$\mathbf{A}$	Matrix notation
$\mathbf{A}^T$	Transpose of a matrix $\mathbf{A}$
$\mathbf{A}^H$	Conjugate transpose of a matrix $\mathbf{A}$
$*$	Complex conjugate
$|\mathbf{A}|$ or $\det(\mathbf{A})$	Determinant of a matrix $\mathbf{A}$
$\lambda_{max}\{\mathbf{A}\}$	Largest eigenvalue of a matrix $\mathbf{A}$
$\rho(\mathbf{A})$	Spectral radius of a matrix $\mathbf{A}$
$\mathbf{I}_N$	$N \times N$ identity matrix
$\mathbf{0}_N$	Matrix of all zeros
$\mathbf{a}$	Vector notation
$\sum$	Summation
$\prod$	Product
$\otimes$	Kronecker product
$\delta(.)$	Dirac delta function
$\arctan(.)$	Arc tangent function
$\cos(.)$	Cosine function
$\sin(.)$	Sine function
$\mathcal{N}(x, w)$	Gaussian density
$\mathcal{CN}(x, w)$	Circular-symmetric, complex Gaussian distribution
$I_0(.)$	Zeroth-order modified Bessel function
$J_0(.)$	Zeroth-order Bessel function of the first kind
$\mathbb{R}$	Real numbers
$\mathbb{C}$	Complex numbers

$\mathbb{Z}$	Integers
$\|\mathbf{p}\|_p$	ℓ_p-norm of the sequence $\mathbf{p}$
$\ell_{p,e}^K$	Extended normed function space

Chapter 1

Introduction and Summary of Contributions

1.1 Wireless Communications

Since the deployment of first generation (1G) cellular networks in the early 1980s, there has been a substantial increase in the development of wireless communication technologies. This tremendous boost in the cellular industry reflects the growing demands for higher data rate and better quality services, and the increase in the number of subscribers to mobile phone services. According to latest statistics [19], there are over 1.3 billion subscribers worldwide in 2003 as compared to only 10 million subscribers in 1990 [13]. In Europe, the current average market penetration of mobile phones is 70% of the total population [19], while in other countries like Iceland and Finland, the market penetration is as high as 90%. Thus, wireless communication does not only complement the mature wireline network but may become a dominant method of communication in the near future.

Since the early 1980s, wireless communications have gone through two generations of technology overhaul. The first generation of public cellular networks was

established to provide basic voice telephony services to mobile subscribers over a wide area. These first generation systems were analog and based on frequency-division multiplexing technologies. Examples of first generation systems are the Advance Mobile Phone System (AMPS) in North America, the Nordic Mobile Telephone/Total Access Communication System (NMT/TACS) in Europe, and the Nippon Telephone and Telegraph-800/Japanese Total Access Communication System (NTT-800/JTACS) in Japan. In the early 1990s, second generation (2G) systems based on digital transmission techniques were introduced to provide more robust communications. They provided basic services such as voice, facsimile, low-rate circuit and packet data (9.6 and 14.4 kb/s), and medium-rate packet data (up to 76.8 kb/s). Examples of 2G wireless systems are the Global System for Mobile Communication (GSM), Personal Digital Cellular (PDC), IS-136 and cdmaOne/IS-95. Due to the growing demands for a variety of multimedia communication services such as high-speed Internet access and video/high-quality image transmission, third generation (3G) wireless systems (e.g. Universal Mobile Telecommunication System/International Mobile Telecommunications-2000 (UMTS/IMT-2000)) are now under development to address these needs. These 3G wireless systems will evolve gracefully from mature 2G networks and offer true packet access at significant higher speeds. It is expected that 3G wireless communication systems will support user data rate at 144kb/s for vehicular applications, 384kb/s for outdoor pedestrian applications, and up to 2 Mb/s for indoor applications. For more details on the developments and standards of wireless communication systems, readers are referred to the following references [3, 13, 26, 58, 77, 111].

1.2   Motivation and Background

Unlike wireline communications, transmissions of information signals in wireless medium suffer several impairments that can significantly degrade their performance.

Among these impairments, multiple access interference and channel fading are the two major problems that limit the high-data rate transmission in wireless communication systems. In addition, wireless transmissions must also deal with the difficulty of limited availability of the radio spectrum at the frequency of interest where the propagation conditions are favourable. Thus, in order to effectively utilise this precious bandwidth and to overcome the impairments in wireless communications, there is a need to develop efficient methods of transmission and coding together with sophisticated signal processing techniques. The research in this thesis is therefore aimed achieving the above objectives. In this thesis, we address the problems of multiple access interference and channel fading in wireless communications and focus on two particular techniques that recently received enormous attention, namely multiuser detection and multiple transmit multiple receive antennas.

1.2.1 Multiple Access Interference - Multiuser Detection

There are a number of multiple access schemes that allow many wireless users to share simultaneously a finite amount of radio spectrum. Frequency division multiple access (FDMA), time division multiple access (TDMA) and code-division multiple access (CDMA) are the three major multiple access techniques for multiplexing wireless users. In FDMA systems, the frequency spectrum is partitioned into distinct bands of frequencies (or channels) and each user is allocated a dedicated frequency band in which information may be transmitted. In contrast to FDMA systems, TDMA systems divide the radio spectrum into time slots. Each user is assigned a time slot and during the time slot that user can access the entire available bandwidth. Since TDMA transmissions are time slotted, strict synchronism between the transmitter and receivers are required and guard slots are necessary to separate different users. Unlike FDMA and TDMA systems, each user in CDMA systems can transmit information over the entire available radio spectrum at all times. Each

user is distinguished from each other by multiplying its data with a unique signature code sequence. At the receiver, the message signal of a user can be detected by performing a time correlation operation of the received signal with that user's code waveform. Multiplexing wireless users using the CDMA technique has many advantages over the FDMA and TDMA techniques. These advantages include the following (but not limited to):

- *Potential Capacity Increase* - In the FDMA and TDMA systems, in order to avoid co-channel interference spectral guards and time guards respectively, are required. In addition, since each user in the FDMA system is assigned a frequency band, FDMA channels that are not in use cannot be used by other users to increase the capacity. Similarly for TDMA system where each user is allocated a time slot, when a time slot that is not in use, it cannot be used by other users. Thus, both FDMA and TDMA systems do not efficiently utilise the available spectrum. In contrast, CDMA systems allow users to access the entire available spectrum for all times. Therefore they fully utilise the available spectrum and have higher potential capacity over the other multiple access methods.

- *Soft Capacity Limit* - As the number of users in a CDMA system is increased, the level of multiple access interference also raises and causes the system performance to decrease accordingly. When the number of users is decreased, the system performance is then improved. Thus, there is no absolute limit on the number of users in a CDMA system.

- *Access Flexibility* - Since each user can use the entire bandwidth for transmission at any time, a new user can be added to the system without the channel resources having to be re-sliced.

- *Diversity* - Since transmitted signals are spread over a large spectrum, frequency diversity will arise if the spread spectrum bandwidth is greater than the channel coherent bandwidth. Hence, the effects of small-scale fading can be substantially reduced. In addition, the chip duration is very short and much less than the channel delay spread. Thus, multiple delayed versions of the chip signal will appear in the received signal. Therefore, a Rake receiver can be used to exploit this multipath diversity to improve reception.

- *Soft Handoff* - In CDMA systems, the spread spectrum mobiles share the same channel in every cell. Thus, two or more base stations can simultaneously monitor a particular user and allow the Mobile Switching Center (MSC) to choose the best version of the signal at any moment in time.

Due to the attractive features of the CDMA technique, it has been chosen as the main multiple access scheme for the 3G systems [3, 26, 111]. However, a major problem in CDMA systems is the presence of the multiple access interference (MAI). The MAI arises due to the non-orthogonality of the spreading code sequences employed by users and the fact that all users in a CDMA system share the same frequency spectrum for transmission at any given time. If the spreading code sequences are orthogonal to each other, the MAI can be completely suppressed after performing the time correlation operation of the received signal with the user's signature code waveform. However in practice, the orthogonality of the spreading code sequences cannot be maintained after they are transmitted through the wireless medium. Hence the MAI is always present. This MAI can severely limit the multiple access capability of CDMA systems if it is not properly exploited at the receiver.

A simple detection technique for CDMA systems is to correlate the received signal with the user's spreading code and pass it through a threshold device [33]. This is a single user detection strategy where each user is treated separately and the signal of other users are considered as noise. As a result, it performs poorly

when a large amount of MAI is present. A better approach is to perform multiuser detection where the information of multiple users is used jointly [100]. In [99], a maximum likelihood sequence estimator (MLSE) is proposed which involves finding an output sequence that has the maximum conditional probability. This is the optimal multiuser detector. However its complexity grows exponentially with the number of users and therefore can not be implemented in practical systems. Due to the high complexity of the optimal multiuser detector, a number of suboptimal multiuser detectors with a much lesser complexity have been proposed. The first class is known as linear multiuser detectors where a linear transformation is applied to the correlator outputs to reduce the MAI. The linear decorrelator [62] can completely remove the MAI by multiplying the correlator outputs with the inverse of the crosscorrelation matrix of the spreading codes, albeit at the cost of background noise enhancement. An advantage of the decorrelator is that it does not require the knowledge of the transmitted signal amplitudes. As it turns out, in the absence of any prior knowledge of the transmitted signal amplitudes, the decorrelator is the optimal detector [100]. The minimum mean squared error (MMSE) detector [63] is another example of linear suboptimal multiuser detectors. Unlike the decorrelator, the MMSE detector takes the background noise into account in the detection. It is found that when the background noise approaches zero, the MMSE detector converges to the decorrelator and when the background noise approaches infinity, it approaches the conventional matched filter [100, pp. 296]. Thus, the decorrelator and the conventional matched filter are the two limiting cases of the MMSE detector. A drawback of these two linear multiuser detectors is that they need to compute the inverse of a square matrix whose elements depended on the crosscorrelation of the spreading codes. This poses a major problem in terms of processing complexity when long codes are employed because the computation must be done in real time. In addition, due to the linear structure of the linear multiuser detectors, they have

limited capacity. An alternative approach to mitigate the MAI is to perform interference cancellation where each user estimates the interference contributed by other users and subtracts it from the received signal to give an estimate of the desired signal. Thus, a correct decision on a particular user's data symbol will cancel that user's interference, while an incorrect decision will enhance the contribution of that interferer. The interference cancellation process can be performed in a sequentially (serial) order [75, 102] or in parallel [29, 30, 31, 76, 97, 98, 109]. As compared to linear multiuser detection, interference cancellation techniques have lower complexity and can provide better performance. For this reason, some interference cancellation techniques have been proposed in the W-CDMA [3] and the CDMA I [111] proposals. Due to these attractive characteristics of the interference cancellation methods, the first part of this thesis will focus on this class of suboptimal multiuser detection.

1.2.2 Multipath Fading - Multiple Transmit and Receive Antennas

Due to the unguided nature of the wireless communication channels, radio frequency waves emitted by the transmitter propagate through different paths and undergo different reflection, refraction, diffraction and attenuation. They arrive at the receiver from different directions with different propagation delays and strengths. These waves are then combined, constructively or destructively depending on their phases, at the receiver to give a resulting signal which varies widely in amplitude and phase thereby inducing fading. Depending on the relative magnitude of the time required for the main portion of the transmitted signal to reach the receiver, termed delay spread, and the symbol period, multipath fading can be classified as flat fading or frequency selective fading. If the delay spread is less than the symbol period, the channels undergo flat fading and frequency selective fading otherwise. When frequency selective fading arises, it gives rise to intersymbol interference because the

transmitted signal arrives at the receiver over several symbol periods.

Another factor in wireless propagation channels that influences fading is the presence of Doppler shift which is induced by the motion of the receiver, transmitter or surrounding objects. This Doppler shift may cause the signal level recorded at the receiver to vary widely with time. The temporal variation in the received signal is termed fast fading if the channel impulse response changes rapidly within the symbol duration. Conversely, if the channels impulse response changes at a rate much slower than the transmitted symbol rate, it is called slow fading. It should be noted that a slow or fast fading channel can be either flat or frequency-selective depending on the time delay spread. Thus, a fading channel can be classified into one of the following four types: *flat slow* fading, *flat fast* fading, *frequency-selective slow* fading and *frequency-selective fast* fading.
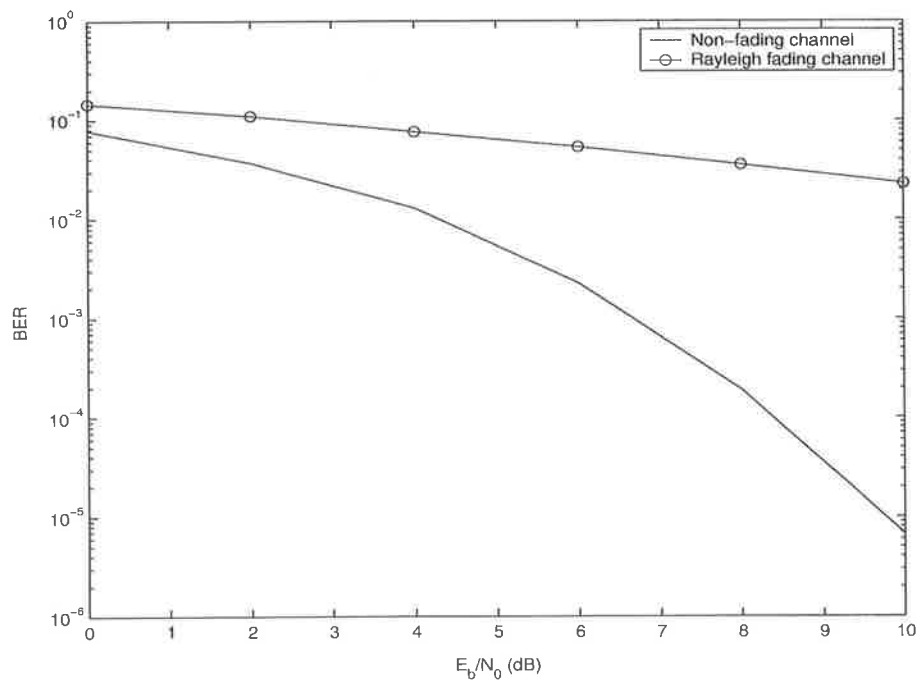
Figure 1.1: Performance degradation as a result of Rayleigh fading

Maintaining a reliable communication in wireless channels can become very dif-

ficult as a result of the random fluctuation in the received signal amplitude. In a situation where the receiver experiences deep fades, it may require up to 20-30dB more transmitted power in order to achieve the same bit error rate as systems operating over non-fading channels. Figure 1.1 illustrates the performance degradation of BPSK transmissions over Rayleigh fading channels. An effective method of combating channel fading is to introduce diversity into the system which would allow the transmission and/or reception of the information signal over multiple fading branches. If the multiple fading branches are highly uncorrelated, it would be unlikely that they experience deep fades simultaneously. Hence, there would be some branches that have acceptable signal quality which allow the receiver to recover the transmitted information correctly. One way of providing diversity is to use channel coding in conjunction with time interleaving. This provides temporal diversity for the system as it introduces redundancy of the transmitted signal in the temporal domain. An alternative diversity technique is to send information over more than one carrier frequency to induce different multipaths. Thus, if the carrier frequencies are separated by more than the coherent bandwidth of the channel, the transmitted signals will not experience the same fades and hence we have frequency diversity. Both temporal and frequency diversity techniques induce loss of bandwidth efficiency as they have to introduce redundancy in the time and frequency domain respectively. The third type of diversity is spatial diversity which can be obtained by employing multiple transmit and/or multiple receive antennas. This technique is more advantageous than the previous two as it can introduce redundancy of the signal without having to sacrifice the precious bandwidth resources.

Recently, receive antenna diversity has been exploited in the uplink of mobile systems by deploying multiple antennas at the base station. Provided that the antennas at the base station are well separated (i.e. around 10 wavelengths), the multiple received signals at the antennas will be reasonably uncorrelated and they

can be intelligently combined at the receiver to improve the performance of the system in the presence of channel fading. Common processing techniques that can be used to utilise this form of diversity are: switch diversity, equal gain and maximum ratio combining (MRC) [69, 80]. Switch diversity is the simplest diversity technique in which the signal from the antenna branch with the best quality is selected. In the equal gain combining, received signals from all antennas are co-phased and summed together. In the MRC, before the signals are co-phased and added together, they must be weighted to provide the optimal signal to noise ratio (SNR).

Another form of spatial diversity is transmit diversity which is obtained by deploying multiple antennas at the transmitter. A number of transmit diversity techniques for the downlink of mobile systems have been recently adopted or under consideration for the third generation standards [27]. Systems employing transmit diversity can be classified into one of the following three categories:

- Feedback schemes,

- Feedforward schemes,

- Blind schemes.

In the feedback schemes, explicit information of the fading channels is fedback from the receiver to the transmitter and the transmitter uses this knowledge to its advantage. Examples of the transmit diversity schemes that involve feedback are switched transmit diversity (STD) and transmit adaptive array (TXAA) [27]. In the STD scheme, information is transmitted on only one antenna at any given times. To determine which antenna to transmit, pilot symbols are sent from each antenna. Based on the average received power, the receiver then decides which antenna it would like to transmit. In the TXAA scheme, information signals are multiplied with a set of weights before being transmitted on all antennas simultaneously. These transmit weights are optimised to maximise the signal power at the receiver. They are

periodically being sent back from receiver to the transmitter via a feedback channel.

On the contrary, transmit diversity systems employing feedforward only require the transmitter to send training information so that the receiver can estimate the fading channels. There is no information being fedback to the transmitter. A simple diversity scheme of this type is the delay diversity [84] where a flat fading channel is made to become frequency selective by sending simultaneously the data symbol on one antenna and the delayed versions of the previous data symbols on the remaining antennas. Another scheme with feedforward information is space-time coding. In this method, coding is performed not only in the temporal domain but also across the spatial-domain created by the multiple antennas. Unlike the conventional time-domain channel coding in which the coding gain is achieved at the expense of bandwidth expansion, by taking the advantage of the spatial-domain, space-time coding achieves the coding gain without having to sacrifice the precious bandwidth. Space-time coding can be implemented in either block [7, 91] or trellis [89] forms.

The third category of transmit diversity does not require any information about the fading channels and hence no training sequence or feedback information is required. The capacity of communication systems belonged to this category has been analysed in [67, 112]. It is found that for a fixed number of antennas, as the length of the coherent interval of the fading channels increases the capacity approaches the capacity obtained as if the receiver has perfect channel estimates. It is also found that the space-time codes that attain capacity have a unitary structure (i.e. the signals are mutually orthogonal with respect to time among the transmit antennas) [4, 43, 44]. Another type of transmit diversity belonged to the blind scheme is differential space-time coding [45, 48, 92].

Transmit diversity schemes discussed so far can also be applied when multiple receive antennas are used. Thus, the resulting systems are multiple-input-multiple-

output (MIMO) systems which provide both transmit and receive diversity. As being shown in [36, 93], simultaneous deployment of multiple antennas at the transmitter and receiver not only provides diversity over fading channels but can significantly boost the channel capacity. Provided that the fading paths between all pairs of transmit and receive antennas are independent, it is found that the channel capacity increases linearly with the smaller of the number of transmit and receive antennas [35]. However, most previous works in this area assume that i) the channels are statistically independent and ii) the channels undergo quasi-static fading where the fading coefficients remain constant during the transmissions of a frame and change independently from one frame to another. The assumption of independent fading is the ideal case where we have a rich scattering environment and that the antennas within the transmitter or receiver sides can be sufficiently spaced apart. This could hardly be met in practice and spatial correlation will present. The assumption of quasi-static fading is only valid if the mobile unit is stationary or moving at low velocity. If the mobile unit is moving at high velocity, the channels will undergo fast fading where the fading coefficients can change from symbol to symbol. These fading coefficients, however, will be temporally correlated to a certain extend. In this thesis, we will examine MIMO systems with both spatial and temporal correlation factors being taken into account.

1.3 Overview of the Thesis and Contributions

The focus of this thesis is to develop efficient coding and signal processing techniques for the transmitter and the receiver to overcome the two major impairments in wireless communication systems, namely multiple access interference and channel fading. In chapter 2, we examine different multiuser detection techniques for mitigating the MAI in CDMA systems. We particularly focus on the interference cancellation techniques as they have low processing complexity and potential ca-

pacity gains. In chapter 3, we analyse the convergence behaviour of an interference cancellation technique known as parallel interference cancellation (PIC). In chapter 4, we investigate the performance of multiple transmit multiple receive antennas systems and design space-time trellis codes that can exploit the additional spatial diversity created by the multiple antennas. In chapter 5, we propose a receiver structure which can jointly decode the space-time trellis codes and estimate the fading channels. Since there are many different types of diversity which can be exploited, we develop a generalised and unified taxonomy for system modelling and signal processing for such systems in chapter 6. Conclusions of the thesis and future research direction are given in chapter 7.

The thesis is organised such that the materials presented in each chapter are self contained. We now give more detailed summary of the main contributions of the thesis.

Multiuser Detectors for CDMA Systems (Chapter 2) - In this chapter, we present two new interference cancellation techniques which are hybrid of the successive and parallel interference cancellation methods. Computer simulation results show that the performance of the proposed techniques are in general superior to the known successive and parallel interference cancellation techniques. Since interference cancellation techniques that use nonlinear tentative decisions require the knowledge of the signal amplitudes of all active users in the system, we propose an adaptive algorithm for performing joint parameter estimation and symbol detection. This proposed adaptive multiuser detector operates on-line where the estimates of the unknown parameters are updated for each incoming observation. Computer simulations are used to compare the performance of this proposed adaptive multiuser detector with that of the nonadaptive version where the signal amplitudes are perfectly known.

Convergence Behaviour Analysis of the PIC Technique (Chapter 3) - Since there are many different types of tentative decision functions that can be employed by the parallel interference cancellation (PIC) technique for estimating the MAI, a general framework for analysing the convergence behaviour of the PIC detector is developed. This framework permits the derivation of the sufficient condition for convergence of the PIC detector for a wide range of tentative decision functions. Several well-known conditions for convergence of the PIC detector with linear decisions and clip decisions can be obtained using this general framework. Computer simulations are also used to investigate the convergence behaviour of the PIC detector with hyperbolic tangent decisions and simulation results are compared with analytical results.

Space-Time Coding (Chapter 4) - The performance of space-time coded systems with multiple transmit multiple receive antennas are analysed. We derive two new upper bounds for the pairwise error probability of space-time coded systems in spatially correlated Rayleigh fading environments. The traditional design criteria (eg. rank determinant and trace) for constructing space-time trellis codes are to minimise the pairwise error probability of the dominant error event. In this chapter, we present a number of new space-time trellis codes based on the design criterion of minimising the sum of the pairwise error probability of all distinct pairs of codewords. Simulation results support the claim that these new codes are superior to other known codes constructed using the traditional rank determinant and the trace criteria.

Joint Space-time Decoding and Channel Estimation (Chapter 5) - A receiver for joint space-time trellis decoding and channel estimation in time-varying fading channels that are spatially and temporally correlated is proposed. By approximating

the physical channel model of the multiple transmit multiple receive antennas system with a statistical channel model, we incorporate per-survivor processing with Kalman filtering into the Viterbi algorithm to allow the receiver to suboptimally decode the space-time trellis codes and simultaneously track the channel variations. Simulation results demonstrate that a performance close to the maximum likelihood receiver with perfect channel state information can be obtained.

Recursive Receivers for General Diversity Channels (Chapter 6) - We introduce a general state-space model for a general diversity communication system with time and diversity correlated flat fading. Examples of diversity systems which fall within this framework include space-time coded systems, orthogonal frequency division multiplex (OFDM) systems, code division multiple access systems and hybrids of these systems. We develop a number of time-recursive receiver structures based on sequence estimation or on symbol by symbol estimation to exploit the temporal correlations in the channel. Such time-recursive receivers offer some advantages over block processing schemes such as computational and memory requirement reductions and the easier incorporation of adaptivity in the receiver structures. The receivers considered include: Per-survivor processing, M-algorithm, two *a posteriori* probability techniques and a per-symbol iterative technique based on EM algorithm. Using an OFDM system as example, the performance of these proposed receivers are compared with conventional designs which do not exploit the channel time correlations. Simulation results suggest that there can be significant gains in performance by incorporating time correlation into the signal model and the resulting receiver designs.

Chapter 2

Multiuser Detectors for CDMA Systems

This chapter considers the problem of multiuser detection in synchronous code-division multiple access (CDMA) systems. It focuses mainly on interference cancellation detectors as they can provide good performance with a relatively low computational complexity. It proposes two interference cancellation techniques which are hybrid of the successive interference cancellation (SIC) detector and the parallel interference cancellation (PIC) detector. Computer simulations are used to examine the performance of these detectors.

Since all interference cancellation detectors with nonlinear tentative decision function require the knowledge of the signal amplitudes of all active users in the systems, this chapter addresses the issue of joint parameter estimation and symbol detection of multiple users in the CDMA systems. It presents an adaptive multiuser detector which iteratively performs joint symbol detection and estimation of the unknown parameters using the Expectation-Maximization (EM) approach. Simulation results show that the performance of this adaptive multiuser detector is very close to that of the nonadaptive version where the signal amplitudes of all users are perfectly

16

known at the receiver.

2.1   Introduction

In this chapter, we review two popular interference cancellation detectors: the Successive Interference Cancellation (SIC) detector and the Parallel Interference Cancellation (PIC) detector. In the former detector, interference is successively removed from the received signal in the descending order of user's strength while the latter removes the interference for all users simultaneously. By performing the interference cancellation successively, the SIC detector has an extremely good performance when the powers of the users in the system are unequal. However, when all users in the system have equal power it performs poorly, especially for the users being processed first. The PIC detector on the other hand performs extremely well in both cases, but it requires several interference cancellation stages. This motivates us to propose two hybrid PIC-SIC detectors, which are combination of the SIC and PIC detectors. These combined detectors inherit the behaviours of the SIC detector when no power control is used and that of the PIC detector when ideal power control is used.

We particularly focus on the use of nonlinear tentative decision functions for estimating the MAI as they offer superior performance than the linear counterparts [29]. The main disadvantage with using nonlinear tentative decision functions is that the detector must have the knowledge of the signal amplitudes of all active users in the system. Most work on interference cancellation detectors with nonlinear tentative decision functions often assumes that the signal amplitudes are available at the receiver. In this chapter, we present an adaptive multiuser detector that performs joint parameter estimation and symbol detection for code division multiple access (CDMA) systems. In the literature, there are proposed receivers which also perform parameter estimation and symbol detection. The proposed receivers in [10, 53] perform parameter estimation and symbol detection using the Baum-Welch

version of the Expectation-Maximization (EM) method. The unknown parameters are estimated using the maximum likelihood criterion. These methods operate off-line and update the parameter estimates based on an entire block of observations using the forward-backward algorithm. These methods have the advantage that they can estimate the unknown parameters without the need of a training sequence [65]. Hence they improve the throughput of the system. However, they have the drawback that they require large amounts of memory for storage of the forward and backward variables. This motivates the use of an on-line algorithm [59] in our proposed receiver which updates the parameter estimates for each incoming observation. This method not only reduces the memory requirement, it also allows the receiver to perform in real-time.

The layout of this chapter is organised as follows: In section 2.2, a synchronous CDMA system model is presented. In section 2.3, we review some known interference cancellation detectors and describe the two proposed detectors. The performance of these detectors with different nonlinear tentative decision functions are examined by using computer simulations. In section 2.4, we describe the adaptive multiuser detector which performs joint symbol detection and parameter estimation for CDMA systems. The performance of this detector is also examined via computer simulation.

2.2   The System Model

We consider a synchronous CDMA system with K users transmitting simultaneously to a common receiver over an additive white Gaussian noise (AWGN) channel. Each user $k \in \{1, ..., K\}$ is assigned a normalized signature sequence $\mathbf{s}_k$ of length M, $\mathbf{s}_k \in \{\frac{-1}{\sqrt{M}}, \frac{1}{\sqrt{M}}\}^M$ and transmits at symbol rate the information $d_k(i) \in \{\pm 1\}$. We use the time indices t and i to denote the chip index and symbol index, respectively.

The transmitted signal of the kth user during the tth chip interval is

$$y_k(t) = A_k(i) \, d_k(i) \, s_k(j) \, , \tag{2.1}$$

where $A_k(i) > 0$ is the signal amplitude of user k over symbol period i, $d_k(i)$ is the ith symbol and $t = Mi + j$, for $0 \le j \le M - 1$.

Thus, the baseband received signal at the chip rate is the noisy sum of all signals, which can be written as

$$r(t) = \sum_{k=1}^{K} y_k(t) + v(t) \, , \tag{2.2}$$

where $v(t)$ is the realization at time t of a zero mean white Gaussian noise with variance σ^2.

Now, let $\mathbf{r}(i)$ consists of M consecutive observations of the received data at the chip rate $t = Mi$ to $M(i+1) - 1$ i.e. $\mathbf{r}(i) = [r(Mi) \, r(Mi+1) \, \cdots \, r(M(i+1)-1)]^T$ where T denotes the transpose operation. Similarly, we use

$$\mathbf{v}(i) = [v(Mi) \, v(Mi+1) \, \cdots \, v(M(i+1)-1)]^T \, , \tag{2.3}$$

$$\mathbf{y}_k(i) = [y_k(Mi) \, y_k(Mi+1) \, \cdots \, y_k(M(i+1)-1)]^T \, , \tag{2.4}$$

for the corresponding noise vector and vector of user k contribution signals, respectively.

Therefore, the symbol rate version of (2.2) is:

$$\mathbf{r}(i) = \sum_{k=1}^{K} \mathbf{y}_k(i) + \mathbf{v}(i) \text{ with } \mathbf{y}_k(i) = A_k(i) \, d_k(i) \, \mathbf{s}_k \, . \tag{2.5}$$

2.3 Interference Cancellation Techniques

2.3.1 Tentative Decision Functions

Before we review the SIC and PIC detectors, and describe the proposed detectors, we firstly describe some tentative decision functions that can be used by the detectors

for estimating the MAI. In literature, there are several tentative decision functions have been used such as the hard decision [29, 32, 97, 98, 109], the infinitely soft decision (linear) [29, 75], the hyperbolic tangent decision [30, 31, 76], and the null-zone [29]. For interference cancellation detectors that use these tentative functions, with the exception of the linear decision, they all require the knowledge of the signal amplitudes of all users in the system in order to reconstruct the MAI. The estimation of the signal amplitude can be done using the techniques described in [53, 70] and section 2.4. Studies in [40] have shown that imperfect estimation of the signal amplitude may significantly reduce the performance of the detectors. For the interference cancellation detectors with linear tentative decision function, the MAI can be estimated using the output of the tentative decision function since the signal component at the output is linearly proportional to the user's signal amplitude. This in effect provides a joint estimate of the signal amplitude and the user's data bit, and hence no longer requires the estimation of the signal amplitude separately. However, the disadvantage of using the linear tentative decision function is that additive noise is now introduced into the cancellation process. Studies in [29] found that the PIC detector with linear decisions is inferior to one with the hard decisions, hyperbolic tangent decisions or null-zone decisions.

In this paper, we compare the performance of the SIC, PIC and the proposed detectors using the hard decision and the hyperbolic tangent decision functions. These two tentative decision functions can be described mathematically by the following equations

- Hard-decision:

$$g_k(x) = \text{sign}(x) \ . \tag{2.6}$$

- Hyperbolic tangent decision:

$$g_k(x) = \tanh(\alpha_k \, x) \ , \tag{2.7}$$

where α_k are positive constants.

2.3.2 Successive Interference Cancellation Detector

The idea of successive interference cancellation (SIC) detector is to perform interference cancellation in descending order of the user's strength. Here we assume the signal amplitudes are ordered as follow: $A_1(i) \geq A_2(i) \geq \cdots \geq A_K(i)$. There are two reasons for cancelling in this descending order. Firstly, users with stronger power can be detected more reliably and secondly, the removal of stronger signals have the most benefit for the remaining users as they contribute significant interference. The successive detection statistics using the nonlinear tentative decision functions are computed according to

$$
\begin{aligned}
\hat{\mathbf{y}}_k(i) &= \mathbf{r}(i) - \sum_{j=1}^{k-1} A_j(i)\, \tilde{d}_j(i)\, \mathbf{s}_j \ ; \\
\tilde{d}_k(i) &= g_k(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k(i)) \,,
\end{aligned}
\tag{2.8}
$$

where $\hat{\mathbf{y}}_k(i)$ is the kth user's estimate of the desired signal and $g_k(x)$ is the tentative decision function as previously described. From (2.8) one can notice that the first user sees all the interference while later users see less and less interference as the process progresses. Thus, there is no benefit for the first user and the detection is equivalent to using the conventional matched filter. The most beneficial user is the last user as it utilises the decisions of all other users and ideally (i.e., when the decisions of the stronger power users are correct) achieves the single user performance.

2.3.3 Parallel Interference Cancellation Detector

An alternative to successive cancellation is to perform parallel cancellation in which all users simultaneously subtract off all the interference from the received signal. Thus all users receive equal treatment. The parallel cancellation scheme can be

done in multiple stages as in [29, 97, 98, 109]. The idea here is that as the number of stages increases, the accuracy of the tentative decisions will be improved and hence the PIC detector can remove more and more interference. Now, suppose the PIC detector performs L interference cancellation stages, then at the nth stage where $n = 1, ..., L$ we have the following estimate for the desired signal

$$
\begin{aligned}
\hat{\mathbf{y}}_k^{(n)}(i) &= \mathbf{r}(i) - \sum_{j \neq k} A_j(i)\, \tilde{d}_j^{(n-1)}(i)\, \mathbf{s}_j \\
\tilde{d}_k^{(n)}(i) &= g_k \left(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(n)}(i) \right) ,
\end{aligned}
\tag{2.9}
$$

where $\tilde{d}_k^{(0)}(i) = g_k \left(\mathbf{s}_k^T \mathbf{r}(i) \right)$. Note that with this notation, the initial stage which involves correlating the spreading sequence with the received signal is not counted as an interference cancellation stage. For the final stage, a hard decision is used to determine the transmitted information bit.

2.3.4   Iterative Multiuser Detector

In this section, we describe the iterative multiuser detector presented in [76] where it iteratively separates the received signal into individual user signals and estimates the *a posteriori* probabilities (APPs) of the user's data symbol. The iterative process is per-symbol basis. This iterative multiuser detector can also be viewed as a PIC detector where each iteration corresponds to one interference cancellation stage. In fact, we will show that for the simple CDMA system considered in this chapter, where there is no interchip interference, this iterative multiuser detector is exactly the same as the PIC detector with hyperbolic tangent decision function. However the advantage of this iterative multiuser detector implementation is its ease in incorporating adaptive processing algorithms into the detector as will be shown later in section 2.4.

This iterative multiuser detector works as follow: Let $\hat{\pi}_{k,\phi}^{(n)}$ denotes the *a posteriori* at the nth iteration that user k transmits $d_k(i) = \phi$ for $\phi \in \{\pm 1\}$. For the

initial iteration, we initialise $\hat{\pi}_{k,1}^{(0)} = \hat{\pi}_{k,-1}^{(0)} = 0.5$ since $d_k(i) \in \{\pm 1\}$ are transmitted with equal probabilities. Assuming at the end of iteration n these $\hat{\pi}_{k,\phi}^{(n)}$ have been calculated, then at iteration $n + 1$, we then apply the following steps:

Signal Separation: Using the estimated *a posteriori* probabilities from previous iteration, the iterative multiuser detector separates the received signal into individual user signal component according to

$$\hat{\mathbf{y}}_k^{(n+1)}(i) = \mathbf{r}(i) - \sum_{j \neq k} \tilde{\mathbf{y}}_j^{(n)}(i) \ , \tag{2.10}$$

where $\tilde{\mathbf{y}}_j^{(n)}(i)$ is the conditional expectation of $\mathbf{y}(i)$ computed by

$$\tilde{\mathbf{y}}_j^{(n)}(i) = \sum_{\phi \in \{\pm 1\}} A_j(i)\, \mathbf{s}_j\, \pi_{j,\phi}^{(n)}(i)\, \phi$$
$$= A_j(i)\, \mathbf{s}_j\, (2\pi_{j,1}^{(n)}(i) - 1) \ . \tag{2.11}$$

In the equation above, we have used the fact that $\pi_{j,1}^{(n)}(i) + \pi_{j,-1}^{(n)}(i) = 1$, $\forall i$, n. Let the soft estimate of the user data be defined as $\tilde{d}_k^{(n)}(i) = 2\pi_{k,1}^{(n)}(i) - 1$, then (2.10) can be rewritten as

$$\hat{\mathbf{y}}_k^{(n+1)}(i) = \mathbf{r}(i) - \sum_{j \neq k} A_j(i)\, \mathbf{s}_j\, \tilde{d}_j^{(n)}(i) \ , \tag{2.12}$$

Calculation of a posteriori probabilities: By defining $\hat{z}_k^{(n+1)} = \mathbf{s}_k^T \hat{\mathbf{y}}_k^{(n+1)}$, we can write

$$\hat{z}_k^{(n+1)} = A_k(i)d_k(i) + \mathbf{s}_k^T \mathbf{v}(i) + \sum_{j \neq k} \rho_{j,k} A_j(i) \left(d_j(i) - \tilde{d}_j^{(n)}(i) \right) \ , \tag{2.13}$$

where $\rho_{j,k} = \mathbf{s}_k^T \mathbf{s}_j$ is the crosscorrelation between the spreading codes of user k and user j. The advantage of using $\hat{z}_k^{(n)}$ instead of $\hat{\mathbf{y}}_k^{(n)}$ in the calculation of the *a posteriori* probability is that the uncancelled MAI doesn't have to be assumed as white Gaussian (i.e. the variance of $\sum_{j \neq k} \mathbf{y}_j(i) - \tilde{\mathbf{y}}_j^{(n)}(i)$ does not have to be proportional to an identity matrix). This is important as this assumption can not be justified for

general CDMA systems. In the expression above, the first term is the desired signal, the second term is the background noise and the last term is the unsuppressed multiple access interference (MAI). The uncanceled interference term of (2.13) can be approximated as a zero-mean Gaussian random variable [31]. Hence we can model $\hat{z}_k^{(n+1)}(i)$ as a Gaussian random variable with mean $A_k(i)d_k(i)$ and covariance γ^2 where $\gamma^2 > \sigma^2$, recalling that σ^2 is true additive Gaussian noise variance. Assuming the user data $d_k(i) \in \{\pm 1\}$ are transmitted with equal probability, the *a posteriori* can then be updated according to

$$\hat{\pi}_{k,\phi}^{(n+1)}(i) = p\left(d_k(i) = \phi \,|\, \hat{z}_k^{(n+1)}(i)\right)$$
$$= \frac{p\left(\hat{z}_k^{(n+1)}(i) \,|\, d_k(i) = \phi\right)}{\sum_{\theta \in \{\pm 1\}} p\left(\hat{z}_k^{(n+1)}(i) \,|\, d_k(i) = \theta\right)}. \tag{2.14}$$

This iterative process is terminated when *a posteriori* probabilities do not change significantly from one iteration and the next.

To show that this iterative multiuser detector corresponds to the PIC detector with hyperbolic tangent decision, we first derive the expression for the *a posteriori* probability that $d_k(i) = 1$ is transmitted. Using (2.14),

$$\hat{\pi}_{k,1}^{(n+1)}(i) = p\left(d_k(i) = 1 \,|\, \hat{z}_k^{(n+1)}(i)\right)$$
$$= \frac{\mathcal{N}\left(\hat{z}_k^{(n+1)}(i) - A_k; \gamma^2\right)}{\mathcal{N}\left(\hat{z}_k^{(n+1)} - A_k; \gamma^2\right) + \mathcal{N}\left(\hat{z}_k^{(n+1)}(i) + A_k; \gamma^2\right)}$$
$$= \frac{1}{1 + \exp\left\{-\frac{2}{\gamma^2} A_k(i)\, \hat{z}_k^{(n+1)}(i)\right\}} \tag{2.15}$$

where $\mathcal{N}(x, w)$ denotes the Gaussian density

$$\mathcal{N}(x, w) = \frac{1}{\sqrt{2\pi w}} \exp\left\{-\frac{x^2}{2w}\right\}. \tag{2.16}$$

Now, substituting (2.15) into the definition of $\tilde{d}_k^{(n)}(i)$ gives

$$
\begin{aligned}
\tilde{d}_k^{(n)}(i) &= \frac{2}{1 + \exp\left\{-\frac{2}{\gamma^2} A_k(i)\, \hat{z}_k^{(n)}(i)\right\}} - 1 \\
&= \frac{1 - \exp\left\{-\frac{2}{\gamma^2} A_k(i)\, \hat{z}_k^{(n)}(i)\right\}}{1 + \exp\left\{-\frac{2}{\gamma^2} A_k(i)\, \hat{z}_k^{(n+1)}(i)\right\}} \\
&= \tanh\left(\frac{A_k(i)}{\gamma^2} \hat{z}_k^{(n)}(i)\right)
\end{aligned}
\tag{2.17}
$$

Hence, we can rewrite (2.12) as

$$
\hat{\mathbf{y}}_k^{(n+1)}(i) = \mathbf{r}(i) - \sum_{j \neq k} A_j(i)\mathbf{s}_j \tilde{d}_j^{(n)}(i)
\tag{2.18}
$$

$$
\tilde{d}_k^{(n)} = \tanh\left(\frac{A_k(i)}{\gamma^2} \mathbf{s}_k^T \hat{\mathbf{y}}_k^{(n)}(i)\right) .
\tag{2.19}
$$

This expression is the same as (2.9) of the PIC detector where $g_k(x) = \tanh(\alpha_k x)$ and $\alpha_k = A_k(i)/\gamma^2$.

2.3.5   Successive-Parallel IC Detector: Scheme 1

This proposed detector is a hybrid between the SIC and the PIC detectors. It has a multistage structure similar to the PIC detector and at each stage it also attempts to cancel the interference of all interferers. However, the interference cancellation process is performed successively in the descending order of user's strength similar to the SIC detector. In estimating the amount of interference presented in the received signal, it utilises the decisions of the stronger power users in the current stage as well as the tentative decisions of weaker users in previous stage. Hence it always uses the most up-to-date available estimate of the interference. Thus, unlike the SIC detector which totally ignores the contribution of weaker users and treats them as background noise, this proposed detector will take them into account. This hybrid detector is also different from the PIC detector in the way that at each stage

the PIC detector only uses the tentative decisions from the previous stage, while the proposed detector utilises the decisions of the current stage as well as from the previous stage and successively cancels the interference.

Again, we assume the signal amplitudes are ordered as follow: $A_1(i) \geq A_2(i) \geq \cdots \geq A_K(i)$. Thus, the estimate of the individual signal at the nth interference cancellation stage can be computed according to

$$
\begin{aligned}
\hat{\mathbf{y}}_k^{(n)}(i) &= \mathbf{r}(i) - \sum_{j=1}^{k-1} A_j(i)\, \tilde{d}_j^{(n)}\, \mathbf{s}_j - \sum_{j=k+1}^{K} A_j(i)\, \tilde{d}_j^{(n-1)}\, \mathbf{s}_j \ ; \\
\tilde{d}_k^{(n)}(i) &= g_k \left(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(n)}(i) \right) \ .
\end{aligned}
\tag{2.20}
$$

for $2 \leq n \leq L$. For the first interference cancellation stage, we will include a weighting factor to the tentative estimates of the matched filter outputs. The reason behind this inclusion is that there is no interference cancellation being performed on these tentative estimates. As a result, they may not be reliable, especially for the low power users. Since using an incorrect estimate of the MAI of an interferer in the cancellation process will enhance the contribution of that interferer, the role of the weighting factor is therefore to reduce this enhancement when the tentative estimate is incorrect. Thus, the weighting factor is a measure of the reliability of the tentative estimates of the matched filter, i.e. $\tilde{d}_k^{(0)}(i) = g_k \left(\mathbf{s}_k^T \mathbf{r}(i) \right)$. For a very reliable tentative decision $\tilde{d}_k^{(0)}(i)$, this weighting factor will be adjusted close to 1 while for a less reliable estimate, this weighting factor will be reduced accordingly. Consequently, we will fully use the correct tentative estimates to cancel the interference while for a less accurate tentative estimates, we only partially make use of them. We don't address the issue of finding the optimal weighting factor in this chapter. Instead, as an example we use the probability that the decision $\tilde{d}_k^{(0)}(i)$ is correct under the worse case condition as the weighting factor. That is

$$
W_k(i) = \Pr\!\left(A_k(i) - \sum_{j \neq k} A_j(i)\, |\rho_{j,k}| \geq v \right) ,
\tag{2.21}
$$

where v is the background noise. Thus, the 1st stage's estimate of the desired signal with the weighting factor is

$$
\begin{aligned}
\hat{\mathbf{y}}_k^{(1)}(i) &= \mathbf{r}(i) - \sum_{j=1}^{k-1} A_j(i)\, \tilde{d}_j^{(1)}\, \mathbf{s}_j - \sum_{j=k+1}^{K} W_j(i)\, A_j(i)\, \tilde{d}_j^{(0)}\, \mathbf{s}_j \; ; \\
\tilde{d}_k^{(0)}(i) &= g_k \left(\mathbf{s}_k^T\, \mathbf{r}(i) \right) \; .
\end{aligned}
\tag{2.22}
$$

After the publication of [71], we become aware that this hybrid PIC-SIC detector is similar to the detectors examined in [81, 108] where they are shown to be equivalent to the Gauss-Seidel algorithm. However, there is a slight different to our proposed detector as those detectors do not have the weighting factor at the first stage.

2.3.6 Successive-Parallel IC Detector: Scheme 2

In [98], Varanasi and Aazhang looked at the performance of a two-stages PIC detector (i.e. L=1) in which the first stage is the linear decorrelator. The performance of this detector is found to be extremely high as compared to the two stages PIC detector where the conventional matched filter is used as the first stage. However, it has a disadvantage in term of complexity. That is, it requires to compute the inverse of the crosscorrelation matrix of the spreading codes. This can be a major disadvantage in situations where the CDMA system employs long code in which each symbol is spread by a random code and is different from symbol to symbol. Thus, in this situation the detector must compute the inverse of the crosscorrelation matrix for every symbol time interval. In this section we propose an interference cancellation detector where the first stage performs SIC and subsequent stages perform PIC. There are two reasons that we use the SIC in the first stage. Firstly, SIC has a lower complexity than the linear decorrelator while provides a better estimate of the individual user's signal than the conventional matched filter, especially in the system where no power control is exercised. Secondly, using the SIC in the first

stage has a potential to help the detector to obtain a faster convergence when power control is not used.

Since the first stage uses SIC, it will perform interference cancellation in the descending order of signal power. We will again assume the signal amplitudes are ordered as follow: $A_1(i) \geq A_2(i) \geq \cdots \geq A_K(i)$ for the first stage. Thus, the individual user signal is estimated by the following equations :

At stage $n = 1$:

$$
\begin{aligned}
\hat{\mathbf{y}}_k^{(1)}(i) &= \mathbf{r}(i) - \sum_{j=1}^{k-1} A_j(i)\, \tilde{d}_j^{(1)}(i)\, \mathbf{s}_j \ ; \\
\tilde{d}_k^{(1)}(i) &= g_k(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(1)}(i)) \ ,
\end{aligned}
\tag{2.23}
$$

and at stage n for $2 \leq n \leq L$

$$
\begin{aligned}
\hat{\mathbf{y}}_k^{(n)}(i) &= \mathbf{r}(i) - \sum_{j \neq k} A_j(i)\, \tilde{d}_j^{(n-1)}(i)\, \mathbf{s}_j \ ; \\
\tilde{d}_k^{(n)}(i) &= g_k\left(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(n)}(i)\right) \ .
\end{aligned}
\tag{2.24}
$$

Note that we count the initial SIC stage as one interference stage since interference being cancelled within this stage.

2.3.7 Performance Evaluation of IC Techniques

We evaluate the Bit Error Rates (BER) performance of these interference cancellation detectors through extensive computer simulations since the exact analytical evaluation of the BER is very complex, especially when nonlinear tentative decision function is used. As shown in [97], the computation of the BER of a two-stages PIC detector is only possible when the number of users is small as it involves integrating multidimensional normal distribution. This difficulty arises due to the fact that the tentative decisions are not independent of one another. Similarly for the SIC detector, the computation of the BER is difficult because the tentative decisions of

the stronger users (i.e., users being processed before the user currently considered) are dependent.

We compare the performance of the proposed detectors with the conventional matched filter, the SIC detector, the PIC detector and the single user performance bound for three different scenarios. We consider a CDMA system with 6 active users and each user is assigned a pseudorandom code sequence of length $N = 7$. We examine the performance of these detectors using the hard decision and the hyperbolic tangent decision functions. The three scenarios that we examine are:

- Ideal power control: where all users have the same power with $A_k^2 = 1 \ \forall k$.

- No power control: where all users have different powers with the following power distribution: $A_1 = 30$, $A_2 = 20$, $A_3 = 10$, $A_4 = 5$, $A_5 = 2$ and $A_6 = 1$.

- Non-ideal power control: where the power distribution is $A_1 = 10$ and $A_k = 1$ for $k = 2, ..., 6$.

In the simulation results, we use the postfixes -HD and -HT to indicate whether the detectors use the hard decisions or hyperbolic tangent decisions, respectively.

Ideal Power Control

We obtained the BER for all six users at the signal to noise ratio $SNR = 9dB$ where the SNR for user k is defined as A_k^2/σ^2. Fig. 2.1 shows the BER of different detectors using the hard decision function while Fig. 2.2 shows the results when the hyperbolic tangent decision function is used. The BER of the PIC and the proposed detectors are obtained using three interference cancellation stages (i.e. L = 3). We find that there is no significant performance improvement after this third stage. As from the two graphs, we can see that the performance of the proposed detectors outperforms the conventional matched filter and is closed to the single user bound. Comparing the performance of the proposed detectors with the SIC

Figure 2.1: Ideal power control - Hard decision function
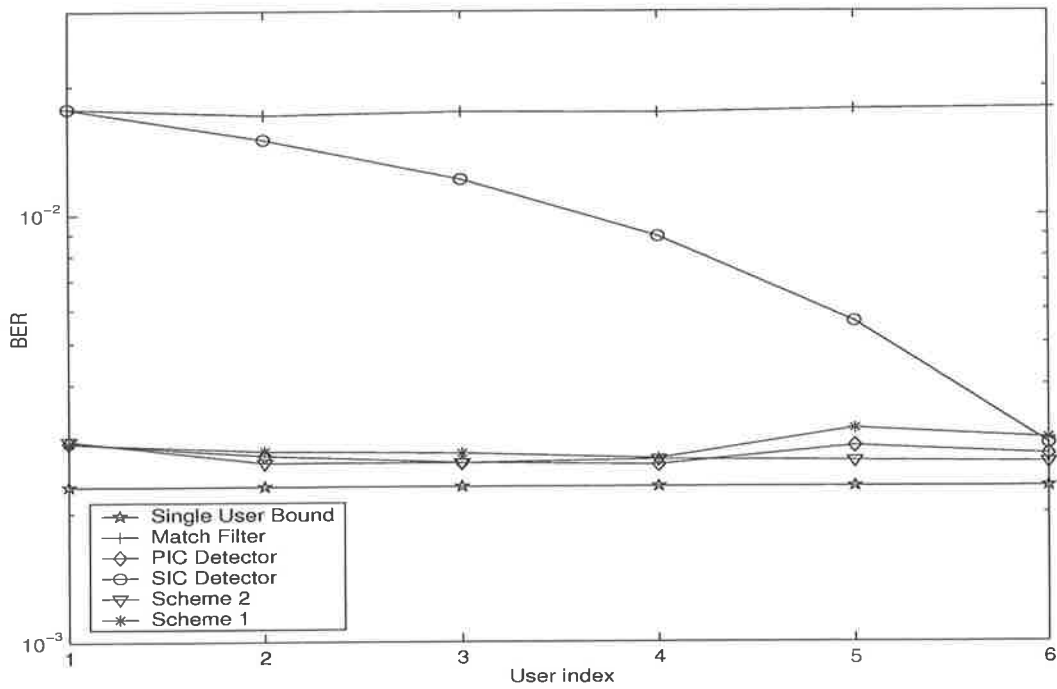
Figure 2.2: Ideal power control - Hyperbolic tangent decision function

detector, there is a considerable improvement for the users that are being processed first. The performance of the proposed detectors is approximately the same as the PIC detector, except for scheme 2 that uses hard decisions which has a slightly higher BER. These results demonstrate that in the ideal power control environment the proposed detectors behave similar to the PIC detector.
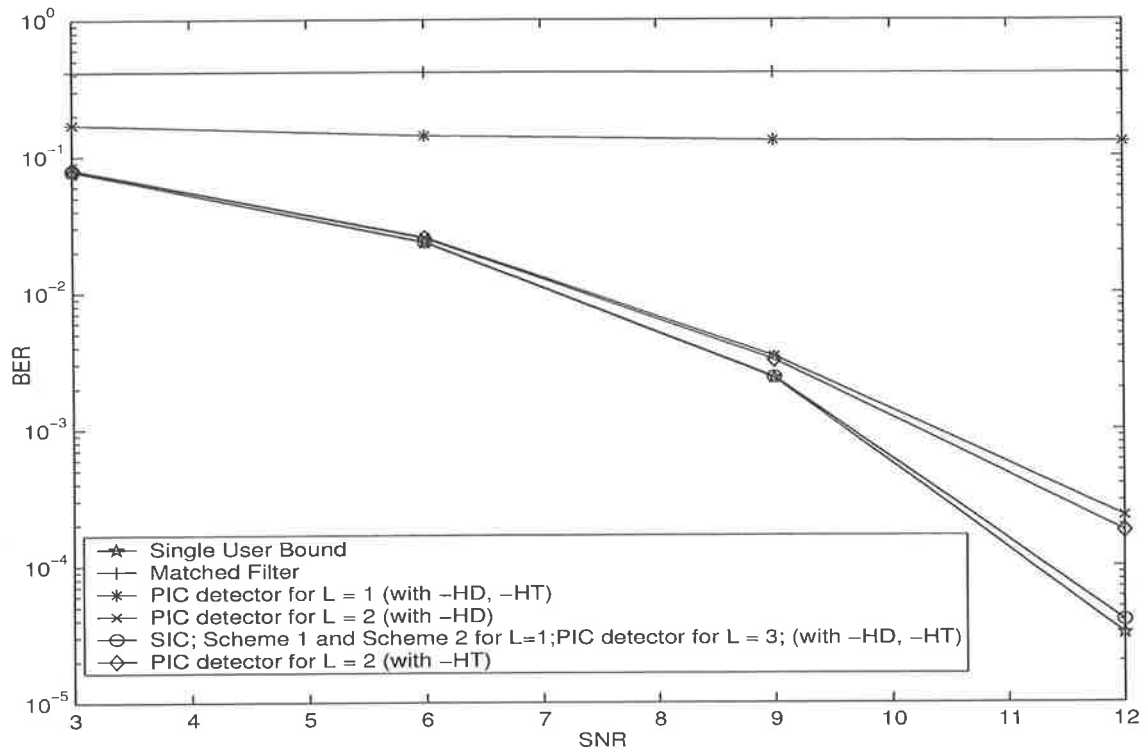
Figure 2.3: Performance of IC detectors with no power control

No Power Control

We examine the performance of these detectors for the no power control scenario using both hard decision and hyperbolic tangent decision functions. Here we obtain the BER of the lowest power user for different values of SNR. Fig. 2.3 shows that all the interference cancellation detectors that we examined achieve near single user performance bound. However, the PIC detector requires up to three interference

cancellation stages in order to obtain the same performance as that of the SIC and the proposed detectors where only one interference cancellation stage is required. Thus, in the system where no power control is exercised, the proposed detectors are now behaved similar to the SIC detector. The matched filter performs extremely poor in this case. The result shows that increasing the SNR does not improve the BER of matched filter. This is the inherent behavior of the matched filter due to the near-far effect.

Non-ideal Power Control

In this section, we investigate the performance of these detectors under the non-ideal power control condition. We examine the performance of the five lower power users at the SNR = 9dB (SNR of the lower power users). Fig. 2.4 shows the performance of the five lower power users (user 2 to user 6) when hard decision function is used. The BER of the PIC and proposed detectors are obtained using three interference cancellation stages (i.e. $L = 3$). As one can see the matched filter under this situation performs extremely poor. This is due to the presence of the strong interference from user 1. Simulation results show that the SIC detector significantly improves the BER of lower power users as compared to the matched filter. This is because the strong signal of user 1 can be reliably detected and removed from the received signal. However, the performance of those users being processed in the early stage is still quite far from the single user bound. The BER of the PIC detector is much better than conventional matched filter. However it is still higher than that of the proposed detectors. The reason that the PIC detector has a lower performance than the proposed detectors is due to the unreliable tentative decisions from the matched filter (as a result of the presence of the strong interference of user 1) in the initial stage. This unreliable initial estimate is the main cause for this performance lost since it is well-known that for any multistage interference cancellation (or iterative)
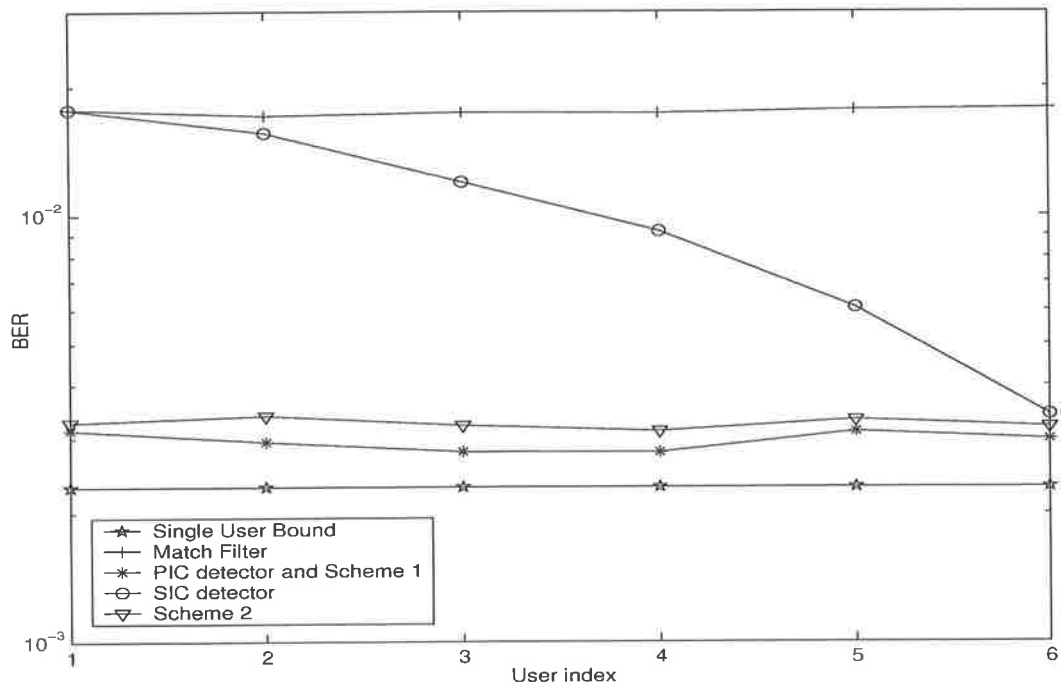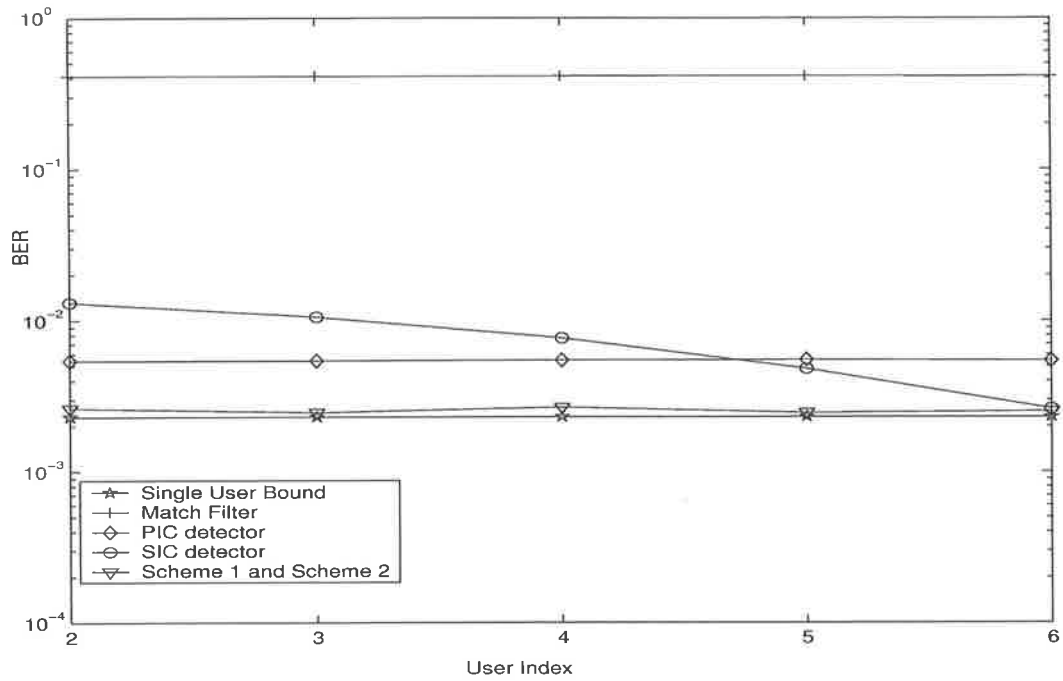
Figure 2.4: Non-ideal power control - Hard decision function
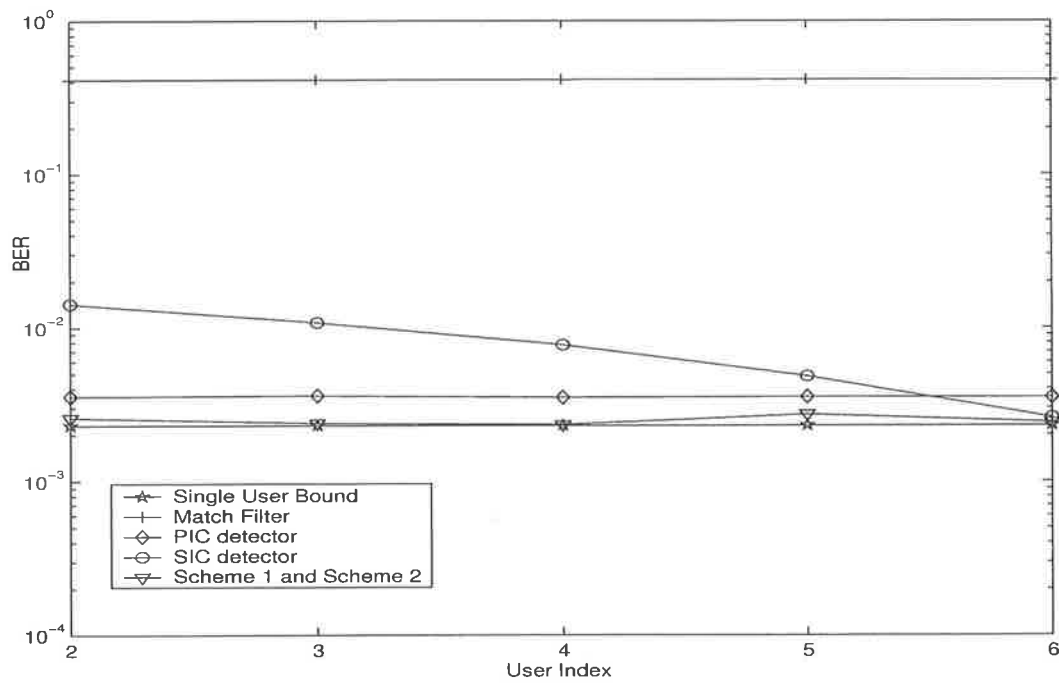
Figure 2.5: Non-ideal power control - Hyperbolic tangent decision function

process, the accuracy of the initial estimates is very important. The performance of these detectors when employing hyperbolic tangent decision function is illustrated in Fig. 2.5. It shows a similar performance order as in the case when hard decision function is used.

2.4 Iterative Multiuser Detection with Parameter Estimation

In this section we describe an iterative multiuser detector which performs joint parameter estimation and symbol detection. This detector is an extension of the iterative multiuser detector presented in section 2.3 where we now include an extra step into the detection algorithm to estimate the unknown parameters. It will perform on-line parameters estimation based on the recursive approach described in [59]. The idea of this method is to use the estimates of the individual signal contribution and the *a posteriori probabilities* (APPs) to perform parameter estimation. These estimated parameters will then be used for symbol detection which involves separating the received signal $\mathbf{r}(i)$ into individual user signal contribution $\mathbf{y}_k(i)$ and computing the APPs. This process is iterated several times per symbol until the APPs do not change significantly from one iteration to the next.

2.4.1 Detailed Algorithm Description

Step 1 - Initialisation: At the start of the iterative process of each symbol, we need to perform some initialisations. Let $\gamma_k^{2(n)}(i)$ and $A_k^{(n)}(i)$ denote the estimates of the noise variance (taking into account the unsuppressed interference) and the signal amplitude of user k at the nth iteration, respectively. Assuming that the signal amplitude does not change significantly from one symbol period to the next, we set $A_k^{(0)}(i) = A_k^{(n^*)}(i-1)$, where n^* is the terminating iteration index of symbol

$i - 1$. This is a reasonable assumption when the mobile user is moving at low speed. Although the noise variance $\gamma_k^{(n)}(i)$ varies from one symbol period to the next due to different combination of the transmitted bits of the interferers, we still set $\gamma_k^{2(0)}(i) = \gamma_k^{2(n^*)}(i-1)$ as $\gamma_k^{2(n^*)}(i-1)$ is the only available information that we have about the noise variance at the beginning of the symbol interval i. For the first symbol $i = 1$, we initialize $A_k^{(0)}(1) = \alpha$ and $\gamma_k^{2(0)}(1) = \beta$ where α and β are arbitrary positive values.

Since each user k transmits the information symbols $d_k(i) \in \{\pm 1\}$ with equal probability, we initialize the APPs $\hat{\pi}_{k,\phi}^{(0)}(i) = 0.5$, where $\hat{\pi}_{k,\phi}^{(n)}(i)$ denotes the estimate of the APP at the nth iteration that user k transmits $d_k(i) = \phi$ for $\phi \in \{\pm 1\}$. Thus, the initial conditional expectation of $\mathbf{y}_k(i)$ given the initial estimates for the signal amplitudes and APPs is

$$\tilde{\mathbf{y}}_k^{(0)}(i) = \sum_{\phi \in \{\pm 1\}} A_k^{(0)}(i)\, \mathbf{s}_k\, \hat{\pi}_{k,\phi}^{(0)}(i)\, \phi = 0 \, . \tag{2.25}$$

Step 2 - Signal Separation: Based on the available estimate of the conditional expectation $\tilde{\mathbf{y}}_k^{(n)}(i)$ at iteration n from step 5, we estimate the kth user signal contribution

$$\hat{\mathbf{y}}_k^{(n+1)}(i) = \mathbf{r}(i) - \sum_{j \neq k} \tilde{\mathbf{y}}_j^{(n)}(i) \, , \tag{2.26}$$

which can be rewritten as

$$\hat{\mathbf{y}}_k^{(n+1)}(i) = A_k(i)\, d_k(i)\, \mathbf{s}_k + \mathbf{v}(i) + \sum_{j \neq k} (\mathbf{y}_j(i) - \tilde{\mathbf{y}}_j^{(n)}(i)) \, . \tag{2.27}$$

Step 3 - Computation of the APPs: Let's define $\hat{z}_k^{(n+1)}(i) = \mathbf{s}_k^T \hat{\mathbf{y}}_k^{(n+1)}(i)$, then

$$\hat{z}_k^{(n+1)}(i) = A_k(i)\, d_k(i) + \mathbf{s}_k^T \mathbf{v}(i) + \sum_{j \neq k} \rho_{j,k} \left(A_j(i)\, d_j(i) - \sum_{\phi \in \{\pm 1\}} A_j^{(n)}\, \hat{\pi}_{j,\phi}^{(n)}(i)\, \phi \right) . \tag{2.28}$$

As discussed in the previous section, we can approximate $\hat{z}_k^{(n+1)}(i)$ as a Gaussian random variable of mean $A_k(i) d_k(i)$ and covariance $\gamma_k^{2(n)}(i)$, where $\gamma_k^{2(n)}(i) > \sigma^2$,

recalling that σ^2 is the variance of the background noise. Since the symbols $d_k(i) \in \{\pm 1\}$ are transmitted with equal probability, the *a posteriori* probability $\hat{\pi}_{k,\phi}^{(n+1)}(i)$ can be calculated as follows

$$
\begin{aligned}
\hat{\pi}_{k,\phi}^{(n+1)}(i) &= p\left(d_k(i) = \phi \mid \hat{z}_k^{(n+1)}(i)\right) \\
&= \frac{p\left(\hat{z}_k^{(n+1)}(i) \mid d_k(i) = \phi\right)}{\sum_{\theta \in \{\pm 1\}} p\left(\hat{z}_k^{(n+1)}(i) \mid d_k(i) = \theta\right)} .
\end{aligned}
\tag{2.29}
$$

Step 4 - Parameters Estimation: In this step, the signal amplitude and the noise variance are estimated using the EM algorithm. The E-step involves calculating the average log likelihoods with respect to the APPs. Thus we have the following cost function which is then maximized over the unknown parameters $\gamma_k^{2(n)}(i)$ and $A_k(i)$:

$$
\begin{aligned}
J_k^{(n+1)}(i) &= \sum_{\phi \in \{\pm 1\}} \hat{\pi}_{k,\phi}^{(n+1)}(i) \log \mathcal{N}\left(\hat{z}_k^{(n+1)}(i) - z_{k,\phi}(i); \gamma_k^{2(n)}(i)\right) \\
&= -\frac{1}{2} \log\left(2\pi\gamma_k^{2(n)}(i)\right) - \frac{1}{2\gamma_k^{2(n)}(i)} \sum_{\phi \in \{\pm 1\}} \hat{\pi}_{k,\phi}^{(n+1)}(i) \left(\hat{z}_k^{(n+1)}(i) - z_{k,\phi}(i)\right)^2 ,
\end{aligned}
\tag{2.30}
$$

where $z_{k,\phi}(i) = A_k(i)d_k(i)|d_k(i) = \phi$. For the M-step, we use the recursive algorithm introduced by Titterington [94] for parameter estimation using the incomplete data. The unknown parameter can be estimated according to the following formula

$$
\varphi_k^{(n+1)}(i) = \varphi_k^{(n)}(i) + \mu R_k^{(n)-1}(i)\frac{dJ_k^{(n+1)}(i)}{d\varphi_k(i)}
\tag{2.31}
$$

$$
R_k^{(n+1)}(i) = (1-\mu)R_k^{(n)}(i) + \mu\frac{dJ_k^{(n+1)}(i)}{d\varphi_k(i)}\frac{dJ_k^{(n+1)}(i)}{d\varphi_k(i)}^T ,
\tag{2.32}
$$

where $\varphi_k^{(n)}(i)$ is the estimate of the unknown parameter such as the signal amplitude and noise variance, and $1 - \mu$ is a forgetting constant, usually chosen to be close to, but slightly less than unity.

Step 5 - Calculation of the conditional expectation of $\mathbf{y}_k(i)$: If the APPs don't change significantly compared to the previous iteration, we terminate the iterative

process. Otherwise, we update the approximate conditional expectation of $\mathbf{y}_k(i)$ using the newly estimated signal amplitudes and APPs according to

$$\tilde{\mathbf{y}}_k^{(n+1)}(i) = \sum_{\phi \in \{\pm 1\}} A_k^{(n+1)}(i)\, \mathbf{s}_k\, \hat{\pi}_{k,\phi}^{(n+1)}(i)\, \phi\ . \tag{2.33}$$

2.4.2   Performance Evaluation

We again use simulations to obtain the BER of the iterative multiuser detectors since the exact analytical evaluation of the BER is very complex as shown in [97]. We examine the performance of the iterative multiuser detector with parameter estimation and compare it with that of the nonadaptive version where the signal amplitudes of all users are known at the receiver. In addition, we also compare it with the performance of the conventional matched filter and the single user bound. Simulation results are obtained for a CDMA system with 7 active users and each is assigned a pseudorandom sequence of length 7. In the simulation, the termination condition is chosen such that if the APPs of all users change by less than 0.005, the iterative process will be stopped.

Ideal Power Control

We investigate the performance of the iterative multiuser detectors when all users have equal power. We obtain the BER for different values of SNR where the SNR for user k is defined as $\frac{A_k^2}{\gamma^2}$. Fig. 2.6 shows that the performance of the iterative multiuser detector with parameter estimation degrades only slightly as compared to that of the nonadaptive version. It is evident that both iterative multiuser detectors outperform the conventional matched filter and achieve a performance which is close to the single user bound. Fig. 2.7 also shows the average number of iterations that the iterative multiuser detectors iterate per symbol. As one can see, on average the iterative multiuser detectors require less than 4.7 iterations/symbol at the SNR = 3 dB and 2.9 iterations/symbol at the SNR = 12 dB.
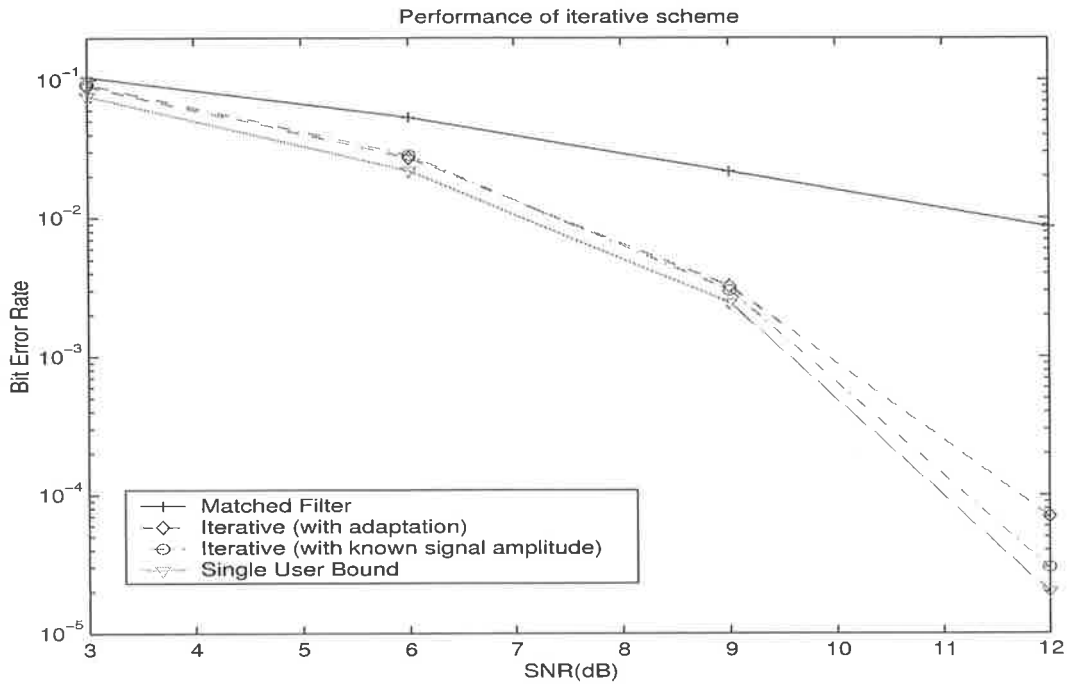
Figure 2.6: The performance of iterative multiuser detectors
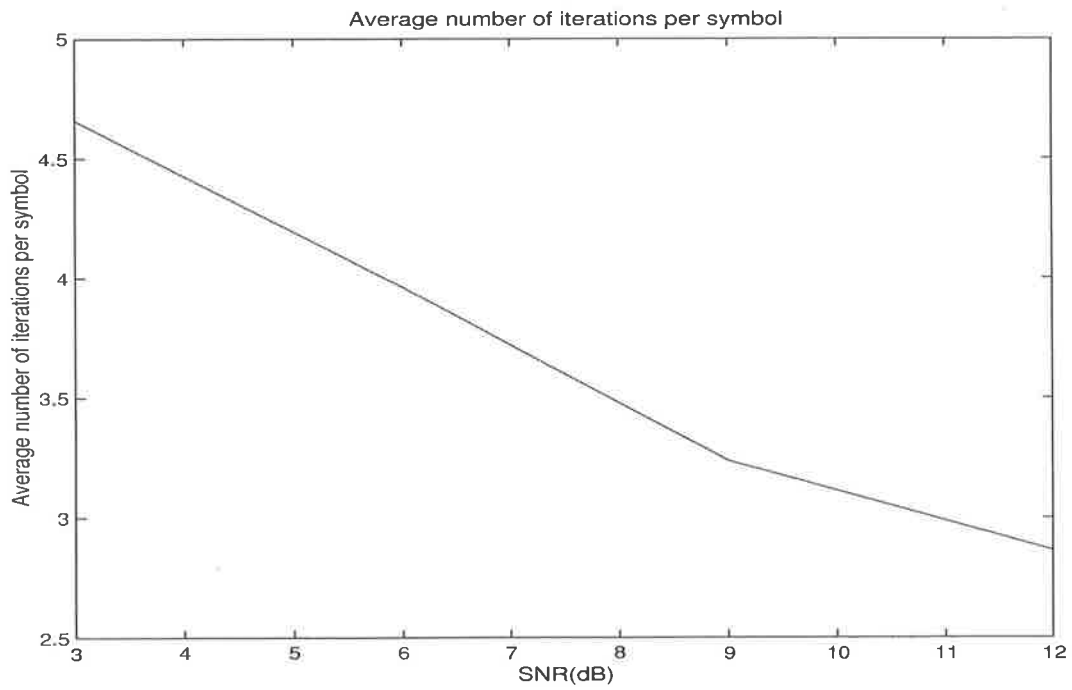
Figure 2.7: Average number of iterations per symbol

Fig. 2.8 and Fig. 2.9 are plots of the estimates of the signal amplitude and the noise variance (background noise plus unsuppressed MAI) respectively at the SNR = 12dB for the first 5000 symbols. The result shows that the estimated signal amplitude converges to the true amplitude $A_k(i) = 1$ and $\gamma_k^{2(\ell)}(i)$ converges to approximately 0.13 within about 500 symbols. A faster rate of convergence can be obtained by increasing the value of μ however it is found that this would lead to an estimated amplitude with larger variation.
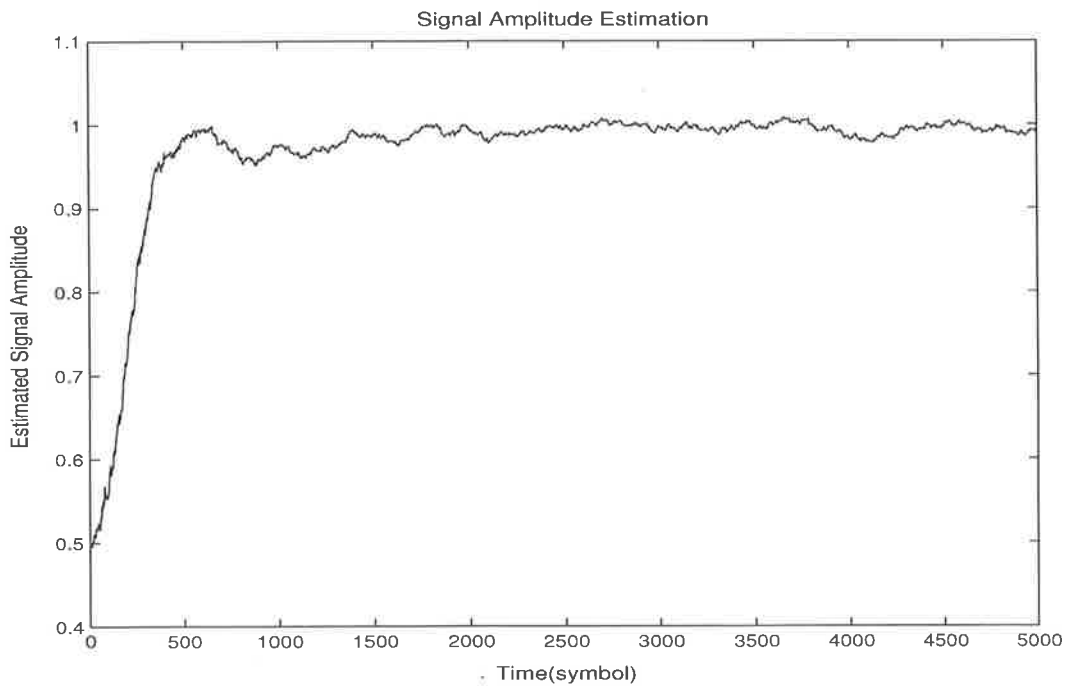
Figure 2.8: Estimated signal amplitude

No Power Control

The near-far resistance property of the iterative multiuser detectors is examined in this section. We obtain the BER of the lowest power user in the system for two different cases. In the first case the power of the weakest user is at -20dB relatives to the others. As shown in Fig. 2.10, the performance of the matched filter is extremely
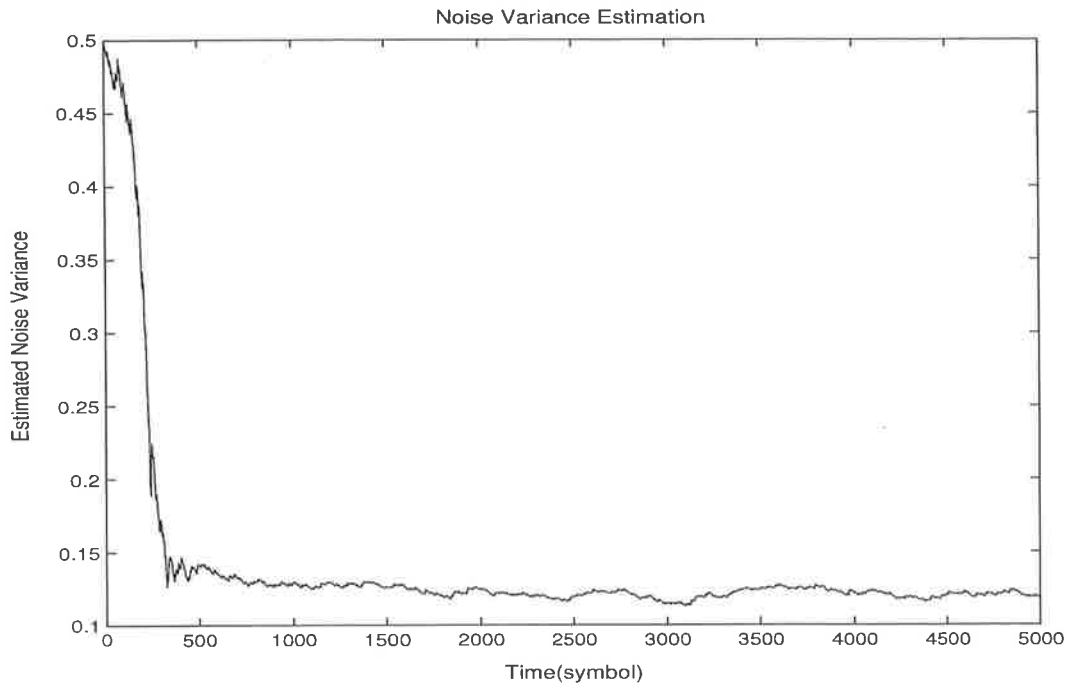
Figure 2.9: Estimated noise variance

poor in this case. The results show that increasing the SNR does not improve the BER of the matched filter. This is the inherent property of the matched filter due to the near-far effect. The BER of the iterative detector with known signal amplitudes is very close to the single user bound despite the present of strong interferers. This suggests that the iterative multiuser detector with known signal amplitude is near far resistant. The iterative multiuser detector with parameter estimation performs poorly in this case. This is because when the power of the interferers is large, inaccurate estimation of the signal amplitudes of the interferers results in a large amount of interference uncanceled. In the second case, the power of the weakest user is at -6dB relatives to the others. The results in Fig. 2.11 show that the BER of the matched filter is still unacceptably high while the iterative detector with known signal amplitude again performs extremely well as expected. The BER of iterative multiuser detector with parameter estimation, however, is very close to the non-
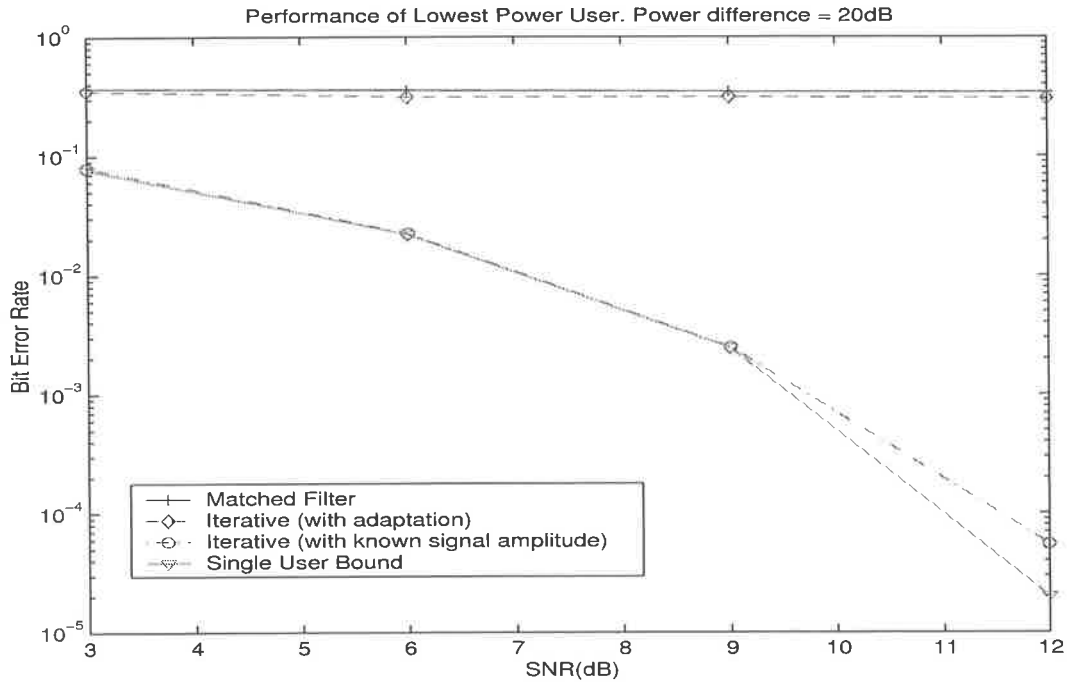
Figure 2.10: Performance of the lowest power user: -20dB relative to other users
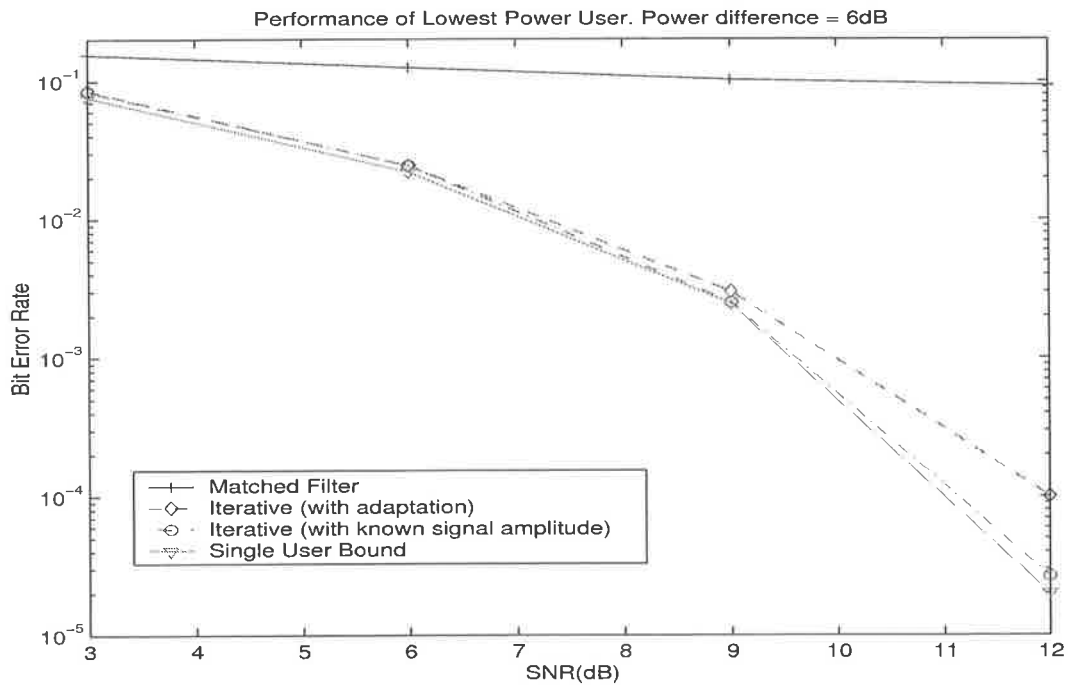
Figure 2.11: Performance of the lowest power user: -6dB relative to other users

adaptive version in this case. This suggests that, although the iterative detector with parameter estimation is not near-far resistant, it does not require tight power control in order to perform well.

2.5 Conclusions

In this chapter we have presented a number of interference cancellation detectors. These include the successive interference cancellation (SIC) detector, the parallel interference cancellation (PIC) detector and two proposed hybrid PIC-SIC detectors. Through computer simulations we showed that by combining the PIC and the SIC techniques, the hybrid PIC-SIC detectors inherit many good behaviours of these two detectors. We have also presented in this chapter an adaptive multiuser detector which performs joint symbol detection and parameter estimation in CDMA systems. Simulation results show that if the relative power difference between users in the system is not too large, the performance of this adaptive multiuser detector is very close to the nonadaptive version (i.e where the signal amplitudes are perfectly known at the receiver).

Chapter 3

Convergence Behaviour Analysis of the PIC Technique

This chapter analyses the convergence behaviour of the parallel interference cancellation (PIC) technique in code division multiple access (CDMA) systems. It introduces a general PIC detector model where it can be used to describe the interference cancellation process of the PIC detector with any type of tentative decisions. With this model, the PIC detector can be viewed as either a feedback system or an iterated-map neural network. Using the known results from previous stability analysis in these areas, the chapter develops a general framework for analysing the convergence behaviour of the PIC detector. This framework permits the derivation of the sufficient condition for convergence of the PIC detector for a wide range of tentative decision functions. As examples, the chapter derives the sufficient conditions for convergence of the PIC detector with linear decision, clip decision and hyperbolic tangent decision functions. The chapter shows that some well-known conditions for convergence of the PIC detector with linear decision and clip decision functions can be derived using this general framework. The chapter also examines the convergence behaviour of the PIC detector with hyperbolic tangent decision

function via computer simulation and compares it with analytical results.

3.1 Introduction

Since the PIC detector performs interference cancellation in multiple stages, it is of interest to ascertain what the behaviour would be as the number of interference cancellation stages L approaches infinity. This behaviour is largely dependent on the types of tentative decision function that the PIC detector employed for estimating the multiple access interference (MAI) at the end of each interference cancellation stages. In recent years there are a number of published works addressing the convergence behaviour issue of the PIC detector for various types of tentative decision function. However, these works analyse the convergence behaviour of the PIC detector for each type of tentative decision function separately. The convergence behaviour of the PIC detector with linear decisions (also known as the linear PIC detector), receives the most attention [15, 17, 18, 39, 68]. It is found that when the spectral radius of the matrix $(\mathbf{R} - \mathbf{I})$ is less than 1, where $\mathbf{R}$ is the crosscorrelation matrix of the spreading codes, the linear PIC detector converges to the decorrelator detector as $L \to \infty$. The convergence behaviour of the PIC detector with clip decision function is analysed in [108] using a nonlinear programming approach. This PIC detector is found to converge to a fixed point if the maximum eigenvalue of the matrix $(\mathbf{R} - \mathbf{I})$ is less than 1. In [100, pp. 363], it has been observed that the PIC detector with hard decisions does not always converge to a fixed point as $L \to \infty$. Later, it has been shown in [56] that the PIC detector with hard decisions corresponds to a special case of a Hopfield neural network [46]. The existence of the period-two limit cycle for this special case is proved in [16]. This cyclic behaviour is also being observed in [76] for the iterative multiuser detector that utilises the *a posteriori* probabilities to estimate the transmitted symbol. For a system with no interchip interference, this iterative multiuser detector corresponds to the PIC de-

tector with hyperbolic tangent decisions. Using the theory of contraction mappings, a sufficient condition for convergence of this particular detector is derived. In [76], it is shown that when the initialisation is carefully chosen such that the sufficient condition is satisfied, convergence to a fixed point will always occur. On the other hand when the sufficient condition is not satisfied, it may lead to the cyclic behaviour.

Given that there are many different types of tentative decision function that can be used by the PIC detector for estimating the MAI at the end of each interference cancellation stage, our aim in this study is to develop a general framework that would allow us to analyse the convergence behaviour of the PIC detector for any type of tentative decision function. In this chapter, we approach the parallel interference cancellation problem from two different perspectives. First, we view the PIC detector as a feedback system and apply some known results from the stability analysis of the latter to investigate the convergence behaviour of the PIC detector. We derive a general condition from which the sufficient condition for convergence of the PIC detector for a wide range of tentative decision functions can be calculated. Second, we establish a one to one correspondence between the PIC detector and an iterated-map neural network. Using this relationship, we apply the stability analysis of the latter to the convergence behaviour study of the PIC detector. We prove that the PIC detector with any nonlinear tentative decision function that is monotonically increasing at a sublinear rate will either converge to a fixed point or enter a period-two limit cycle. In addition, we derive the sufficient condition which guarantees that the PIC detector with these types of tentative decision function will always converge. Our analysis from both approaches reveals that by placing a bound on the maximum slope of the tentative decision function, the PIC detector will always converge to a fixed point for a given input and noise realisation. To the best of our knowledge, there is no such general framework similar to ours has been known. In [56], a connection between the PIC detector with hard decisions and a

neural network has been established similar to our second approach. However, the convergence analysis in [56] is only applied to the case of hard decision function whereas the analysis in our second approach is more general.

The remaining of this chapter is organised as follows: In section 3.2, a synchronous CDMA system model is presented. In section 3.3, we introduce a general PIC detector model where it can be used to describe the interference cancellation process of the PIC detector with any tentative decision function. Sections 3.4 and 3.5 are the main parts of this chapter where we analyse the convergence behaviour of the PIC detector from both feedback system and iterated-map neural network perspectives. In section 3.6, computer simulation is used to examine the convergence behaviour of the PIC detector with hyperbolic tangent decision function.

3.2   The System Model

We consider a synchronous CDMA system with K users transmitting simultaneously to a common receiver over an additive white Gaussian noise (AWGN) channel. Each user $k \in \{1, ..., K\}$ is assigned a normalized signature sequence $\mathbf{s}_k$ of length M, $\mathbf{s}_k \in \{\frac{-1}{\sqrt{M}}, \frac{1}{\sqrt{M}}\}^M$ and transmits at symbol rate the information $d_k(i) \in \{\pm 1\}$. We use the time indices t and i to denote the chip index and symbol index, respectively. The transmitted signal of the kth user during the tth chip interval is

$$y_k(t) = A_k(i)\, d_k(i)\, s_k(j) \, , \tag{3.1}$$

where $A_k(i) > 0$ is the signal amplitude of user k over symbol period i, $d_k(i)$ is the ith symbol and $t = Mi + j$, for $0 \leq j \leq M - 1$.

Thus, the baseband received signal at the chip rate is the noisy sum of all signals, which can be written as

$$r(t) = \sum_{k=1}^{K} y_k(t) + v(t) \, , \tag{3.2}$$

where $v(t)$ is a realization at time t of the zero mean white Gaussian noise with variance σ^2.

Now, let $\mathbf{r}(i)$ consists of M consecutive observations of the received data at the chip rate $t = Mi$ to $M(i+1) - 1$ i.e. $\mathbf{r}(i) = [r(Mi)\ r(Mi+1)\ \cdots\ r(M(i+1)-1)]^T$ where T denotes the transpose operation. Similarly, we use

$$\mathbf{v}(i) = [v(Mi)\ v(Mi+1)\ \cdots\ v(M(i+1)-1)]^T\ , \tag{3.3}$$

$$\mathbf{y}_k(i) = [y_k(Mi)\ y_k(Mi+1)\ \cdots\ y_k(M(i+1)-1)]^T \tag{3.4}$$

for the corresponding noise vector and vector of user k contribution signals, respectively. Therefore, the symbol rate version of (3.2) is:

$$\mathbf{r}(i) = \sum_{k=1}^{K} \mathbf{y}_k(i) + \mathbf{v}(i) \text{ with } \mathbf{y}_k(i) = A_k(i)\, d_k(i)\, \mathbf{s}_k\ . \tag{3.5}$$

3.3 The PIC Detector

The underlying principle of interference cancellation methods is that each user estimates the multiuser interference contributed by other users and subtracts it from the received signal to produce a better estimate of the desired signal. The PIC detector can be implemented with multiple interference cancellation stages [97, 109] where the tentative decisions at the prior stage (sometime called iteration) are used to generate the MAI estimates for the current stage. The idea here is that the accuracy of the tentative decisions are presumably improved as the number of stages increases and hence the PIC detector can suppress more and more interference. A common feature of this type of PIC detection is that at each stage the detector performs total interference cancellation where it attempts to completely remove the interference caused by other users by subtracting all the estimated MAI from the received signal. An alternative approach is to perform partial interference cancellation as presented in [31] where only part of the estimated MAI is subtracted from

the received signal at each stage. The idea is that for the earlier stages, the detector cancels only a small fraction of the MAI since at these earlier stages the tentative decisions are less reliable. For later stages, the tentative decisions are presumably more reliable and the detector will attempt to cancel more MAI. It is shown in [31] that this detector has a significant capacity gain as compared to the conventional matched filter. In this chapter, we only examine the convergence behaviour of the PIC detector which performs total interference cancellation.

Let's consider a PIC detector that employs any arbitrary tentative decision function for estimating the MAI at the end of each interference cancellation stage. Suppose the PIC detector performs L interference cancellation stages, then at the nth stage where $n = 1, ..., L$ the estimate for the desired signal can be calculated by one of the following set of equations depending on the type of tentative decision function that is employed.

First form:

$$\hat{\mathbf{y}}_k^{(n)}(i) = \mathbf{r}(i) - \sum_{j \neq k} A_j(i)\, \tilde{d}_j^{(n-1)}(i)\, \mathbf{s}_j \,, \tag{3.6}$$

$$\tilde{d}_k^{(n)}(i) = g_k\left(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(n)}(i)\right) \,, \tag{3.7}$$

where $\tilde{d}_k^{(0)}(i) = g_k\left(\mathbf{s}_k^T \mathbf{r}(i)\right)$ and $g_k(x)$ is the tentative decision function of user k. Examples of the tentative decision functions that can be represented in this form are: the hard decision function $g_k(x) = \text{sgn}(x)$ and the hyperbolic tangent decision function $g_k(x) = \tanh(\alpha_k x)$ where α_k are positive constants.

Second form:

$$\hat{\mathbf{y}}_k^{(n)}(i) = \mathbf{r}(i) - \sum_{j \neq k} \tilde{b}_j^{(n-1)}(i)\, \mathbf{s}_j \,, \tag{3.8}$$

$$\tilde{b}_k^{(n)}(i) = h_k\left(\mathbf{s}_k^T\, \hat{\mathbf{y}}_k^{(n)}(i)\right) \,, \tag{3.9}$$

where $\tilde{b}_k^{(0)}(i) = h_k\left(\mathbf{s}_k^T \mathbf{r}(i)\right)$ and $h_k(x)$ is the tentative decision function of user k. This second form can be described using the first form by setting $g_k(x) = \frac{1}{A_k} h_k(x)$,

and vice versa. Since the analysis in this chapter can be applied to any type of tentative decision functions, it is advantageous to have two different forms as we find that for some tentative decision functions one form leads to a stronger sufficient condition for convergence than the other. It should be noted that even though the second form (3.8) does not explicitly include the signal amplitude A_k, it does not imply that second form can only be used for PIC detection technique which requires no knowledge of the signal amplitude since the information of the signal amplitude can be embedded into the function $h_k(x)$. Examples of tentative decision functions that can be described in this form are: the linear decision function with $h_k(x) = x$ and the clip decision function with

$$h_k(x) = \begin{cases} A_k & x > A_k \\ x & A_k \geq x \geq -A_k \\ -A_k & x < -A_k \end{cases} \tag{3.10}$$

In the following sections, we investigate the convergence behaviour of the PIC detector with an arbitrary tentative decision function from both feedback system and iterated-map neural network perspectives. To simplify the notation in the analysis, we drop the symbol time index i since we are only interested in the behaviour of the PIC detector in relation to the interference cancellation stage index n.

3.4 Convergence Behaviour: A Feedback System Perspective

In the following, we employ a general method for analysing the stability of feedback systems which is known as the small gain approach. This method can be used to study the ℓ_p-stability of the feedback systems for any value $p \in [1, \infty]$.

3.4.1 Background on Feedback Systems

Let S be the space of sequences in $\mathbb{R}^K$, i.e., the sequence $\mathbf{p} \in S$, iff $\mathbf{p} = (\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \dots)$ with $\mathbf{p}^{(n)} \in \mathbb{R}^K$ for $n = 1, 2, \dots$. The ℓ_p-norms of the sequence $\mathbf{p}$ are defined as

$$\|\mathbf{p}\|_p = \left(\sum_{n=1}^{\infty} |\mathbf{p}^{(n)}|^p \right)^{1/p} \quad , \text{ for } 1 \le p < \infty \tag{3.11}$$

$$\|\mathbf{p}\|_\infty = \sup_{n \ge 1} |\mathbf{p}^{(n)}| \, , \tag{3.12}$$

where $|.|$ denotes any norms on $\mathbb{R}^K$ or the absolute value of a number in $\mathbb{R}$. Let $\mathbf{p}_N = (\mathbf{p}^{(1)}, \mathbf{p}^{(2)}, \dots, \mathbf{p}^{(N)})$ denotes a truncated sequence of $\mathbf{p}$. The corresponding ℓ_p-norms of the truncated sequence are defined as

$$\|\mathbf{p}_N\|_p = \left(\sum_{n=1}^{N} |\mathbf{p}^{(n)}|^p \right)^{1/p} \quad , \text{ for } 1 \le p < \infty \tag{3.13}$$

$$\|\mathbf{p}_N\|_\infty = \sup_{1 \le n \le N} |\mathbf{p}^{(n)}| \, . \tag{3.14}$$

Let the extended space ℓ_{pe}^K be defined by

$$\ell_{pe}^K = \{ \mathbf{p} \in S \,|\, \forall N \in \mathbb{Z}_+, \|\mathbf{p}_N\|_p < \infty \} \, . \tag{3.15}$$

Thus, ℓ_{pe}^K is the space of all $\mathbf{p} \in S$ with the property that $\forall N \in \mathbb{Z}_+$, the truncated sequence $\mathbf{p}_N$ has finite ℓ_p-norm.

We now state the small gain and the incremental small gain theorems. The proofs of these theorems are omitted here and readers are referred to [28] for detailed proofs.

Theorem 3.1 (Small Gain Theorem) *Consider the feedback system as shown in Fig. 3.1 and suppose $p \in [1, \infty]$ is specified. Let $\mathbf{H}_1$, $\mathbf{H}_2 : \ell_{pe}^K \to \ell_{pe}^K$. Suppose there are constants $\gamma_1 \ge 0$ and $\gamma_2 \ge 0$ such that $\forall N \in \mathbb{Z}_+$ and $\forall \mathbf{v}_i \in \ell_{pe}^K$ with $i = 1, 2$*

$$\|(\mathbf{H}_i \mathbf{v}_i)_T\|_p \le \gamma_i \|\mathbf{v}_{iT}\|_p \, . \tag{3.16}$$
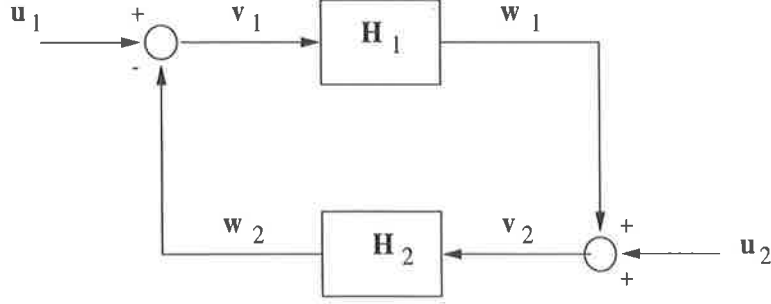
Figure 3.1: General feedback system

Under these conditions, if

$$\gamma_1 \gamma_2 < 1 \quad and \tag{3.17}$$

$$\|\mathbf{u}_1\|_p , \|\mathbf{u}_2\|_p < \infty , \tag{3.18}$$

then the outputs $\mathbf{w}_1$, $\mathbf{w}_2$ *of the feedback system as shown in Fig. 3.1 have finite* ℓ_p-*norms.*

Theorem 3.2 (Incremental Small Gain Theorem) *Consider the feedback system as shown in Fig. 3.1 and suppose* $p \in [1, \infty]$ *is specified. Let* $\mathbf{H}_1$, $\mathbf{H}_2 : \ell_{pe}^K \to \ell_{pe}^K$. *Suppose there are constants* $\tilde{\gamma}_1 \geq 0$ *and* $\tilde{\gamma}_2 \geq 0$ *such that* $\forall N \in \mathbb{Z}_+$ *and* $\forall \mathbf{v}_i \in \ell_{pe}^K$ *with* $i = 1, 2$

$$\|(\mathbf{H}_i \mathbf{v}_i)_T - (\mathbf{H}_i \mathbf{v}'_i)_T\|_p \leq \tilde{\gamma}_i \|\mathbf{v}_{iT} - \mathbf{v}'_{iT}\|_p . \tag{3.19}$$

Under these conditions, if

$$\tilde{\gamma}_1 \tilde{\gamma}_2 < 1 , \tag{3.20}$$

then $\forall \mathbf{u}_1, \mathbf{u}_2 \in \ell_{pe}^K \exists$ *a unique solution* $\mathbf{w}_1, \mathbf{w}_2 \in \ell_{pe}^K$ *which can be obtained by iteration.*

3.4.2 Application to the PIC Detector

To facilitate the analysis in this section, we will assume that there is no background noise in the received signal. Now, by substituting $\mathbf{r}$ from (3.5) into (3.6), we can

rewrite the PIC detector of the first form (3.6) as

$$\mathbf{s}_k^T \hat{\mathbf{y}}_k^{(n)} - A_k d_k = \sum_{j \neq k} A_j \rho_{j,k} \left(d_j - \tilde{d}_j^{(n-1)} \right) . \tag{3.21}$$

Let's denote the estimation error of the kth user's desired signal on the l.h.s. of (3.21) by $e_k^{(n)}$. Then

$$e_k^{(n)} = \sum_{j \neq k} \rho_{j,k} A_j \left(d_j - g_j(e_j^{(n-1)} + A_j d_j) \right) , \tag{3.22}$$

or in vector form

$$\mathbf{e}^{(n)} = (\mathbf{R} - \mathbf{I}) \, \mathbf{A} \, \left(\mathbf{d} - \mathbf{g}(\mathbf{e}^{(n-1)} + A \, \mathbf{d}) \right) , \tag{3.23}$$

where $\mathbf{g}(.)$ acts componentwise. We thus have a feedback (on index n) system as shown in Fig. 3.1 with

$$\mathbf{H}_1 = (\mathbf{R} - \mathbf{I})\mathbf{A} \; ; \quad \mathbf{H}_2 = \mathbf{g}(.) \; ; \tag{3.24}$$

$$\mathbf{u}_1^{(n)} = \mathbf{d} \; ; \qquad \mathbf{u}_2^{(n)} = \mathbf{A} \, \mathbf{d} \; ; \tag{3.25}$$

and the output $\mathbf{w}_1^{(n)} = \mathbf{e}^{(n)}$.

Since $\mathbf{u}_1^{(n)} = \mathbf{d}$ and $\mathbf{u}_2^{(n)} = \mathbf{A} \, \mathbf{d} \; \forall n \in \mathbb{Z}_+$, the ℓ_p-norms of the inputs sequence $\mathbf{u}_1, \mathbf{u}_2$ are not finite for any $p \in [1, \infty)$. Hence $\mathbf{u}_1, \mathbf{u}_2 \notin \ell_{p,e}^K$ and therefore we cannot analyse the ℓ_p-stability of the PIC detector using the small gain theorem for any $p \in [1, \infty)$. However, their ℓ_∞-norms are finite. Thus, we can use the small gain theorem to analyse the ℓ_∞-stability of the PIC detector.

Proposition 3.1 *Let's define*

$$\varphi_k^g = \sup_{x \neq 0} \frac{|g_k(x)|}{|x|} . \tag{3.26}$$

For the PIC detector of the first form (3.6), the error in estimating each user's signal is bounded if

$$\max_k \varphi_k^g < \frac{1}{\|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_\infty} , \tag{3.27}$$

where the ℓ_∞ operator norm is given by

$$\|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_\infty = \max_k \sum_{j \neq k} A_j |\rho_{j,k}| \ . \tag{3.28}$$

Proof: By Lemma 3.1 and Lemma 3.2 in the Appendix, the corresponding values of γ_1 and γ_2 which satisfy the conditions

$$\|(\mathbf{R} - \mathbf{I})\mathbf{A}\,\mathbf{x})_N\|_\infty \leq \gamma_1 \|\mathbf{x}_N\|_\infty \ , \tag{3.29}$$

$$\|(\mathbf{g}(\mathbf{x}))_N\|_\infty \leq \gamma_2 \|\mathbf{x}_N\|_\infty \ , \tag{3.30}$$

$\forall \mathbf{x} \in \ell_{\infty e}^K$ and $\forall N \in \mathbb{Z}_+$ are

$$\gamma_1 = \|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_\infty \ , \tag{3.31}$$

$$\gamma_2 = \max_k \varphi_k^g \ . \tag{3.32}$$

By the small gain theorem, if $\gamma_1 \gamma_2 < 1$, the PIC detector is then ℓ_∞-stable and thus the ℓ_∞-norm of the sequence of errors $\mathbf{e}^{(n)}$ is finite. $\triangle$

The above result only ensures that the error in estimating each user's signal is bounded and does not guarantee that the PIC detector will converge. We now show that by imposing an additional restriction on the tentative decision function, we can ensure that the PIC detector will always converge. As it turns out, this restriction is to have the maximum slope of the tentative decision function to be less than (or equal to) its maximum gain.

Proposition 3.2 *Let's define*

$$\psi_k^g = \sup_x \frac{dg_k(x)}{dx} \ . \tag{3.33}$$

Under the conditions that

$$\max_k \varphi_k^g < \frac{1}{\|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_\infty} \quad and \tag{3.34}$$

$$|\psi_k^g| \leq \varphi_k^g \quad \forall k, \tag{3.35}$$

the PIC detector of the first form (3.6) will converge to a fixed point where the soft estimates $\tilde{\mathbf{d}}^{(n)} \to \mathbf{d}^*$ *as* $n \to \infty$.

Proof: By Lemma 3.3 in the Appendix, a value of $\tilde{\gamma}_1$ which satisfies the condition

$$\|((\mathbf{R} - \mathbf{I})\mathbf{A}\,\mathbf{x})_N - ((\mathbf{R} - \mathbf{I})\mathbf{A}\,\mathbf{x}')_N\|_\infty \leq \tilde{\gamma}_1 \|\mathbf{x}_N - \mathbf{x}'_N\|_\infty \,, \tag{3.36}$$

$\forall \mathbf{x},\ \mathbf{x}' \in \ell^K_{\infty e}$ and $\forall N \in \mathbb{Z}_+$, is

$$\tilde{\gamma}_1 = \|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_\infty \,. \tag{3.37}$$

As proved in Lemma 3.4 in the Appendix, provided that $|\psi_k^g| \leq \varphi_k^g\ \forall k$ the condition

$$\|(\mathbf{g}(\mathbf{x}))_N - (\mathbf{g}(\mathbf{x}'))_N\|_\infty \leq \tilde{\gamma}_2 \|\mathbf{x}_N - \mathbf{x}'_N\|_\infty \,, \tag{3.38}$$

is satisfied $\forall \mathbf{x},\ \mathbf{x}' \in \ell^K_{\infty e}$ and $\forall N \in \mathbb{Z}_+$, if

$$\tilde{\gamma}_2 = \max_k \varphi_k^g \,. \tag{3.39}$$

Hence, by the incremental small gain theorem, the errors $\mathbf{e}^{(n)} \to \mathbf{e}^*$ as $n \to \infty$ if $\tilde{\gamma}_1 \tilde{\gamma}_2 < 1$. From the definition of $\mathbf{e}^{(n)}$ in (3.23), this implies that the soft estimates $\tilde{\mathbf{d}}^{(n)}$ approaches a fixed point $\mathbf{d}^*$. $\triangle$

In order to obtain a more intuitive interpretation of the results derived previously, we will show that these conditions are equivalent to having the tentative decision function to satisfy a certain sector condition. Let's define $\phi : \mathbb{R} \to \mathbb{R}$ with $\phi(0) = 0$. The function $\phi \in$ **sector** $(\eta_1,\ \eta_2)$, where $\eta_1, \eta_2 \in \mathbb{R}$ with $\eta_1 \leq \eta_2$, iff

$$\eta_1 x^2 < x\,\phi(x) < \eta_2 x^2 \,, \quad \forall x \in \mathbb{R} \ \text{with} \ x \neq 0 \tag{3.40}$$

Thus, the condition (3.27) given in Proposition 3.1 is equivalent to

$$g_k(x) \in \ \textbf{sector} \ (-\varphi_{max}^g,\ \varphi_{max}^g) \ \forall k \,, \tag{3.41}$$

where $\varphi_{max}^g = \frac{1}{\|(\mathbf{R} - \mathbf{I})\,\mathbf{A}\|_\infty}$. A geometric interpretation of the above result is illustrated in Fig. 3.2. It says that if the tentative decision function for all users lies
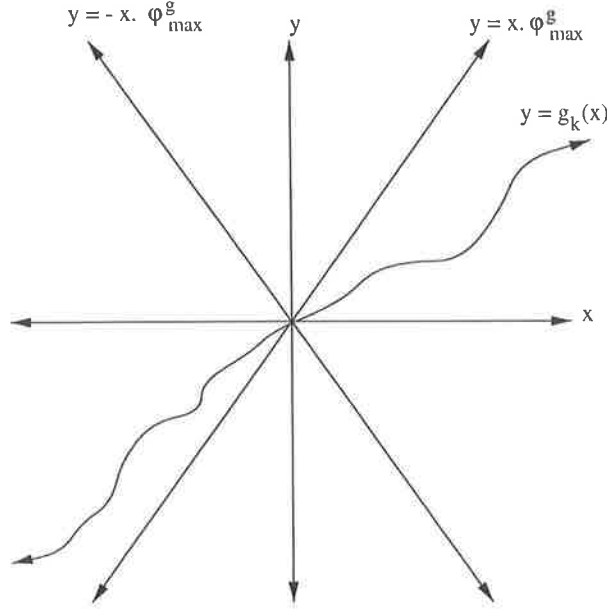
Figure 3.2: Geometric interpretation of the sufficient condition for convergence

within the region that includes the x-axis and bounded by the functions $g(x) = -\varphi^g_{max} x$ and $g(x) = \varphi^g_{max} x$, then the sequence of errors $\mathbf{e}$ is bounded. In addition, if the tentative decision function of all users has a maximum slope less than (or equal to) its maximum gain, i.e. $|\psi_k| \leq \varphi_k \; \forall k$, then the soft estimates $\tilde{\mathbf{d}}^{(n)} \to \mathbf{d}^*$ as $n \to \infty$. It should be noted that if a function $g_k(x) \in \mathbf{sector} \; (0, \varphi^g_{max})$ leads to $\tilde{d}^{(n)}_k \to d^*_k$, then there exists another function in $\mathbf{sector} \; (-\varphi^g_{max}, 0)$, which is a mirror reflection of the function $g_k(x) \in \mathbf{sector} \; (0, \varphi^g_{max})$ on the x-axis, that would lead to $\tilde{d}^{(n)}_k \to -d^*_k$.

Following the previous steps, the PIC detector in the second form (3.8) can be rewritten as

$$e^{(n)}_k = \sum_{j \neq k} \rho_{j,k} \left(A_j \, d_j - h_j(e^{(n-1)}_j + A_j \, d_j) \right) , \qquad (3.42)$$

or in vector form

$$\mathbf{e}^{(n)} = (\mathbf{R} - \mathbf{I}) \left(\mathbf{A} \, \mathbf{d} - \mathbf{h}(\mathbf{e}^{(n-1)} + \mathbf{A} \, \mathbf{d}) \right) , \qquad (3.43)$$

where $\mathbf{h}(.)$ acts componentwise. This is in the feedback form as given by Fig. 3.1

with

$$\mathbf{H}_1 = (\mathbf{R} - \mathbf{I}) \; ; \qquad \mathbf{H}_2 = \mathbf{h}(.) \; ; \tag{3.44}$$

$$\mathbf{u}_1^{(n)} = \mathbf{A}\,\mathbf{d} \; ; \qquad \mathbf{u}_2^{(n)} = \mathbf{A}\,\mathbf{d} \; ; \tag{3.45}$$

and the output $\mathbf{w}_1^{(n)} = \mathbf{e}^{(n)}$. We now state the stability and convergence results of the PIC detector of the second form in the next two propositions, respectively. The proofs for these results are similar to those in Proposition 3.1 and Proposition 3.2 and therefore are omitted here.

Proposition 3.3 *Let's define*

$$\varphi_k^h = \sup_{x \neq 0} \frac{|h_k(x)|}{|x|} \; . \tag{3.46}$$

For the PIC detector of the second form (3.8), the error in estimating each user's signal is bounded if

$$\max_k \varphi_k^h < \frac{1}{\|\mathbf{R} - \mathbf{I}\|_\infty} \; . \tag{3.47}$$

Proposition 3.4 *Let's define*

$$\psi_k^h = \sup_x \frac{dh_k(x)}{dx} \; . \tag{3.48}$$

Under the conditions that

$$\max_k \varphi_k^h < \frac{1}{\|\mathbf{R} - \mathbf{I}\|_\infty} \quad and \tag{3.49}$$

$$|\psi_k^h| \le \varphi_k^h \; \forall k \; , \tag{3.50}$$

the PIC detector of the second form (3.8) will converge to a fixed point where the soft estimates $\tilde{\mathbf{b}}^{(n)} \to \mathbf{b}^*$ as $n \to \infty$.

3.4.3 Examples

As examples, we apply the results derived previously to obtain the sufficient condition for convergence of the PIC detector for a number of tentative decision functions. These tentative decision functions include the hyperbolic tangent decision, the linear decision and the clip decision functions. In addition, we also compare the new sufficient conditions for convergence of the PIC detector that employs linear decision and clip decision functions with some well-known conditions.

Hyperbolic Tangent Decision Function

For the PIC detector with hyperbolic tangent decision function $\tanh(\alpha_k x)$, we can put it into the first form (3.6) by having $g_k(x) = \tanh(\alpha_k x)$. For $g_k(x) = \tanh(\alpha_k x)$, it can be shown that the maximum gain φ_k^g and the maximum slope ψ_k^g are equal

$$\psi_k^g = \varphi_k^g = \sup_{x \neq 0} \frac{|g_k(x)|}{|x|} \, , \tag{3.51}$$

$$= \alpha_k \, . \tag{3.52}$$

Thus, by Proposition 3.2 the sufficient condition for convergence to a fixed point is therefore

$$\max_k \, \alpha_k < \frac{1}{\|(\mathbf{R} - \mathbf{I}) \, \mathbf{A}\|_\infty} \, . \tag{3.53}$$

This condition is quite different to that presented in [76]. The sufficient condition for convergence derived in [76] involves finding the initialisations such that the requirement for contraction mapping is met.

Linear PIC detector

The convergence behaviour of the linear PIC detector is studied in [15, 17, 18, 39, 68]. Suppose the linear PIC detector has L interference cancellation stages, the estimate

of the desired signal of user k at the nth stage is

$$\hat{\mathbf{y}}_k^{(n)} = \mathbf{r} - \sum_{j \neq k} \mathbf{s}_j^T \, \hat{\mathbf{y}}_j^{(n-1)} \, \mathbf{s}_j \; . \tag{3.54}$$

Let's define $z_k^{(n)} = \mathbf{s_k}^T \hat{\mathbf{y}}_k^{(n)}$ and $\mathbf{z}^{(n)} = [\; z_1^{(n)} \; z_2^{(n)} \; ... \; z_K^{(n)} \;]^T$, the linear PIC detector can be described in matrix form as [15]

$$\mathbf{z}^{(n)} = \tilde{\mathbf{r}} - (\mathbf{R} - \mathbf{I})\mathbf{z}^{(n-1)} \; ,$$

$$\mathbf{z}^{(0)} = \tilde{\mathbf{r}} \; ,$$

$$\hat{\mathbf{d}}_{LPIC} = \mathrm{sgn}(\mathbf{z}^{(L)}) \tag{3.55}$$

where $\tilde{\mathbf{r}} = \mathbf{R}\,\mathbf{A}\,\mathbf{d} + \tilde{\mathbf{v}}$ is the K-dimensional matched filter output, $\mathbf{R} \in \mathbb{R}^{K \times K}$ is the symmetric correlation matrix with elements $[\mathbf{R}]_{j,k} = \rho_{j,k} = \mathbf{s}_k^T \mathbf{s}_j$, $\mathbf{A} \in \mathbb{R}^{K \times K}$ is the diagonal matrix of signal amplitudes, $\mathbf{d} \in \{\pm 1\}^K$ is the vector of binary user symbols and $\tilde{\mathbf{v}}$ is the zero mean Gaussian noise with variance $E[\tilde{\mathbf{v}}\tilde{\mathbf{v}}^T] = \mathbf{R}$. Using (3.54), the decision statistic of the linear PIC detector at the final stage is

$$\mathbf{z}^{(L)} = \sum_{\ell=0}^{L} (-1)^{\ell} (\mathbf{R} - \mathbf{I})^{\ell} \tilde{\mathbf{r}} \; . \tag{3.56}$$

It can be shown that

$$\lim_{L \to \infty} \sum_{\ell=0}^{L} (-1)^{\ell} (\mathbf{R} - \mathbf{I})^{\ell} = \mathbf{R}^{-1} \; , \tag{3.57}$$

if the spectral radius of the matrix $(\mathbf{R} - \mathbf{I})$ is less than 1. The spectral radius of the matrix $\boldsymbol{\Theta}$ is defined as $\rho(\boldsymbol{\Theta}) = \max\{|\lambda| : \lambda \in \sigma(\boldsymbol{\Theta})\}$ where $\sigma(\boldsymbol{\Theta})$ denotes the spectrum of $\boldsymbol{\Theta}$ [47, pp. 35]. Thus, as the number of interference cancellation stages approaches infinity the linear PIC detector converges to the decorrelator detector if $\rho(\mathbf{R} - \mathbf{I}) < 1$.

The linear PIC detector (3.54) can be expressed in the second form (3.8) by letting $h_k(x) = x \; \forall k$. For $h_k(x) = x$, the maximum gain φ_k^g and the maximum slope

ψ_k^g are also equal

$$\psi_k^h = \varphi_k^h = \sup_{x \neq 0} \frac{|h_k(x)|}{|x|} \; , \tag{3.58}$$

$$= 1. \tag{3.59}$$

Hence, using Proposition 3.4 the sufficient condition for convergence to a fixed point is

$$\|\mathbf{R} - \mathbf{I}\|_\infty < 1 \; . \tag{3.60}$$

This condition is the same as the diagonal dominance condition

$$\sum_{\substack{j \neq k}}^{K} |\rho_{j,k}| < |\rho_{k,k}| = 1 \quad \forall k \tag{3.61}$$

reported in [17]. For any $\boldsymbol{\Theta} \in \mathbb{R}^{K \times K}$, it is shown in [28, pp. 27] that

$$\inf_{p \in [1,\infty]} \sup_{\mathbf{x} \neq 0} \frac{\|\boldsymbol{\Theta}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \rho(\boldsymbol{\Theta}) \; . \tag{3.62}$$

Thus for any operator norm, $\rho(\boldsymbol{\Theta}) \leq \|\boldsymbol{\Theta}\|_p$. This suggests that the well known condition for convergence $\rho(\mathbf{R} - \mathbf{I}) < 1$ is weaker (less restrictive) than that derived using this approach.

However, we can obtain the same condition for convergence as the well-known condition $\rho(\mathbf{R} - \mathbf{I}) < 1$ by noting that for the linear PIC detector case, (3.43) is equivalent to

$$\mathbf{e}^{(n)} = -(\mathbf{R} - \mathbf{I})\mathbf{e}^{(n-1)} \; . \tag{3.63}$$

Thus, the errors $\mathbf{e}^{(n)} \to 0$ as $n \to \infty$ if $\rho(\mathbf{R} - \mathbf{I}) < 1$. This implies that when there is no background noise, the soft estimates $\tilde{\mathbf{b}}^{(n)} \to \mathbf{Ad}$ as $n \to \infty$ if $\rho(\mathbf{R} - \mathbf{I}) < 1$. This is what one would expect because for $\rho(\mathbf{R} - \mathbf{I}) < 1$ the linear PIC detector converges to the decorrelator and it is known that in the absence of background noise the decorrelator can perfectly recover the transmitted information bits.

Alternatively, the well-known condition for convergence of linear PIC detector can also be obtained by realising that (3.63) is in the same feedback form as shown in Fig. 3.1 with

$$\mathbf{H}_1 = (\mathbf{R} - \mathbf{I}) ; \quad \mathbf{H}_2 = \mathbf{I} ; \tag{3.64}$$

$$\mathbf{u}_1^{(n)} = \mathbf{0} ; \quad \mathbf{u}_2^{(n)} = \mathbf{0} ; \tag{3.65}$$

and $\mathbf{w}_1^{(n)} = \mathbf{e}^{(n)}$. Since the ℓ_2-norms of the sequence $\mathbf{u}_1$, $\mathbf{u}_2$ are finite (equal zero in this case), we can therefore analyse the ℓ_2-stability of this detector. Using the small gain theorem, it can be shown that if

$$\|\mathbf{R} - \mathbf{I}\|_2 < 1 , \tag{3.66}$$

the linear PIC detector is ℓ_2-stable and thus the error sequence $\mathbf{e}$ will have finite ℓ_2-norm. From the definition of the ℓ_2-norm, if the error sequence $\mathbf{e} = (\mathbf{e}^{(1)}, \mathbf{e}^{(2)}, \ldots)$ has finite ℓ_2-norm, then $\mathbf{e}^{(n)} \to 0$ as $n \to \infty$. Hence by showing that the linear PIC detector is ℓ_2-stable, we can further conclude that the sequence of soft estimates $\tilde{\mathbf{b}}^{(n)} \to \mathbf{Ad}$ as $n \to \infty$. Since the matrix $(\mathbf{R} - \mathbf{I})$ is real symmetric, $\rho(\mathbf{R} - \mathbf{I}) = \|\mathbf{R} - \mathbf{I}\|_2$. Thus, we have the condition for convergence which is the same as the well-known result.

Clip decision function

The convergence behaviour of the PIC detector with clip decision function has been studied in [108]. It is found that this detector converges to a fixed point if the maximum eigenvalue of the matrix $(\mathbf{R} - \mathbf{I})$ is less than 1. For the PIC detector with clip decision function, we can express it in the second form (3.8) with $h_k(x)$ defined as in (3.10). Similar to linear decision function, the clip decision function has equal

maximum gain and maximum slope

$$\psi_k^h = \varphi_k^h = \sup_{x \neq 0} \frac{|h_k(x)|}{|x|} \ , \tag{3.67}$$

$$= 1. \tag{3.68}$$

Thus, by Proposition 3.4 the sufficient condition for convergence to a fixed point is therefore

$$\|\mathbf{R} - \mathbf{I}\|_\infty < 1 \ , \tag{3.69}$$

which is stronger than the known condition $\lambda_{max}(\mathbf{R} - \mathbf{I}) < 1$ derived in [108] since

$$\|\mathbf{R} - \mathbf{I}\|_\infty \geq \rho(\mathbf{R} - \mathbf{I}) \tag{3.70}$$

$$\geq \lambda_{max}(\mathbf{R} - \mathbf{I}) \ . \tag{3.71}$$

We will obtain this known condition using the analysis presented in the next section.

3.5 Convergence Behaviour Analysis: A Neural Network Perspective

The analysis in the previous section assumes that there is no background noise. In this section, the presence of background noise is taken into account. We will firstly prove that the PIC detector with any nonlinear tentative decision function that is monotonically increasing at a sublinear rate, i.e. magnitude increases less than linearly, will either converge to a fixed point or enter a period-two limit cycle. We then derive a sufficient condition which guarantees that the PIC detector with these types of tentative decision function always converge to a fixed point. Due to the nonlinearity of the tentative decision function the exact calculation of the fixed point is intractable. It should also be noted that the results in this section do not apply to the PIC detector with linear decision function nor hard decision function as they do not satisfy the condition of monotonically increasing at a sublinear rate.

3.5.1 Iterated-Map Neural Networks
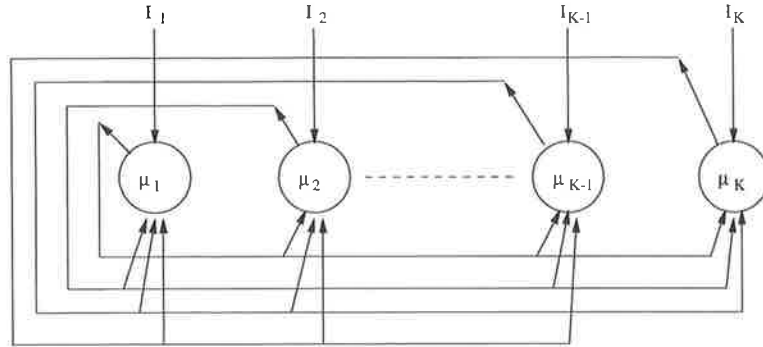
Figure 3.3: An iterated-map neural network

Consider an iterated-map neural network with K neurons as depicted in Fig. 3.3. The state of neuron k at time n is denoted by $\mu_k^{(n)}$. The input to each neuron k is the sum of the external input I_k and the weighted sum of the states of other neurons at time $n-1$. For the iterated-map neural network where the state of the neurons are updated in parallel, the state of neuron k is given by

$$\mu_k^{(n+1)} = f_k \left(I_k + \sum_{j \neq k} T_{j,k} \mu_j^{(n)} \right) , \qquad (3.72)$$

where $T_{j,k} = T_{k,j}$ is the weight on the interconnection which couples neuron j and neuron k, and $f_k(x)$ is a nonlinear input-output transfer function.

The stability of this iterated-map neural network is analysed in [66]. We now summarise the main results of [66] in the following theorems. The proofs of these results are omitted here and readers are referred to [66] for the detailed proofs.

Theorem 3.3 *Consider the iterated-map neural network as illustrated in Fig. 3.3 with the states of the neurons updated in parallel as described by (3.72). Provided that the input-output transfer function satisfies the conditions*

1. $f_k(x)$ is monotonically increasing for all k, and

2. $|f_k(x)| \le \zeta_k |x|^{\tau_k}$ *for some constant* $0 \le \tau_k < 1$ *and* $\zeta_k < \infty$, *and all* $x \in \mathbb{R}$

then the iterated-map network will either converge to a fixed point or enter a period-two limit cycle.

Theorem 3.4 *Let denote the maximum slope of the input-output transfer function* $f_k(x)$, *which satisfies conditions 1 and 2 in Theorem 3.3, by* β_k *i.e.*,

$$\beta_k = \sup_x \frac{df_k(x)}{dx} . \tag{3.73}$$

Define a diagonal matrix $\mathbf{B} = diag\{\beta_1, \beta_2, \cdots, \beta_K\}$ *and a symmetric matrix* $\mathbf{T} = [T_{j,k}]$ *with* $T_{k,k} = 0 \; \forall k$. *Under the condition that the matrix* $(\mathbf{T} + \mathbf{B}^{-1})$ *is positive definite, the iterated-map neural network will always converge to a fixed point.*

3.5.2 Application to the PIC Detector

The soft estimate of the information bit of user k in (3.7) can be rewritten as

$$\tilde{d}_k^{(n)} = g_k \left(\mathbf{s}_k^T \mathbf{r} - \sum_{j \neq k} A_j \, \rho_{j,k} \, \tilde{d}_j^{(n-1)} \right) , \tag{3.74}$$

$$= \tilde{g}_k \left(A_k \, \mathbf{s}_k^T \mathbf{r} - \sum_{j \neq k} A_k \, A_j \, \rho_{j,k} \, \tilde{d}_j^{(n-1)} \right) , \tag{3.75}$$

where $\tilde{g}_k(x) = g_k(\frac{x}{A_k})$. Thus, we have arranged the PIC detector of the first form (3.6) into the same form as that of the iterated-map neural network. Each active user in the CDMA system can be viewed as a neuron in the neural network with $\tilde{d}_k^{(n)}$ represents the state of the neuron k, $A_k \mathbf{s}_k^T \mathbf{r}$ represents the external input to neuron k, and $-A_k A_j \rho_{j,k}$ is the strength of the connection from neuron j to neuron k. Similarly, the PIC detector of the second form can also be rewritten into the form similar to the iterated-map neural network by substituting (3.8) into (3.9). Thus we have

$$\tilde{b}_k^{(n)} = h_k \left(\mathbf{s}_k^T \mathbf{r} - \sum_{j \neq k} \rho_{j,k} \, \tilde{b}_j^{(n-1)} \right) . \tag{3.76}$$

Therefore by using Theorem 3.3 and Theorem 3.4, we obtain the following corollary and propositions which state the convergence behaviour of the PIC detector.

Corollary 3.1 *The PIC detector of the form (3.6) or (3.8) with an arbitrary tentative decision function that is monotonically increasing at a sublinear rate will either converge to a fixed point or enter a limit cycle of period-two.*

Proposition 3.5 (First Form) *Denote the largest positive eigenvalues of the symmetric matrix $(\mathbf{R} - \mathbf{I})\mathbf{A}^2$ by $\lambda_{max}((\mathbf{R} - \mathbf{I})\mathbf{A}^2)$ and the maximum slope of the tentative decision function $g_k(x)$ by ψ_k^g i.e.*

$$\psi_k^g = \sup_x \frac{d\,g_k(x)}{dx} \ . \tag{3.77}$$

Under the condition that

$$\psi_k^g < \frac{A_k}{\lambda_{max}\left((\mathbf{R} - \mathbf{I})\mathbf{A}^2\right)} \quad \forall k, \tag{3.78}$$

the PIC detector of the first form (3.6) with any tentative decision function that is monotonically increasing at a sublinear rate will always converge to a fixed point where the soft estimates $\tilde{\mathbf{d}}^{(n)} \to \mathbf{d}^*$ as $n \to \infty$.

Proof: Since $\tilde{g}_k(x) = g_k(\frac{x}{A_k})$, the maximum slope of $\tilde{g}_k(x)$ is $\frac{\psi_k^g}{A_k}$. As shown previously in (3.75), the PIC detector of the first form (3.6) is equivalent to the iterated-map neural network (3.72). Hence, by Theorem 3.4 this detector will converge to a fixed point if the matrix

$$\mathbf{B}_g^{-1} - \mathbf{A}\,(\mathbf{R} - \mathbf{I})\,\mathbf{A} \tag{3.79}$$

is positive definite where $\mathbf{B}_g = \text{diag}\{\frac{\psi_1^g}{A_1}, \frac{\psi_2^g}{A_2}, \cdots, \frac{\psi_K^g}{A_K}\}$.

A sufficient condition for the matrix $[\mathbf{B}_g^{-1} - \mathbf{A}\,(\mathbf{R} - \mathbf{I})\,\mathbf{A}]$ to be positive definite is

$$\psi_k^g < \frac{A_k}{\lambda_{max}(\mathbf{A}\,(\mathbf{R} - \mathbf{I})\,\mathbf{A})} \quad \forall k. \tag{3.80}$$

The result immediately follows as the eigenvalues of $\mathbf{A}\,(\mathbf{R} - \mathbf{I})\,\mathbf{A}$ are the same as the eigenvalues of $(\mathbf{R} - \mathbf{I})\,\mathbf{A}^2$. $\triangle$

Proposition 3.6 (Second Form) *Denote the largest positive eigenvalues of the symmetric matrix* $(\mathbf{R} - \mathbf{I})$ *by* $\lambda_{max}(\mathbf{R} - \mathbf{I})$ *and the maximum slope of the tentative decision function* $h_k(x)$ *by* ψ_k^h *i.e.,*

$$\psi_k^h = \sup_x \frac{d\,h_k(x)}{dx} \; . \tag{3.81}$$

Under the condition that

$$\psi_k^h < \frac{1}{\lambda_{max}(\mathbf{R} - \mathbf{I})} \quad \forall k, \tag{3.82}$$

the PIC detector of the second form (3.8) with any tentative decision function that is monotonically increasing at a sublinear rate will always converge to a fixed point where the soft estimates $\tilde{\mathbf{b}}^{(n)} \to \mathbf{b}^*$ *as* $n \to \infty$.

Proof: This proposition is proved in a similar way as for Proposition 3.5. By rewriting the PIC detector of the second form (3.9) into (3.76) and recognizing its connection with the iterated-map neural network, we can then apply Theorem 3.4. Thus, the detector will converge if the matrix

$$\mathbf{B}_h^{-1} - (\mathbf{R} - \mathbf{I}) \tag{3.83}$$

is positive definite where $\mathbf{B}_h = \text{diag}\{\psi_1^h, \; \psi_2^h, \; \cdots, \; \psi_K^h\}$. The result immediately follows which concludes the proof. $\triangle$

Similar to the results in the previous section, the sufficient conditions for convergence derived using this approach also demonstrate that by placing a bound on the maximum slope of the tentative decision function, the PIC detector will converge to a fixed point. Moreover, this bound is depended on the correlation between the spreading codes and the signal amplitudes of the users in the system. We now compare the bounds derived here with those derived in the previous section to determine which approach would give a weaker (less restrictive) sufficient condition for convergence.

For the PIC detector of the first form, the upper bounds on the largest allowable slope of the tentative decision function are

$$\psi_{ub1}^{fdback} = \frac{1}{\|(\mathbf{R} - \mathbf{I})\mathbf{A}\|_{\infty}} \qquad (3.84)$$

as given by the feedback approach and

$$\psi_{ub1,k}^{nnet} = \frac{A_k}{\lambda_{max}((\mathbf{R} - \mathbf{I})\mathbf{A}^2)} \qquad (3.85)$$

as given by the iterated-map neural network approach. We cannot conclude which approach would give a weaker sufficient condition for convergence for this general case because the second approach gives a different bound for each user. However, for the special case when all users have equal signal amplitude with $A_k = a \ \forall k$, then we have

$$\psi_{ub1}^{fdback} = \frac{1}{a \, \|(\mathbf{R} - \mathbf{I})\|_{\infty}} \ , \qquad (3.86)$$

$$\psi_{ub1,k}^{nnet} = \frac{1}{a \, \lambda_{max}(\mathbf{R} - \mathbf{I})} \quad \forall k. \qquad (3.87)$$

Since

$$\lambda_{max}(\mathbf{R} - \mathbf{I}) \leq \rho(\mathbf{R} - \mathbf{I}) \ , \qquad (3.88)$$

$$\leq \|\mathbf{R} - \mathbf{I}\|_{\infty} \ , \qquad (3.89)$$

the iterated-map neural network approach gives a weaker sufficient condition for convergence for this special case.

For the PIC detector of the second form, the upper bounds on the largest allowable slope of the tentative decision function are

$$\psi_{ub2}^{fdback} = \frac{1}{\|\mathbf{R} - \mathbf{I}\|_{\infty}} \ , \qquad (3.90)$$

for the feedback approach and

$$\psi_{ub2}^{nnet} = \frac{1}{\lambda_{max}(\mathbf{R} - \mathbf{I})} \ , \qquad (3.91)$$

for the iterated-map neural network approach. For this second PIC detector form, $\psi_{ub2}^{fdback} \leq \psi_{ub2}^{nnet}$ and therefore we can conclude that the iterated-map neural network approach gives a weaker sufficient condition for convergence.

3.5.3 Examples

Hyperbolic Tangent Decision Function

Let's consider the PIC detector with hyperbolic tangent decision function, $\tanh(\alpha_k x)$ where α_k are some constants. This PIC detector can be expressed by the first form (3.6) with $g_k(x) = \tanh(\alpha_k x)$. As previously shown in section 3.4, the maximum slope of the function $\tanh(\alpha_k x)$ is

$$\psi_k^g = \sup_x \frac{d\,g_k(x)}{d\,x} = \alpha_k \ . \tag{3.92}$$

Using Proposition 3.5, the sufficient condition for convergence to a fixed point is therefore

$$\alpha_k < \frac{A_k}{\lambda_{max}((\mathbf{R}-\mathbf{I})\mathbf{A}^2)} \quad \forall k. \tag{3.93}$$

Clip Decision Function

The PIC detector with the clip tentative decision function can be expressed into the second form (3.8) by having $h_k(x)$ defined as in (3.10). The maximum slope of the clip function is

$$\psi_k^h = 1. \tag{3.94}$$

Thus, using Proposition 3.6, the sufficient condition for convergence is

$$\lambda_{max}(\mathbf{R}-\mathbf{I}) < 1 \ . \tag{3.95}$$

This sufficient condition for convergence turns out to be exactly the same as that derived in [108].

3.6 Simulation

We use computer simulation to examine the convergence behaviour of the PIC detector with hyperbolic tangent decision function in this section. We consider a CDMA system with 4 active users. All four users have equal power with $A_k = 1$. Each user is assigned a spreading code sequence of length $M = 5$ where the correlation matrix is

$$\mathbf{R} = \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 \\ 0.6 & 1 & 0.2 & 0.2 \\ 0.6 & 0.2 & 1 & 0.2 \\ 0.6 & 0.2 & 0.2 & 1 \end{bmatrix}. \tag{3.96}$$

With these system parameters, the corresponding sufficient condition for convergence derived using the feedback approach is

$$\max_k \alpha_k < \alpha_{max}^{fdback} = \frac{1}{\|\mathbf{R} - \mathbf{I}\|_\infty} \tag{3.97}$$

and the iterated-map neural network approach is

$$\alpha_k < \alpha_{max}^{nnet} = \frac{1}{\lambda_{max}(\mathbf{R} - \mathbf{I})} \quad \forall k, \tag{3.98}$$

where $\|\mathbf{R} - \mathbf{I}\|_\infty = 1.8$ and $\lambda_{max}(\mathbf{R} - \mathbf{I}) = 1.2583$. As mentioned previously, the sufficient condition derived using the latter approach is weaker than the former.

In order to find the largest allowable α_k for which the PIC detector with hyperbolic tangent decision function still converge to a fixed point via simulation, we ran a series of Monte Carlo simulations for different values of α_k. The following results are obtained at the signal to noise ratio SNR = 0dB, where the SNR is defined as A_k^2/σ^2. Fig. 3.4 shows the probability that the detector enters the period-two limit cycle for different values of α_k. The result is obtained by averaging the number of times the detector enters a limit cycle over $250,000$ different realisations of the received signal vector $\mathbf{r}$. As shown in Fig. 3.4, the detector converges to a fixed
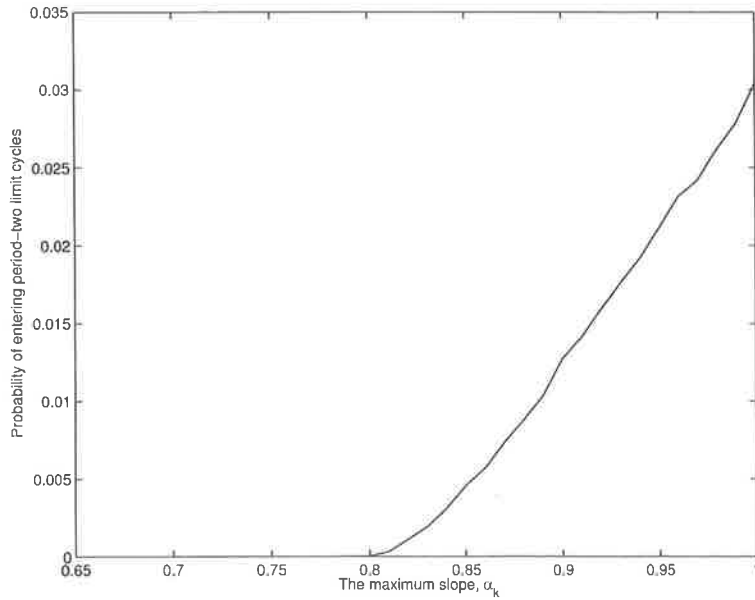
Figure 3.4: Percentage of entering the period-two limit cycle for different α_k

point for all $\alpha_k < 0.8$ and starts to exhibit cyclic behaviour of period two when the value of α_k is greater than 0.8. This suggests that the upper bound on the largest allowable α_k for which the PIC detector still converges to a fixed point is $\alpha_{max}^{sim} = 0.8$. This bound is very close to the analytical bound $\alpha_{max}^{nnet} = 0.795$ derived by using the iterated-map neural network approach. Simulation result also indicates that as the maximum slope of the hyperbolic tangent decision function increases, the probability that the detector enters a limit cycle also increases.

Fig. 3.5 shows the soft estimate of individual user symbol at the end of each interference cancellation stage, $\tilde{d}_k^{(n)}(i)$, when $\alpha_k = 0.5\ \forall k$. With this value, the soft estimates of the transmitted symbol of all users converge to a fixed point after several interference cancellation stages as expected. We also find that for α_k with values that is slightly smaller than α_{max}^{nnet}, it may require more than 100 interference cancellation stages for the estimates of the user symbols to converge to a fixed value as being shown in Fig. 3.6.
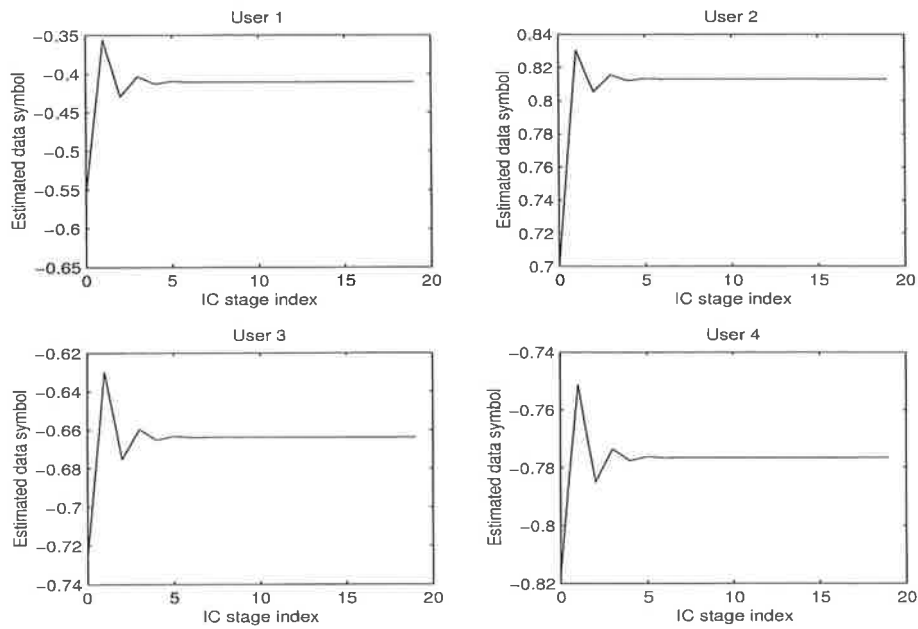
Figure 3.5: Convergence to a fixed point, $\alpha_k = 0.5$
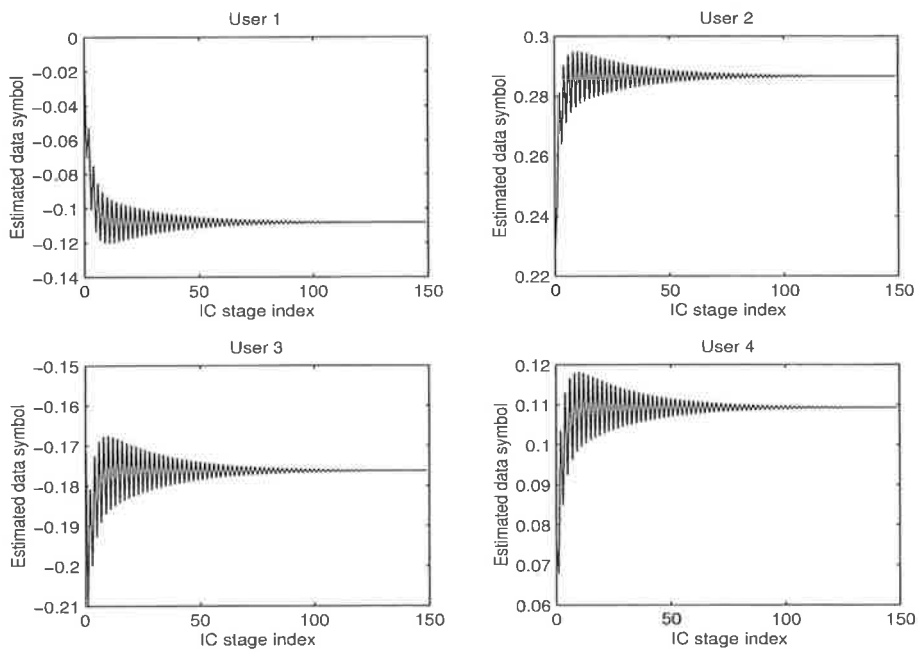
Figure 3.6: Convergence to a fixed point, $\alpha_k = 0.79$
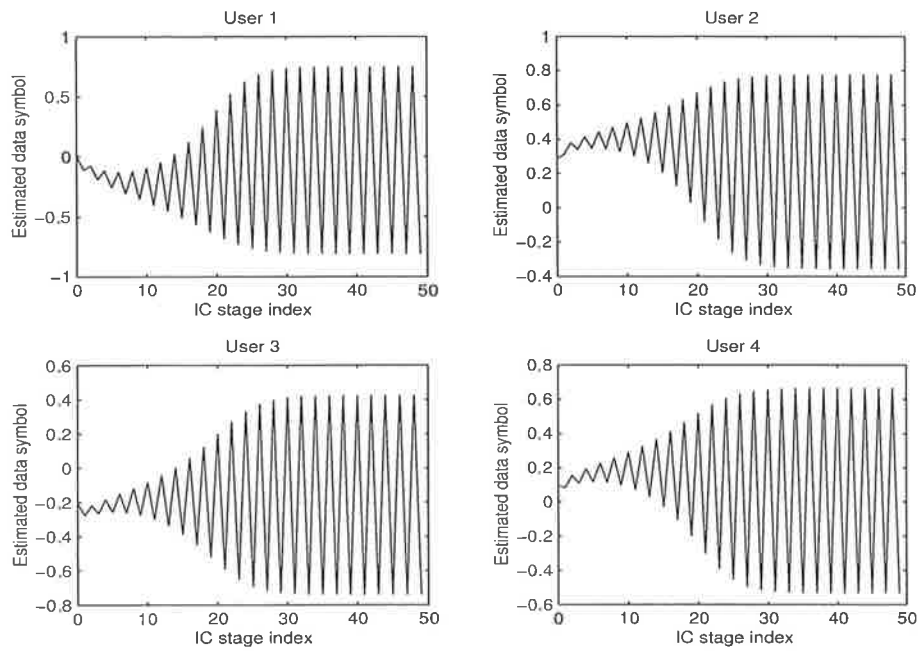
Figure 3.7: Convergence problem - period-two limit cycle: $\alpha_k = 1$
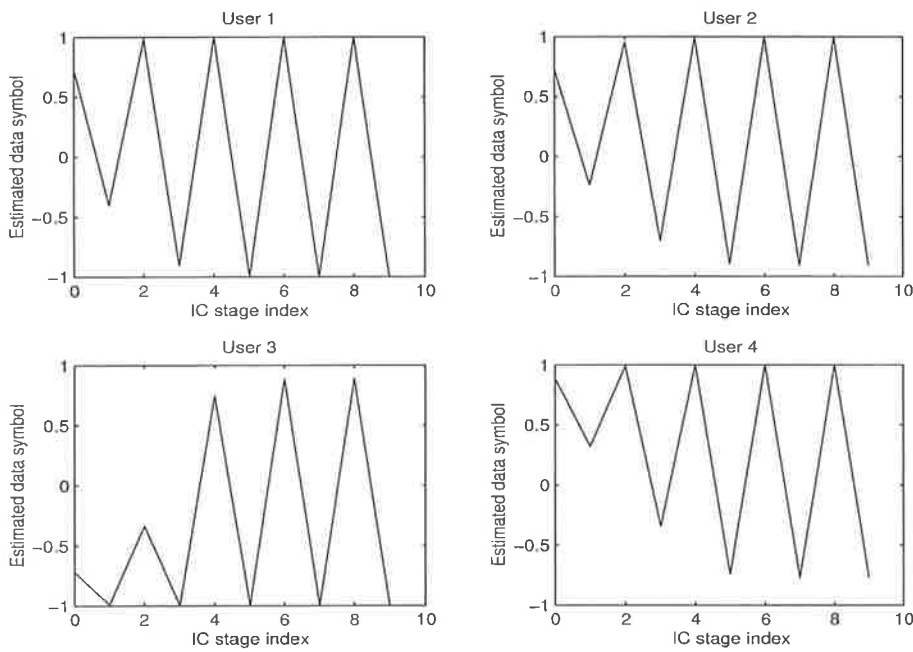
Figure 3.8: Convergence problem - period-two limit cycle: $\alpha_k = 2.5$

The simulation results in Fig. 3.7 and Fig. 3.8 are obtained for the cases when $\alpha_k = 1 \; \forall k$ and $\alpha_k = 2.5 \; \forall k$, respectively. They illustrate the behaviour of the PIC detector when the tentative decision function does not satisfy the sufficient condition for convergence. As one can observe, the detector exhibits a cyclic behaviour of period two for both cases. However, for the case $\alpha_k = 1$, it requires more number of interference cancellation stages before it enters the limit cycles and also the estimated values of the data symbols are not closed to ± 1 as compared to the case $\alpha_k = 2.5$. These observations can be explained by the fact that as α_k becomes large, the hyperbolic tangent decision function approaches the hard decision.

3.7 Conclusions

In this chapter, by introducing a general PIC detector model, we have developed a general framework for analysing the convergence behaviour of the PIC detector. This framework permits the derivation of the sufficient condition for convergence of the PIC detector for a wide range of tentative decision functions. We have shown that some well-known conditions for convergence of the PIC detector with linear decision and clip decision functions can be obtained using this framework. We analysed the convergence behaviour of the PIC detector by using known results from the stability analysis of feedback systems and an iterated-map neural network. Both approaches lead to the same conclusion that that by placing a bound on the maximum slope of the tentative decision function, the soft estimates of the information bits produced by the PIC detectors approach a fixed point as $n \to \infty$. Analytical results derived from the second approach also shows that the PIC detector with any nonlinear tentative decision function that is monotonically increasing at sublinear rate will either converge to a fixed point or enter a limit cycle of period-two. Computer simulation was also used to examine the convergence behaviour of the PIC detector with hyperbolic tangent decision function. The simulation and analytical results

obtained were found to be in close agreement.

3.8   Appendix

Lemma 3.1 *Let* $\mathbf{H} \in \mathbb{R}^{K \times K}$ *and* $\mathbf{x} \in \ell_{pe}^{K}$ *where* ℓ_{pe}^{K} *is the extended space of sequence in* $\mathbb{R}^{K}$ *as defined by (3.15). A value of* γ *which satisfy the condition*

$$\|(\mathbf{Hx})_N\|_p \leq \gamma \|\mathbf{x}_N\|_p \quad \forall N \in \mathbb{Z}_+ \tag{3.99}$$

is

$$\gamma = \|\mathbf{H}\|_p \;, \tag{3.100}$$

where $p \in [1, \infty]$ *and* $\|\mathbf{H}\|_p$ *denotes the operator norm of the matrix* $\mathbf{H}$.

Proof: Note that the value γ is not uniquely defined by (3.99). However, we are interested only in the value of γ that can be calculated from the matrix $\mathbf{H}$. From the definition of the operator norm, we have

$$\|\mathbf{H}\|_p = \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{Hx}\|_p}{\|\mathbf{x}\|_p} \;. \tag{3.101}$$

Hence by choosing

$$\gamma = \|\mathbf{H}\|_p \;, \tag{3.102}$$

the condition in (3.99) will be satisfied. $\triangle$

Lemma 3.2 *Let* $\mathbf{x} \in \ell_{\infty e}^{K}$ *where* $\ell_{\infty e}^{K}$ *is the extended space of sequence in* $\mathbb{R}^{K}$. *Let* $\mathbf{Hx} = \mathbf{f}(\mathbf{x})$ *where* $\mathbf{f}(.) = [f_1(.) \; f_2(.) \; \cdots, \; f_K(.)]^T$ *acts component. Let's define*

$$\varphi_k = \sup_{x \neq 0} \frac{|f_k(x)|}{|x|} \quad \forall x \in \mathbb{R} \;. \tag{3.103}$$

A value of γ *which satisfy the condition*

$$\|(\mathbf{Hx})_N\|_\infty \leq \gamma \|\mathbf{x}_N\|_\infty \quad \forall N \in \mathbb{Z}_+ \tag{3.104}$$

is

$$\gamma = \max_k \varphi_k \;. \tag{3.105}$$

Proof: Now

$$\|(\mathbf{Hx})_N\|_\infty = \|\mathbf{f(x)}_N\|_\infty \tag{3.106}$$

$$= \max_{1 \le n \le N} |\mathbf{f}(\mathbf{x}^{(n)})|_\infty \tag{3.107}$$

$$= \max_{1 \le n \le N} \max_k |f_k(x_k^{(n)})| , \tag{3.108}$$

where $|.|_\infty$ denotes ℓ_∞-norm on $\mathbb{R}^K$ and $|.|$ denotes the absolute value of a number in $\mathbb{R}$. Since

$$|f_k(x)| \le \varphi_k |x| \quad \forall x \in \mathbb{R} , \tag{3.109}$$

we then have

$$\|(\mathbf{Hx})_N\|_\infty \le \max_{1 \le n \le N} \max_k \varphi_k |x_k^{(n)}| \tag{3.110}$$

$$\le \max_j \varphi_j \max_{1 \le n \le N} \max_k |x_k^{(n)}| \tag{3.111}$$

$$= \max_j \varphi_j \|\mathbf{x}_N\|_\infty . \tag{3.112}$$

Thus, by choosing $\gamma = \max_k \varphi_k$ the condition in (3.104) will be satisfied. $\triangle$

Lemma 3.3 *Let* $\mathbf{H} \in \mathbb{R}^{K \times K}$ *and* $\mathbf{x}, \mathbf{x}' \in \ell_{pe}^K$ *where* ℓ_{pe}^K *is the extended space of sequence in* $\mathbb{R}^K$. *A value of* $\tilde{\gamma}$ *which satisfy the condition*

$$\|(\mathbf{Hx})_N - (\mathbf{Hx}')_N\|_p \le \gamma \|\mathbf{x}_N - \mathbf{x}'_N\|_p \quad \forall N \in \mathbb{Z}_+ \tag{3.113}$$

is

$$\gamma = \|\mathbf{H}\|_p . \tag{3.114}$$

where $p \in [1, \infty]$ *and* $\|\mathbf{H}\|_p$ *denotes the operator norm the matrix* $\mathbf{H}$.

Proof: Since $\mathbf{H}$ is a linear mapping,

$$\|(\mathbf{Hx})_N - (\mathbf{Hx}')_N\|_p = \|\mathbf{H}(\mathbf{x} - \mathbf{x}')_N\|_p . \tag{3.115}$$

Now, if $\mathbf{x}, \mathbf{x}' \in \ell_{pe}^K$, then $(\mathbf{x} - \mathbf{x}') \in \ell_{pe}^K$. Thus by Lemma 3.1 the results immediately follows. $\triangle$

Lemma 3.4 *Let* $\mathbf{x}, \mathbf{x}' \in \ell_{\infty e}^K$ *where* $\ell_{\infty e}^K$ *is the extended space of sequence in* $\mathbb{R}^K$. *Let* $\mathbf{Hx} = \mathbf{f}(\mathbf{x})$ *where* $\mathbf{f}(.) = [f_1(.) \ f_2(.) \ \cdots, \ f_K(.)]^T$ *acts component. Let's define*

$$\varphi_k = \sup_{x \neq 0} \frac{|f_k(x)|}{|x|} \ , \tag{3.116}$$

$$\psi_k = \sup_{x} \frac{df_k(x)}{dx} \ . \tag{3.117}$$

Provided that $|\psi_k| \leq \varphi_k \ \ \forall k$, *a value of* $\tilde{\gamma}$ *which satisfy the condition*

$$\|(\mathbf{Hx})_N - (\mathbf{Hx}')_N\|_\infty \leq \tilde{\gamma}\|\mathbf{x}_N - \mathbf{x}'_N\|_\infty \ \ \forall N \in \mathbb{Z}_+ \tag{3.118}$$

is

$$\tilde{\gamma} = \max_k \varphi_k \ . \tag{3.119}$$

Proof: Now

$$\|(\mathbf{Hx})_N - (\mathbf{Hx}')_N\|_\infty = \|\mathbf{f}(\mathbf{x})_N - \mathbf{f}(\mathbf{x}')_N\|_\infty \tag{3.120}$$

$$= \max_{1 \leq n \leq N} |\mathbf{f}(\mathbf{x}^{(n)}) - \mathbf{f}(\mathbf{x}'^{(n)})|_\infty \tag{3.121}$$

$$= \max_{1 \leq n \leq N} \max_k |f_k(x_k^{(n)}) - f_k(x'^{(n)}_k)| \ , \tag{3.122}$$

where $|.|_\infty$ denotes ℓ_∞-norm on $\mathbb{R}^K$ and $|.|$ denotes the absolute value of a number in $\mathbb{R}$. Now, if $|\psi_k| \leq \varphi_k$, we then have

$$|f_k(x) - f_k(x')| \leq \varphi_k |x - x'| \ \ \forall x, \ x' \in \mathbb{R} \ . \tag{3.123}$$

Hence, if $|\psi_k| \leq \varphi_k$

$$\|(\mathbf{Hx})_N - (\mathbf{Hx}')_N\|_\infty \leq \max_{1 \leq n \leq N} \max_k \varphi_k |x_k^{(n)} - x'^{(n)}_k| \tag{3.124}$$

$$\leq \max_j \varphi_j \max_{1 \leq n \leq N} \max_k |x_k^{(n)} - x'^{(n)}_k| \tag{3.125}$$

$$= \max_j \varphi_j \|\mathbf{x}_N - \mathbf{x}'_N\|_\infty \ . \tag{3.126}$$

Thus, if $|\psi_k| \leq \varphi_k$, then by choosing $\tilde{\gamma} = \max_k \varphi_k$ the condition in (3.118) will be satisfied. $\triangle$

Chapter 4

Space-Time Coding

This chapter addresses the issues of designing space-time codes in spatially correlated fading channels. Analytical upper bounds and asymptotic tight expression for the pairwise error probability of space-time coded systems in spatially correlated fading channels are derived. Based on the performance analysis, the chapter presents the rank determinant and the trace design criteria for constructing space-time codes to operate in spatially correlated flat Rayleigh fading channels. Furthermore, some new space-time trellis codes based on a new design criterion, which involves minimising the sum of the pairwise error probability of all distinct pairs of codewords, are also presented. These new space-time trellis codes are found through systematic computer search. Simulation results support the claim that these new codes are superior to other known codes constructed using the rank determinant and the trace criteria.

4.1 Introduction

Wireless communication systems operating in flat fading environments can increase their capacity significantly by employing multiple transmit and multiple receive an-

tennas [36, 93]. In [35], a layered space-time architecture for wireless communication in flat Rayleigh fading environment with multiple antennas is designed. It demonstrates that when the fading coefficients are statistically independent and known at the receiver, the capacity of the link increases linearly with the smaller of the number of transmit and receive antennas. Following these works, Tarokh, Seshadri and Calderbank [89] propose a space-time trellis code modulation scheme to exploit this potential increase in capacity promised by multiple transmit and multiple receive antennas. It is found in [41, 89] that the pairwise error probability of the space-time codes in quasi-static Rayleigh fading channels is determined by the rank and the determinant of the distance matrix. Using this analysis, a rank determinant design criterion is proposed which involves maximising the minimum rank and the minimum determinant of the distance matrix over all distinct pairs of codewords. Based on this criterion, a number of QPSK and 8PSK codes are constructed by hand in [89]. These codes achieve the maximal possible diversity gain, but not the full coding gain. Subsequently, some other QPSK codes with better performance are found by computer search in [12]. Recently another design criterion for constructing space-time trellis codes in quasi-static Rayleigh fading channels is derived in [21, 110]. It is found that when the product of the minimum rank r of the distance matrix and the number of receive antennas is greater or equal to 4, the pairwise error probability is determined by the squared Euclidean distance between two codewords (i.e. trace of the distance matrix). Thus, the code design criterion is then to maximise the minimum squared Euclidean distance. It has been shown by simulation that a code with a smaller determinant can still achieve better performance than the one with a larger determinant, provided that its minimum squared Euclidean distance is larger [110].

The rank determinant and the trace design criteria in [21, 89, 110] are derived based on some upper bounds of the pairwise error probability of the space-time

coded systems in independent fading channels. Such spatially independent fading only occur if the multipath reflections are uniformly distributed around the receiver and that the multiple antennas at the receiver are sufficiently spaced apart. This assumption is often hard to satisfy in practice. At the mobile unit (MU) end, even thought it is frequently immersed in a complex scattering environment where the received signals are linear combination of several multipaths reflected from the nearby local scatterers, the antennas at the MU cannot be sufficiently spaced apart due to the limited size of the MU. In the isotropic scattering environments where the MU receives signals from all directions with equal probabilities, the received signals at two MU antennas can be assumed to be spatially uncorrelated if the two antennas are separated at a distance greater than $\lambda/2$ where λ is the carrier wavelength. However, it has been experimentally demonstrated in [37] that the scattering at the MU is more likely nonisotropic, resulting in a nonuniform distribution of the angle of arrival (AOA) of the multipath components at the MU. This nonuniform distribution of AOA significantly increases the cross-correlation among the antennas at the MU. In addition, the base station (BS) antennas in land radio systems are typically placed highly above the ground and are not surrounded by many local scatterers. As a result, spatial correlation also arises among the antennas at the BS.

In this chapter, we derive new upper bounds for the pairwise error probability of the space-time coded systems in spatially correlated fading channels. Based on these analytical pairwise error probability bounds, we derive the equivalent rank determinant and trace design criteria for constructing space-time codes in correlated fading channels. The rank determinant and the trace design criteria have a common theme in which they minimize only the probability of error of the dominant error event. In particular, for each of these design criteria there exists many space-time trellis codes and they don't necessary have the same performance. For example, it is found in [107] that there are over 250 different 4 states QPSK space-time trellis codes

for 2 transmit antennas that would satisfy the rank determinant design criterion. These observations motivate us to construct new space-time trellis codes based on the criterion of minimising the sum of the pairwise error probability of all possible error events. Since all codewords are likely to be transmitted with equal probability, it is expected that the sum of the pairwise error probability of all possible error events is a better measure of the error performance than the pairwise error probability of the dominant error event. Through computer search, it is found that only a limited number of space-time trellis codes exist that would satisfy this criterion. This sum of pairwise error probability criterion is similar to the distance spectrum criterion reported in [6]. There are some codes constructed based on this distance spectrum criterion have been presented in [54]. However they all use BPSK modulation and achieve a bandwidth efficiency of 1 bit/s/Hz. The space-time trellis codes presented in this chapter all use QPSK modulation and achieve a bandwidth efficiency of 2 bits/s/Hz.

This chapter is organised as follows: In section 4.2, a general space-time coded system with multiple transmit and multiple receive antennas is presented. The performance of the space-time system in spatially correlated Rayleigh fading channels is analysed in section 4.3. Based on the analytical expressions of the pairwise error probability, codes design criteria are derived in section 4.4. In section 4.5, the construction of space-time trellis codes is described. Computer simulations are used to compare the performance of the newly constructed codes with other known codes. The following notation is used throughout the chapter: The superscript ^{T} and ^{H} denote the transpose and conjugate transpose operations, respectively. The Kronecker product operation is denoted by $\otimes$. The zero-mean, circular-symmetric, complex Gaussian distribution with variance of σ^2 is denoted by $\mathcal{CN}(0, \sigma^2)$. The matrix $\mathbf{I}_N$ ($\mathbf{0}_N$) is the $N \times N$ identity matrix (matrix of all zeros).
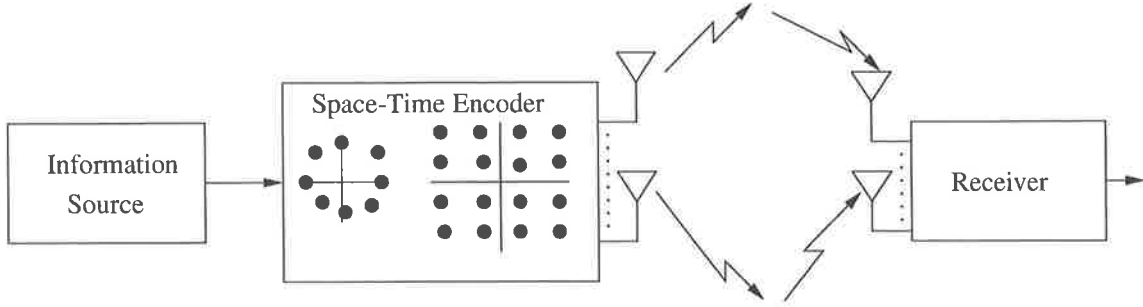
Figure 4.1: The block diagram of the transmitter and receiver

4.2   The System Model

Consider a wireless communication system with n_T transmit and n_R receive antennas as depicted in Fig. 4.1. At each time t, the encoder produces n_T outputs $x_1(t), x_2(t), ..., x_{n_T}(t)$ where $x_i(t)$ is a signal from a certain signal constellation with unit average energy. These outputs are then simultaneously transmitted by n_T antennas. At the receiver, the received signal at the jth receive antennas is

$$r_j(t) = \sum_{i=1}^{n_T} h_{i,j}\ x_i(t)\ \sqrt{E_s} + n_j(t)\ , \tag{4.1}$$

where E_s is the transmitted power per symbol at each transmit antenna and $n_j(t)$ is the additive noise component at the receive antenna j. The additive noise is white and $\mathcal{CN}(0,1)$ distributed. The coefficient $h_{i,j}$ is the fading gain for the path from transmit antenna i to receive antenna j and is assumed to be $\mathcal{CN}(0,[\Sigma]_{g,g})$ distributed where $\mathbf{\Sigma}$ is the channel correlation matrix defined later in this section and the index $g = (j-1)n_T + i$. We assume the channels undergo quasi-static fading (i.e. the fading coefficients remain constant during one frame of T symbol periods and change independently from one frame to another). Thus, the received signals at the receive antenna j during T symbol periods can be written in vector form as:

$$\mathbf{r}_j = \sqrt{E_s}\ \mathbf{X}\ \mathbf{h}_j + \mathbf{n}_j\ , \tag{4.2}$$

where $\mathbf{h}_j = [h_{1,j} \ h_{2,j} \ \cdots \ h_{n_T,j}]^T$, $\mathbf{n}_j = [n_j(1) \ n_j(2) \ \cdots \ n_j(T)]^T$ and

$$\mathbf{X} = \begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_{n_T}(1) \\ x_1(2) & x_2(2) & \cdots & x_{n_T}(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(T) & x_2(T) & \cdots & x_{n_T}(T) \end{bmatrix} . \qquad (4.3)$$

Stacking the T observations per receive antenna yields the Tn_R sufficient statistics

$$\mathbf{r} = [\mathbf{r}_1^T \ \mathbf{r}_2^T \ \cdots \ \mathbf{r}_{n_R}^T]^T = \sqrt{E_s} \, \mathcal{X} \, \mathbf{h} + \mathbf{n} \,, \qquad (4.4)$$

where $\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \cdots \ \mathbf{h}_{n_R}^T]^T$, $\mathbf{n} = [\mathbf{n}_1^T \ \mathbf{n}_2^T \ \cdots \ \mathbf{n}_{n_R}^T]^T$ and $\mathcal{X} = \mathbf{I}_{n_R} \otimes \mathbf{X}$. The additive noise is assumed to be white in space and time (i.e. $E[\mathbf{n}\mathbf{n}^H] = \mathbf{I}_{Tn_R}$). However, we allow the fading process to be spatially correlated with covariance $E[\mathbf{h}\mathbf{h}^H] = \mathbf{\Sigma}$ (i.e. $\mathbf{\Sigma}$ need not be proportional to the identity matrix).

4.3 Performance Analysis

In this section, we first review the derivation of the pairwise error probability presented in [14], which is asymptotically tight to the true pairwise error probability. We then derive two upper bounds for the pairwise error probability of the space-time coded systems in correlated fading channels. These upper bounds are more accurate than the asymptotic expression [14] at low SNRs. The derivations of these upper bounds follow closely to those presented in [89, 110]. However, the difference is that the upper bounds derived in [89, 110] are for independent fading channels while the upper bounds derived here are for spatially correlated fading channels.

Assuming ideal channel state information (CSI) is available at the receiver and a maximum likelihood receiver is employed, the receiver will erroneously select the codeword $\hat{\mathbf{X}}$ when the codeword $\mathbf{X}$ is transmitted if

$$\sum_{t=1}^{T} \sum_{j=1}^{n_R} \left| r_j(t) - \sum_{i=1}^{n_T} h_{i,j} \hat{x}_i(t) \sqrt{E_s} \right|^2 < \sum_{t=1}^{T} \sum_{j=1}^{n_R} \left| r_j(t) - \sum_{i=1}^{n_T} h_{i,j} x_i(t) \sqrt{E_s} \right|^2 . \qquad (4.5)$$

This is equivalent to

$$||\mathbf{r} - \sqrt{E_s}\hat{\mathcal{X}}\mathbf{h}||^2 < ||\mathbf{r} - \sqrt{E_s}\mathcal{X}\mathbf{h}||^2 \;. \qquad (4.6)$$

4.3.1 Asymptotic Tight Pairwise Error Probability

As from (4.6), the pairwise error probability $P(\mathbf{X} \to \hat{\mathbf{X}})$ of mistaking $\hat{\mathbf{X}}$ for $\mathbf{X}$ is

$$P(\mathbf{X} \to \hat{\mathbf{X}}) = Pr(f < 0 \mid \mathbf{X} \text{ is transmitted }) \;, \qquad (4.7)$$

where $f = ||\mathbf{r} - \sqrt{E_s}\hat{\mathcal{X}}\mathbf{h}||^2 - ||\mathbf{r} - \sqrt{E_s}\mathcal{X}\mathbf{h}||^2$. By defining the vector

$$\mathbf{z} = \begin{bmatrix} \sqrt{E_s}\mathbf{h} \\ \mathbf{r} \end{bmatrix} \qquad (4.8)$$

and the Hermitian matrix

$$\mathbf{F} = \begin{bmatrix} \hat{\mathcal{X}}^H\hat{\mathcal{X}} - \mathcal{X}^H\mathcal{X} & \mathcal{X}^H - \hat{\mathcal{X}}^H \\ \mathcal{X} - \hat{\mathcal{X}} & \mathbf{0}_{Tn_R} \end{bmatrix} \;, \qquad (4.9)$$

the test statistic f can be written in the quadratic form as

$$f = \mathbf{z}^H\mathbf{F}\mathbf{z} \;. \qquad (4.10)$$

The characteristic function of the distribution of the test statistic f as derived in [83] is

$$G_f(s) = E[e^{-sf}] = \frac{1}{\prod_{n=1}^{N} \alpha_n^{\mu_n}(s + \alpha_n^{-1})^{\mu_n}} \;, \qquad (4.11)$$

where N is the number the distinct nonzero eigenvalues α_n of $\mathbf{R}_{\mathbf{zz}|\mathbf{X}}\mathbf{F}$ with multiplicity μ_n and $\mathbf{R}_{\mathbf{zz}|\mathbf{X}}$ is the covariance matrix of the random vector $\mathbf{z}$ given that the codeword $\mathbf{X}$ is transmitted

$$\mathbf{R}_{\mathbf{zz}|\mathbf{X}} = E\left[\mathbf{z}\mathbf{z}^H|\mathbf{X}\right] = \begin{bmatrix} E_s\Sigma & E_s\Sigma\mathcal{X}^H \\ E_s\mathcal{X}\Sigma & E_s\mathcal{X}\Sigma\mathcal{X}^H + \mathbf{I} \end{bmatrix} \;. \qquad (4.12)$$

Suppose $\mathbf{R_{zz|X}}\mathbf{F}$ has $\{\alpha_n\}_{n=1}^{L}$ negative eigenvalues and $\{\alpha_n\}_{n=L+1}^{N}$ positive eigenvalues, the exact pairwise error probability is given by [14]

$$
\begin{aligned}
P(\mathbf{X} \to \hat{\mathbf{X}}) &= \int_{-\infty}^{0} \int_{-j\infty+\epsilon}^{j\infty+\epsilon} e^{sf} G_f(s) \frac{ds\, df}{2\pi j} \ , \\
&= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} s^{-1} G_f(s) \frac{ds}{2\pi j} \ , \\
&= -\sum_{l=1}^{L} \mathrm{Res}\left(\frac{1}{s \prod_{n=1}^{N} \alpha_n^{\mu_n}\, (s + 1/\alpha_n)^{\mu_n}}, s_l = \frac{-1}{\alpha_l} \right) \ .
\end{aligned}
\tag{4.13}
$$

The residue of a function $f(s)$ at a pole s_i with multiplicity of μ can be calculated as [82]

$$
\mathrm{Res}(f(s), s_i) = \frac{1}{(\mu-1)!} \lim_{s \to s_i} \frac{d^{\mu-1}}{ds^{\mu-1}}[(s-s_i)^\mu f(s)] \ .
\tag{4.14}
$$

The expression of the exact pairwise error probability above does not give any insight into the structure of the codes. In this section, we are interested in the case when the $SNR \to \infty$ (or equivalently $E_s \to \infty$) as this leads to a closed form expression for the pairwise error probability which shows its dependency on the system parameters such as codewords employed, fading correlation and system diversity.

Proposition 4.1 *Assuming Σ is full rank, as $E_s \to \infty$ half of the nonzero asymptotic eigenvalues of $\mathbf{R_{zz|X}}\mathbf{F}$ are equal to minus unity with multiplicity of rn_R where r is the rank of the distance matrix $\mathbf{A_{X,\hat{X}}} = (\mathbf{X} - \hat{\mathbf{X}})^H(\mathbf{X} - \hat{\mathbf{X}})$. The other half are positive and equal to the eigenvalues of $E_s\Sigma(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})$.*

Proof: To see this, let us first begin by expanding the expression of $\mathbf{R_{zz|X}}\mathbf{F}$:

$$
\begin{aligned}
\mathbf{R_{zz|X}}\mathbf{F} &= \begin{bmatrix} E_s\Sigma & E_s\Sigma\mathcal{X}^H \\ E_s\mathcal{X}\Sigma & E_s\mathcal{X}\Sigma\mathcal{X}^H + \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\mathcal{X}}^H\hat{\mathcal{X}} - \mathcal{X}^H\mathcal{X} & \mathcal{X}^H - \hat{\mathcal{X}}^H \\ \mathcal{X} - \hat{\mathcal{X}} & \mathbf{0}_{Tn_R} \end{bmatrix} \\
&= \begin{bmatrix} E_s\Sigma(\hat{\mathcal{X}}^H\hat{\mathcal{X}} - \mathcal{X}^H\mathcal{X}) + E_s\Sigma\mathcal{X}^H(\mathcal{X} - \hat{\mathcal{X}}) & E_s\Sigma(\mathcal{X}^H - \hat{\mathcal{X}}^H) \\ E_s\mathcal{X}\Sigma(\hat{\mathcal{X}}^H\hat{\mathcal{X}} - \mathcal{X}^H\mathcal{X} + E_s\mathcal{X}\Sigma\mathcal{X}^H(\mathcal{X} - \hat{\mathcal{X}}) + \mathcal{X} - \hat{\mathcal{X}} & E_s\mathcal{X}\Sigma(\mathcal{X} - \hat{\mathcal{X}}^H) \end{bmatrix} \ ,
\end{aligned}
\tag{4.15}
$$

and defining the following matrices:

$$\mathbf{U} = \begin{bmatrix} \mathbf{I}_{n_T n_R} \\ \mathcal{X} \end{bmatrix} \tag{4.16}$$

$$\mathbf{V} = \Sigma (\mathcal{X} - \hat{\mathcal{X}})^H \begin{bmatrix} -\hat{\mathcal{X}} & \mathbf{I}_{T n_R} \end{bmatrix} \tag{4.17}$$

$$\mathbf{Z} = \begin{bmatrix} \mathbf{0}_{n_T n_R} & \mathbf{0}_{n_T n_R \times T n_R} \\ (\mathcal{X} - \hat{\mathcal{X}}) & \mathbf{0}_{T n_R} \end{bmatrix} \tag{4.18}$$

Thus, $\mathbf{R}_{\mathbf{zz}|\mathbf{X}}\mathbf{F}$ can be expressed in term of $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{Z}$ as:

$$\mathbf{R}_{\mathbf{zz}|\mathbf{X}}\mathbf{F} = E_s \, \mathbf{U} \, \mathbf{V} + \mathbf{Z} \,. \tag{4.19}$$

Using [14, Appendix B, Theorem 2], the nonzero eigenvalues of $E_s \mathbf{UV} + \mathbf{Z}$ are arbitrary close to the nonzero eigenvalues of $E_s \mathbf{VU}$ and $\mathbf{Z}(\mathbf{I} - \mathbf{U}(\mathbf{VU})^{-1}\mathbf{V})$ as E_s approach infinity if

$$\mathbf{VU} = \Sigma \, (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}}) \tag{4.20}$$

has full rank. Since $\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}} = (\mathbf{X} - \hat{\mathbf{X}})^H (\mathbf{X} - \hat{\mathbf{X}})$ is nonnegative definite Hermitian matrix and $\mathbf{VU}$ is assumed to be full rank, it follows that the number of eigenvalues of $E_s \mathbf{VU}$ are $n_T n_R$. These eigenvalues are real and linearly proportional to the E_s. Let's define $\mathcal{A}_{\mathbf{X},\hat{\mathbf{X}}} = \mathbf{I}_{n_R} \otimes \mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}}$ and through simple manipulation, it can be shown that

$$\mathbf{Z}\left(\mathbf{I} - \mathbf{U}(\mathbf{VU})^{-1}\mathbf{V}\right) =$$
$$\begin{bmatrix} \mathbf{0}_{n_T n_R} & \mathbf{0}_{n_T n_R \times T n_R} \\ (\mathcal{X} - \hat{\mathcal{X}}) - (\mathcal{X} - \hat{\mathcal{X}})\mathcal{A}_{\mathbf{X},\hat{\mathbf{X}}}^{-1}(\mathcal{X} - \hat{\mathcal{X}})^H \hat{\mathcal{X}} & -(\mathcal{X} - \hat{\mathcal{X}})\mathcal{A}_{\mathbf{X},\hat{\mathbf{X}}}^{-1}(\mathcal{X} - \hat{\mathcal{X}})^H \end{bmatrix} . \tag{4.21}$$

Thus, the nonzero eigenvalues of $\mathbf{Z}\left(\mathbf{I} - \mathbf{U}(\mathbf{VU})^{-1}\mathbf{V}\right)$ are the nonzero eigenvalues of $-(\mathbf{X} - \hat{\mathbf{X}})\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}}^{-1}(\mathbf{X} - \hat{\mathbf{X}})^H$ with multiplicity n_R. For a negative projection matrix, its eigenvalues are minus one or zero [64, pp. 14]. Since $-(\mathbf{X} - \hat{\mathbf{X}})\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}}^{-1}(\mathbf{X} - \hat{\mathbf{X}})^H$ is a full rank negative projection matrix, its eigenvalues equal to minus one with

multiplicity of n_T. Therefore, the eigenvalues of $\mathbf{Z}\left(\mathbf{I} - \mathbf{U}(\mathbf{VU})^{-1}\mathbf{V}\right)$ are equal to minus unity with multiplicity of $n_T n_R$.

When $\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}}$ is not full rank (i.e. $r = rank(\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}}) < n_T$), $\mathbf{VU}$ will has rank rn_R. The number of positive eigenvalues of $E_s\mathbf{VU}$ and the multiplicity of the eigenvalue minus one of $\mathbf{Z}\left(\mathbf{I} - \mathbf{U}(\mathbf{VU})^{-1}\mathbf{V}\right)$ will then be rn_R. $\triangle$

Proposition 4.2 *The asymptotic pairwise error probability is equal to*

$$P(\mathbf{X} \to \hat{\mathbf{X}}) = \frac{E_s^{-rn_R} \begin{pmatrix} 2rn_R - 1 \\ rn_R \end{pmatrix}}{\prod_{l=1}^{rn_R} \lambda_l}, \tag{4.22}$$

where λ_l are nonzero eigenvalues of $\Sigma(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})$.

Proof: By the Lebesgue's dominated convergence theorem, the limit and the integral commute. Therefore, using (4.13)

$$\lim_{E_s \to \infty} E_s^{rn_R} P(\mathbf{X} \to \hat{\mathbf{X}}) = \lim_{E_s \to \infty} E_s^{rn_R} \int_{-j\infty+\epsilon}^{j\infty+\epsilon} s^{-1} G_f(s) \frac{ds}{2\pi j},$$

$$= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} \lim_{E_s \to \infty} E_s^{rn_R} s^{-1} G_f(s) \frac{ds}{2\pi j}. \tag{4.23}$$

Since half of the nonzero eigenvalues of $\mathbf{R}_{\mathbf{zz}|\mathbf{X}}\mathbf{F}$ are equal to minus unity with multiplicity of rn_R and the other half are positive which equal to the eigenvalues of $E_s\Sigma(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})$, it is obvious that

$$\lim_{E_s \to \infty} E_s^{rn_R} s^{-1} G_f(s) = \frac{1}{s\,(s-1)^{rn_R}\,(-1)^{rn_R}\,\prod_{l=1}^{rn_R} \lambda_l s} \tag{4.24}$$

where λ_l are the nonzero eigenvalues of $\Sigma(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})$. Thus,

$$
\begin{aligned}
\lim_{E_s \to \infty} E_s^{rn_R} P(\mathbf{X} \to \hat{\mathbf{X}}) &= \int_{-j\infty+\epsilon}^{j\infty+\epsilon} \frac{1}{s\,(s-1)^{rn_R}\,(-1)^{rn_R}\,\prod_{l=1}^{rn_R} \lambda_l s} \frac{ds}{2\pi j} \\
&= -\mathrm{Res}\left(\frac{1}{s\,(s-1)^{rn_R}\,(-1)^{rn_R}\,\prod_{l=1}^{rn_R} \lambda_l}, s_l = 1 \right), \\
&= \frac{-1}{(rn_R-1)!\,\prod_{l=1}^{rn_R} \alpha_l} \frac{d^{rn_R-1}}{ds^{rn_R-1}} \frac{1}{s^{rn_R+1}\,(-1)^{rn_R}} \bigg|_{s=1}, \\
&= \frac{1}{\prod_{l=1}^{rn_R} \lambda_l} \begin{pmatrix} 2rn_R-1 \\ rn_R \end{pmatrix}
\end{aligned}
\tag{4.25}
$$

where the (rn_R-1)th derivative is calculated using Leibniz's rule of differentiation. The asymptotic pairwise error probability in the proposition is then immediately followed. $\triangle$

This expression for the pairwise error probability is asymptotically tight to the true error probability as compared to the standard Chernoff bound at high SNRs. However at low SNRs, this asymptotic expression very loose. A similar asymptotic tight bound on the pairwise error probability is also derived in [87].

4.3.2 Pairwise Error Probability Upper Bound

Through simple manipulation, the condition given by (4.6) can be shown to be equivalent to

$$
\sqrt{E_s}\, \mathbf{h}^H\,(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})\, \mathbf{h} \;<\; 2Re\{\mathbf{n}^H(\hat{\mathcal{X}} - \mathcal{X})\,\mathbf{h}\},
\tag{4.26}
$$

where $Re\{.\}$ means the real part of an argument. The term on the left hand side of (4.26) is a constant and the term on the right hand side of (4.26) is a zero mean Gaussian random variable with variance $4d^2(\mathbf{X}, \hat{\mathbf{X}})$ where

$$
\begin{aligned}
d^2(\mathbf{X}, \hat{\mathbf{X}}) &= \sum_{t=1}^{T} \sum_{j=1}^{n_R} \left| \sum_{i=1}^{n_T} h_{i,j}\,(x_i(t) - \hat{x}_i(t)) \right|^2 \\
&= \mathbf{h}^H\,(\mathcal{X} - \hat{\mathcal{X}})^H(\mathcal{X} - \hat{\mathcal{X}})\mathbf{h}.
\end{aligned}
\tag{4.27}
$$

Assuming ideal CSI is available at the receiver, the pairwise error probability condition on $\mathbf{h}$ is upper bounded by [89]

$$P(\mathbf{X} \to \hat{\mathbf{X}} \mid \mathbf{h}) \leq \exp\left(-\frac{E_s \, d^2(\mathbf{X}, \hat{\mathbf{X}})}{4}\right) . \tag{4.28}$$

The correlated fading channels can be written as

$$\mathbf{h} = \boldsymbol{\Sigma}^{1/2} \, \tilde{\mathbf{h}} , \tag{4.29}$$

where $\boldsymbol{\Sigma}$ is the $n_T n_R \times n_T n_R$ fading channel correlation matrix and $\tilde{\mathbf{h}}$ is an $n_T n_R$ zero mean i.i.d complex vector with covariance $E[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H] = \mathbf{I}_{n_T n_R}$. Thus,

$$d^2(\mathbf{X}, \hat{\mathbf{X}}) = \tilde{\mathbf{h}}^H \, (\boldsymbol{\Sigma}^{1/2})^H \, (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}}) \, \boldsymbol{\Sigma}^{1/2}\tilde{\mathbf{h}} \tag{4.30}$$

Let's define the nonnegative definite Hermitian matrix

$$\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma}) = (\boldsymbol{\Sigma}^{1/2})^H \, (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}}) \, \boldsymbol{\Sigma}^{1/2} . \tag{4.31}$$

Then by using eigenvalue decomposition, we can represent $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ by a unitary matrix $\mathbf{W}$ and a real diagonal matrix $\boldsymbol{\Delta}$

$$\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma}) = \mathbf{W}\boldsymbol{\Delta}\mathbf{W}^H . \tag{4.32}$$

The columns of $\mathbf{W}$, $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{n_T n_R}\}$ are a complete orthogonal basis of $\mathbb{C}^{n_T n_R}$. They are given by the eigenvectors of $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ with the corresponding eigenvalues $\lambda_l \geq 0$, which are the diagonal elements of $\boldsymbol{\Delta}$. Let q denote the rank of $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ and hence the number of nonzero eigenvalues λ_l. Equation (4.30) can be rewritten as

$$d^2(\mathbf{X}, \hat{\mathbf{X}}) = \sum_{l=1}^{q} \lambda_l \, |\beta_l|^2 \tag{4.33}$$

where

$$\beta_l = \tilde{\mathbf{h}}^H \, \mathbf{w}_l \tag{4.34}$$

Since $\tilde{\mathbf{h}}$ is zero mean complex Gaussian random vector with covariance $E[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H] = \mathbf{I}_{n_T n_R}$ and $\{\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_{n_T n_R}\}$ is an orthogonal basis of $\mathbb{C}^{n_T n_R}$, β_l are zero mean independent complex Gaussian random variables with variance one. Thus, $|\beta_l|$ are independent Rayleigh distributions with pdf

$$p(|\beta_l|) = 2 \, |\beta_l| \, \exp(-|\beta_l|^2) \,. \tag{4.35}$$

Hence, by averaging the conditional probability of error

$$P(\mathbf{X} \to \hat{\mathbf{X}} \mid \mathbf{h}) \le \exp\left(-\frac{E_s \sum_{l=1}^{q} \lambda_l \, |\beta_l|^2}{4}\right) \tag{4.36}$$

with respect to independent Rayleigh distributions of $|\beta_l|$, we get the following upper bound of the pairwise error probability

$$
\begin{aligned}
P(\mathbf{X} \to \hat{\mathbf{X}}) &\le \frac{1}{\prod_{l=1}^{q} 1 + (E_s/4) \, \lambda_l} \\
&= \frac{1}{\det\left(\mathbf{I}_{n_T n_R} + (E_s/4) \, (\mathbf{\Sigma}^{1/2})^H \, (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}}) \, \mathbf{\Sigma}^{1/2}\right)} \\
&= \frac{1}{\det\left(\mathbf{I}_{n_T n_R} + (E_s/4) \, \mathbf{\Sigma} \, (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}})\right)}
\end{aligned}
\tag{4.37}
$$

The second and third lines in (4.37) follow from the following matrix properties

$$\det(I + B) = \prod(1 + \rho_l) \,, \tag{4.38}$$

$$\det(I + A\,B) = \det(I + B\,A) \tag{4.39}$$

respectively, where ρ_l are eigenvalues of B. It follows from the inequality in (4.37) that for high SNR

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \le (E_s/4)^{-q} \, \frac{1}{\prod_{l=1}^{q} \lambda_l} \,. \tag{4.40}$$

This asymptotic bound is different from the asymptotic tight expression in (4.22) by only a constant of

$$10 log(4) - \frac{10}{q} log \binom{2q-1}{q} \tag{4.41}$$

decibels. As the diversity $q \to \infty$, this difference approach 0 dB.

When the fading channels are independent, i.e $\boldsymbol{\Sigma} = \mathbf{I}$, the pairwise error probability is

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \frac{1}{\det\left(\mathbf{I}_{n_T n_R} + (E_s/4) (\mathcal{X} - \hat{\mathcal{X}})^H (\mathcal{X} - \hat{\mathcal{X}})\right)} . \qquad (4.42)$$

Since $(\mathcal{X} - \hat{\mathcal{X}}) = I_{n_R} \otimes (\mathbf{X} - \hat{\mathbf{X}})$, equation (4.42) can be rewritten as

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \left(\frac{1}{\det\left(\mathbf{I}_{n_T} + (E_s/4) (\mathbf{X} - \hat{\mathbf{X}})^H (\mathbf{X} - \hat{\mathbf{X}})\right)}\right)^{n_R} , \qquad (4.43)$$

which is the same as [89, Eq. (8)]

4.3.3 Pairwise Error Probability Upper Bound for Large q

When $q = rank(\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma}))$ is large, this implies that $d^2(\mathbf{X}, \hat{\mathbf{X}})$ is the sum of a large number of independent variables. Thus, according to the central limit theorem, the distribution of $D = d^2(\mathbf{X}, \hat{\mathbf{X}})$ approaches a Gaussian distribution with mean

$$\mu_D = \sum_{l=1}^{q} \lambda_l \qquad (4.44)$$

and variance

$$\theta_D^2 = \sum_{l=1}^{q} \lambda_l^2 . \qquad (4.45)$$

Note that since λ_l are nonzero eigenvalues of a nonnegative definite Hermitian matrix, D is always positive. Hence by averaging (4.28) over the Gaussian distribution of D, the pairwise error probability can be upper bounded by

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \int_{D=0}^{+\infty} \exp\left(-\frac{E_s D}{4}\right) p(D) dD \qquad (4.46)$$

where $p(D)$ is the pdf of the Gaussian random variable D. Using the identity

$$\int_{D=0}^{+\infty} \exp(-\gamma D) p(D) dD = \exp\left(\frac{1}{2}\gamma^2 \theta_D^2 - \gamma \mu_D\right) Q\left(\frac{\gamma \theta_D^2 - \mu_D}{\theta_D}\right) , \gamma > 0 \quad (4.47)$$

the upper bound of the pairwise error probability in (4.46) can be rewritten as

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \exp\left(\frac{1}{32}E_s^2 \sum_{l=1}^{q}\lambda_l^2 - \frac{E_s}{4}\sum_{l=1}^{q}\lambda_l\right) Q\left(\frac{(E_s/4)\sum_{l=1}^{q}\lambda_l^2 - \sum_{l=1}^{q}\lambda_l}{\sqrt{\sum_l^q \lambda_l^2}}\right)$$
(4.48)

where $Q(x)$ is the complementary error function. For high SNR (i.e. large E_s)

$$\frac{(E_s/4)\sum_{l=1}^{q}\lambda_l^2 - \sum_{l=1}^{q}\lambda_l}{\sqrt{\sum_l^q \lambda_l^2}} \approx \frac{E_s}{4}\sqrt{\sum_l^q \lambda_l^2} .$$
(4.49)

Consequently, by using the inequality

$$Q(x) \leq \frac{1}{2}\, e^{-x^2/2} , \ x \geq 0$$
(4.50)

the bound in (4.48) can be further approximated by

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \frac{1}{2}\, \exp\left(-\frac{E_s}{4}\sum_{l=1}^{q}\lambda_l\right)$$
(4.51)

When the fading channels are independent, the upper bound of pairwise error probability in (4.51) is reduced to

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq \frac{1}{2}\, \exp\left(-\frac{E_s n_R}{4}\sum_{l=1}^{r}\sigma_l\right)$$

$$= \frac{1}{2}\, \exp\left(-\frac{E_s n_R}{4}\sum_{t=1}^{T}\sum_{i=1}^{n_T}|x_t^i - \hat{x}_t^i|^2\right)$$
(4.52)

where r and σ_l are rank and nonzero eigenvalues of the distance matrix $\mathbf{A}_{\mathbf{X},\hat{\mathbf{X}}} = (\mathbf{X} - \hat{\mathbf{X}})^H(\mathbf{X} - \hat{\mathbf{X}})$, respectively. This expression is the same as one derived in [21, 110].

4.4 Space-Time Code Design Criteria

In this section, we describe the rank determinant, the trace and the sum of pairwise error probabilities design criteria for constructing the space-time codes to operate in spatially correlated fading channels.

4.4.1 The Rank Determinant Criterion

From (4.40), the pairwise error probability is exponentially dependent on the rank q and inversely proportional to the determinant $\prod \lambda_l$ of the matrix $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$. Thus in order to minimised the error probability, we need to maximise the minimum rank q and the minimum determinant $\prod \lambda_l$ of the matrix $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ over all pairs of distinct codewords. When $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ is full rank, the upper bound of the pairwise error probability in (4.40) can be written as

$$P(\mathbf{X} \to \hat{\mathbf{X}}) \leq (E_s/4)^{-n_T n_R} \frac{1}{|\boldsymbol{\Sigma}| \, |(\mathbf{X} - \hat{\mathbf{X}})^H (\mathbf{X} - \hat{\mathbf{X}})|^{n_R}} . \tag{4.53}$$

The above criterion is then equivalent to:

1. Ensure $\mathbf{A}_{\mathbf{X}, \hat{\mathbf{X}}}$ is full rank for all pairs of distinct codewords.

2. Maximise the minimum determinant $\det(\mathbf{A}_{\mathbf{X}, \hat{\mathbf{X}}})$ along all pairs of distinct codewords.

Thus when $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ is full rank, the rank determinant design criterion is independent from the fading channel correlations.

4.4.2 The Trace Criterion

This trace criterion is only applicable when $q = rank(\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma}))$ is large (i.e. $q \geq 4$) since it is based on the performance analysis for large q. As from (4.51), minimising the pairwise error probability is equivalent to maximise the sum of the eigenvalues of $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$. Since

$$\sum_{l=1}^{q} \lambda_l = \text{trace}\{\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})\} , \tag{4.54}$$

the trace criterion is then to maximise the minimum trace of the matrix $\mathbf{G}(\mathbf{X}, \hat{\mathbf{X}}, \boldsymbol{\Sigma})$ over all pairs of distinct codewords. It can be noticed that the trace criterion does not

guarantee full diversity space-time codes to be constructed. In addition, the design criterion is dependent on the channel correlation matrix even when the channel correlation matrix and the distance matrix $\mathbf{A_{X,\hat{X}}}$ are full rank.

4.4.3 Sum of PEP criterion

Both the rank determinant and the trace design criteria aim to miminise the error probability of the dominant event. We now describe a design criterion which involves minimising the sum of all pairwise error probabilities. From (4.40), the pairwise error probability is inversely proportional to $\prod_{l=1}^{q} \lambda_l$. Hence, to minimise the sum of the pairwise error probability of all error events we need to minimise

$$f(E) = \sum_{e=1}^{E-1} \sum_{d>e} \frac{1}{\prod_{l=1}^{q} \lambda_l^{e,d}} , \tag{4.55}$$

where $\lambda_l^{e,d}$ are the nonzero eigenvalues of $\mathbf{G}(\mathbf{X_e}, \mathbf{X_d}, \mathbf{\Sigma}))$ and E is the total number of codewords. For a space-time code with rate of R bit/sec/Hz, the total number of codewords $E = 2^{RT}$, which is generally very large. For example, if the space-time code has a rate of 2 bit/sec/Hz and the frame length is 100 symbols, there would be $E = 2^{200}$ possible codewords. When $\mathbf{G}(\mathbf{X_e}, \mathbf{X_d}, \mathbf{\Sigma}))$ is full rank, the pairwise error probability is upper bounded by (4.53). Thus, minimising $f(E)$ (and hence the sum of the pairwise error probability of all error events) is now equivalent to minimising

$$d(E) = \sum_{e=1}^{E-1} \sum_{d>e} \frac{1}{|(\mathbf{X_e} - \mathbf{X_d})^H (\mathbf{X_e} - \mathbf{X_d})|^{n_R}} . \tag{4.56}$$

Note that $d(E)$ depends on the number of receive antennas but is independent of the fading channel correlation. The codes design criterion is then equivalent to:

- Ensure that $\mathbf{A_{X_e, X_d}}$ is full rank for all pairs of distinct codewords.

- Minimise the metric $d(E)$.

4.5   Space Time Trellis Codes

4.5.1   The Space-Time Trellis Encoder
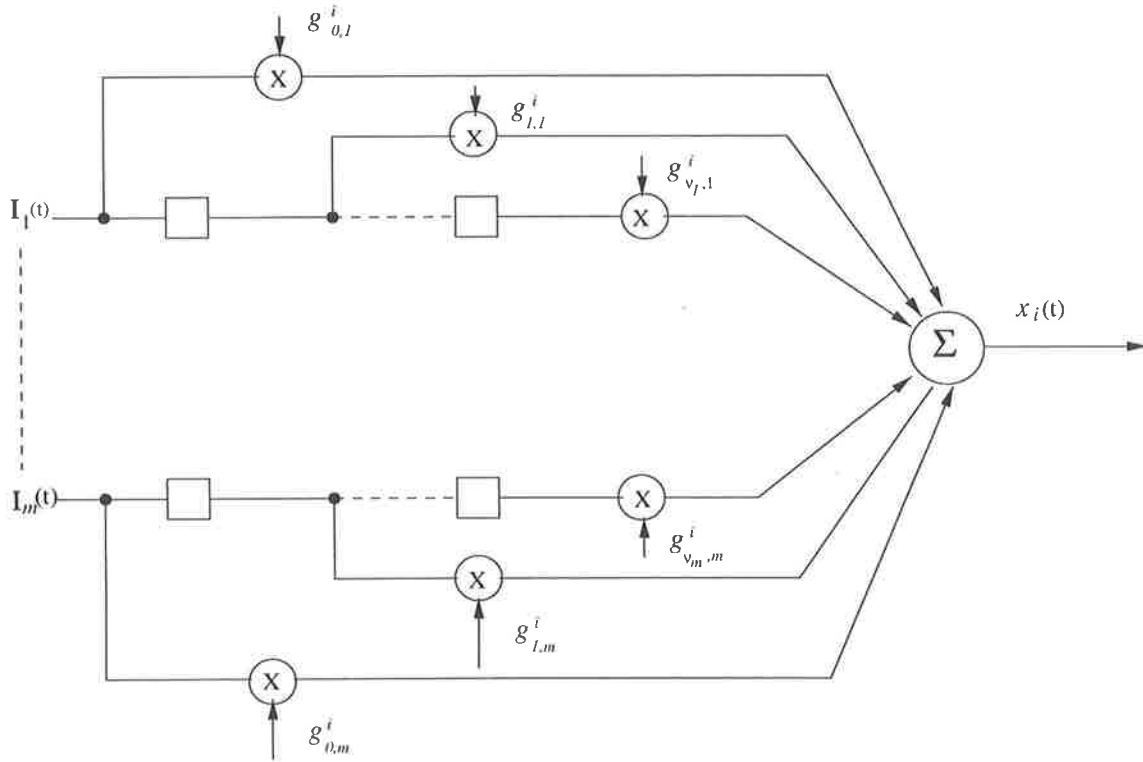
Figure 4.2: Space-time trellis encoder

We consider a space time trellis encoder with MPSK modulation as shown in Fig. 4.2. This space time trellis encoder achieves a bandwidth efficiency of $m = \log_2 M$ bits/s/Hz. We use the same encoder representation as described in [21, 110]. At each time t, m binary inputs $I_1(t), I_2(t), ..., I_m(t)$ are fed into the encoder which consists of m feed forward shift registers, each with a memory order of ν_k. The kth input bit $I_k(t)$, which is fed to the kth shift register, and the delayed bits in the kth shift register's memory are then multiplied with the ith encoder coefficients set

$$\mathbf{g}_k^i = [\ g_{0,k}^i,\ g_{1,k}^i, ..., g_{\nu_k,k}^i\ ],\tag{4.57}$$

where $g_{l,k}^i \in \{0, 1, ..., M-1\}$ with $k = 1, 2, ..., m$; $l = 0, ..., \nu_k$ and $i = 1, ..., n_T$. The multiplier outputs are added modulo M and mapped into signals from the MPSK constellation, giving the encoder output to the ith transmit antenna

$$x_i(t) = \mathcal{M}\left(\sum_{k=1}^m \sum_{l=0}^{\nu_k} g_{l,k}^i \ I_k(t-l) \mod M\right), \tag{4.58}$$

where $\mathcal{M}(x) = \exp(2\pi jx/M)$ is the mapping function that maps integer values to the MPSK constellation. The memory order of the kth shift register, ν_k, is determined from the total memory order ν by

$$\nu_k = \left\lfloor \frac{\nu + k - 1}{\log_2 M} \right\rfloor, \tag{4.59}$$

where $\lfloor x \rfloor$ the maximum integer not larger than x.

4.5.2 Construction of the Space-Time Trellis Codes

Using the sum of pairwise error probabilities design criterion, we perform a systematic search to construct new QPSK space-time trellis codes for two transmit and two receive antennas. We perform a search over all possible pairs of error event with a path length of $T = 4 + \max_k \nu_k$. The values of $d(E)$ in Table 4.1 are computed for this path length. We have tried with a longer path length but the same codes are found. Even though these codes are designed for two receive antennas, simulation results show that they also perform well for other number of receive antennas. The newly found codes with 4, 8, 16 and 32 states are summarized in Table 4.1 where

$$\mathbf{g}_k = [\ (g_{0,k}^1, g_{0,k}^2, ..., g_{0,k}^{n_T}), ..., (g_{\nu_k,k}^1, g_{\nu_k,k}^2, ..., g_{\nu_k,k}^{n_T})\ ]. \tag{4.60}$$

The codes that are found in [89] (TSC) by the rank determinant criterion and in [21] (CYV) by the trace criterion are included as references. All these codes achieve full diversity. Their corresponding minimum determinant and minimum trace (i.e. squared Euclidean distance) are also shown.

Table 4.1: QPSK space-time trellis codes for correlated flat Rayleigh fading

Code	ν	$\mathbf{g}_1$ and $\mathbf{g}_2$	$\det(\mathbf{A}_{\mathbf{x},\hat{\mathbf{x}}})$	$\mathrm{trace}(\mathbf{A}_{\mathbf{x},\hat{\mathbf{x}}})$	$d(E)$
TSC	2	$\mathbf{g}_1 = [(0,2),(2,0)]$ $\mathbf{g}_2 = [(0,1),(1,0)]$	4	4	95.72
CYV	2	$\mathbf{g}_1 = [(0,2),(1,2)]$ $\mathbf{g}_2 = [(2,3),(2,0)]$	4	10	35.97
New	2	$\mathbf{g}_1 = [(2,3),(2,2)]$ $\mathbf{g}_2 = [(1,2),(2,0)]$	8	8	29.41
TSC	3	$\mathbf{g}_1 = [(0,2),(2,0)]$ $\mathbf{g}_2 = [(0,1),(1,0),(2,2)]$	12	8	17.76
CYV	3	$\mathbf{g}_1 = [(2,2),(2,1)]$ $\mathbf{g}_2 = [(2,0),(1,2),(0,2)]$	8	12	15.18
New	3	$\mathbf{g}_1 = [(2,3),(0,2)]$ $\mathbf{g}_2 = [(0,2),(1,0),(2,2)]$	16	10	9.96
TSC	4	$\mathbf{g}_1 = [(0,2),(2,0),(0,2)]$ $\mathbf{g}_2 = [(0,1),(1,2),(2,0)]$	12	8	11.50
CYV	4	$\mathbf{g}_1 = [(1,2),(1,3),(3,2)]$ $\mathbf{g}_2 = [(2,0),(2,2),(2,0)]$	8	16	7.89
New	4	$\mathbf{g}_1 = [(2,2),(3,2),(0,2)]$ $\mathbf{g}_2 = [(2,0),(2,3),(2,2)]$	32	12	5.40
TSC	5	$\mathbf{g}_1 = [(0,2),(2,2),(3,3)]$ $\mathbf{g}_2 = [(0,1),(1,1),(2,0),(2,2)]$	12	12	5.28
CYV	5	$\mathbf{g}_1 = [(0,2),(2,3),(1,2)]$ $\mathbf{g}_2 = [(2,2),(1,2),(2,3),(2,0)]$	20	16	3.83
New	5	$\mathbf{g}_1 = [(0,2),(3,2),(2,0)]$ $\mathbf{g}_2 = [(2,2),(2,1),(0,3),(2,2)]$	36	14	2.98

4.5.3   Performance of Space-Time Trellis Codes

This section compares the performance of the new codes with the TSC and CYV codes for both independent and correlated flat Rayleigh fading. All these codes achieve a bandwidth efficiency of 2 bits/s/Hz. At the receiver, a maximum likelihood Viterbi decoder with ideal channel state information is employed. We simulate for different number of receive antennas. In the simulation, each frame consists of 130 coded symbols transmitted out of each transmit antenna. The frame error rate
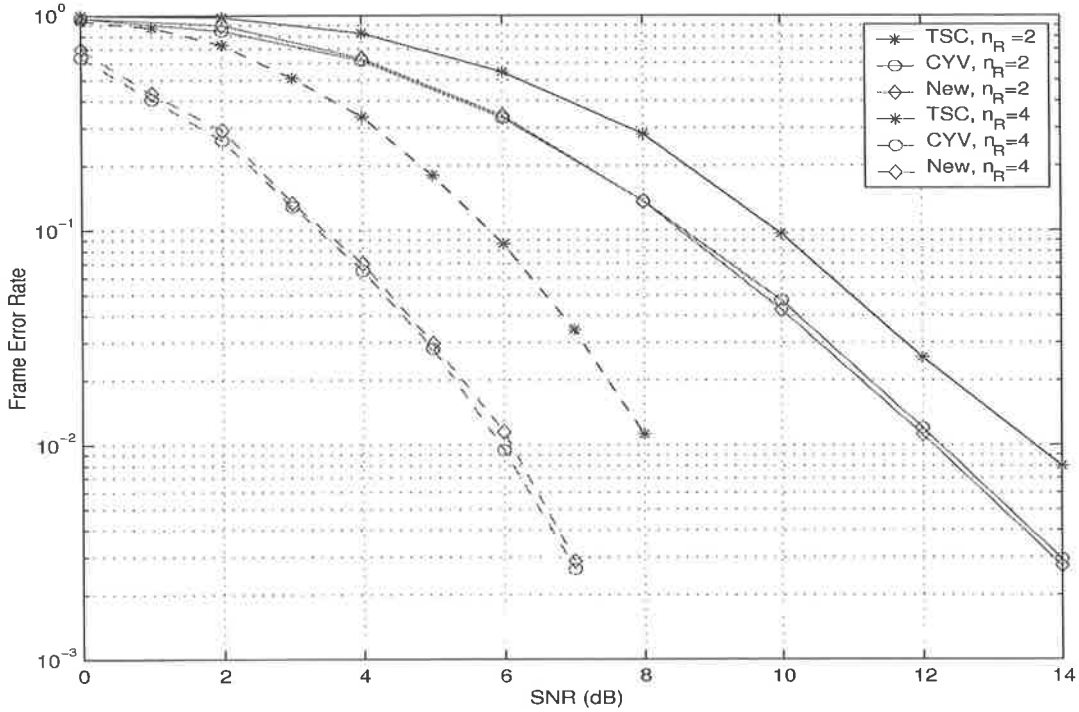
Figure 4.3: Performance of 4 states QPSK codes with 2 and 4 receive antennas

(FER) performance of these codes are obtained and plotted against the signal to noise ratio (SNR). We define $SNR = n_T E_s/N_0$ where N_0 is the variance of the additive noise at the receiver. For independent fading channels where the fading coefficient $h_{i,j}$ is a complex random variable with variance of 0.5 per dimension, this SNR is equal to the average received signal to noise ratio per receive antenna.

The performance of the 4, 8 and 16 states QPSK space-time trellis codes with 2 and 4 receive antennas in quasi-static Rayleigh fading is shown in Fig. 4.3, Fig. 4.4 and Fig. 4.5, respectively. The transmitted signals from different antennas are assumed to undergo independent fades. For the 2 receive antennas case, the newly found codes outperform the TSC codes by approximately 0.6 dB to 1.5 dB and the CYV codes by 0.1 dB to 0.5 dB. For the 4 receive antennas case, the new QPSK codes perform better than the TSC codes by 0.7 dB to 2.2 dB. The new 8 and 16
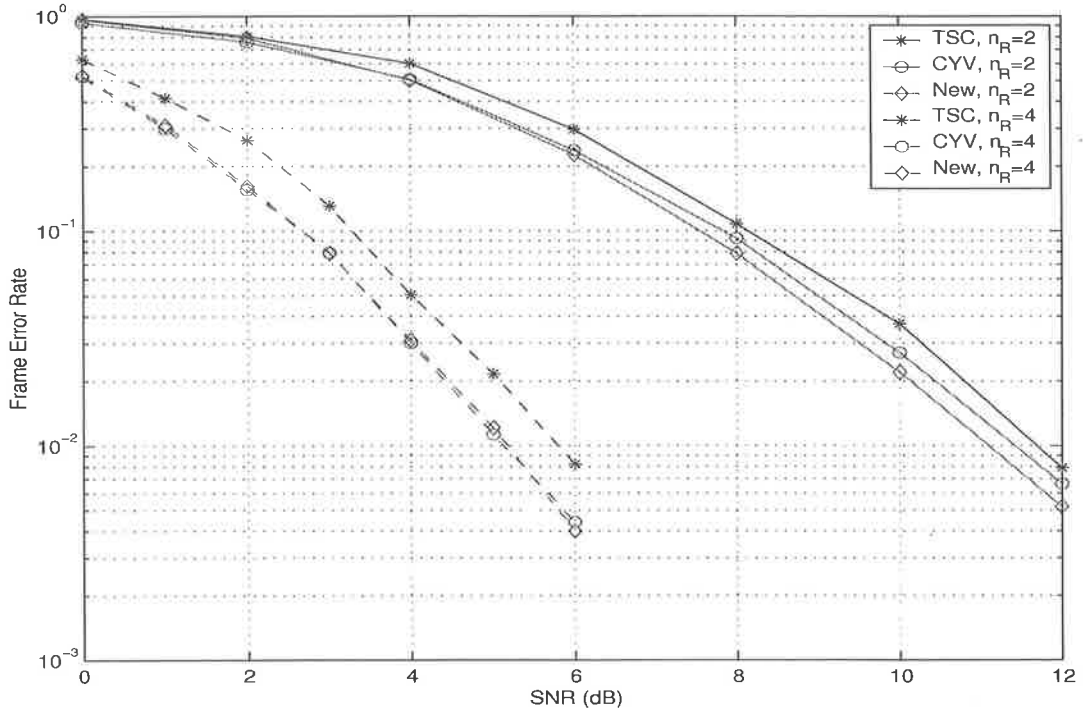
Figure 4.4: Performance of 8 states QPSK codes with 2 and 4 receive antennas

states QPSK codes perform equally well as the CYV codes while the new 4 state code performs slightly poorer than the corresponding CYV code by 0.1 dB.

The simulation results in Fig. 4.3, Fig. 4.4 and Fig. 4.5 all show that the CYV codes achieve a better performance than the TSC codes despite having smaller (or equal) minimum determinant. This further demonstrates the weakness of the rank determinant criterion. As shown in Table 4.1, all new codes have smaller minimum square Euclidean distances (i.e. minimum trace) than the corresponding CYV codes. However simulation results show that the new codes still perform better than the CYV codes for 2 receive antennas case in spite of smaller minimum squared Euclidean distances. This suggests that by taking into account all the possible error events into the design process, additional performance gain can be obtained, and that the metric $d(E)$ is a better measure of the code error performance than the
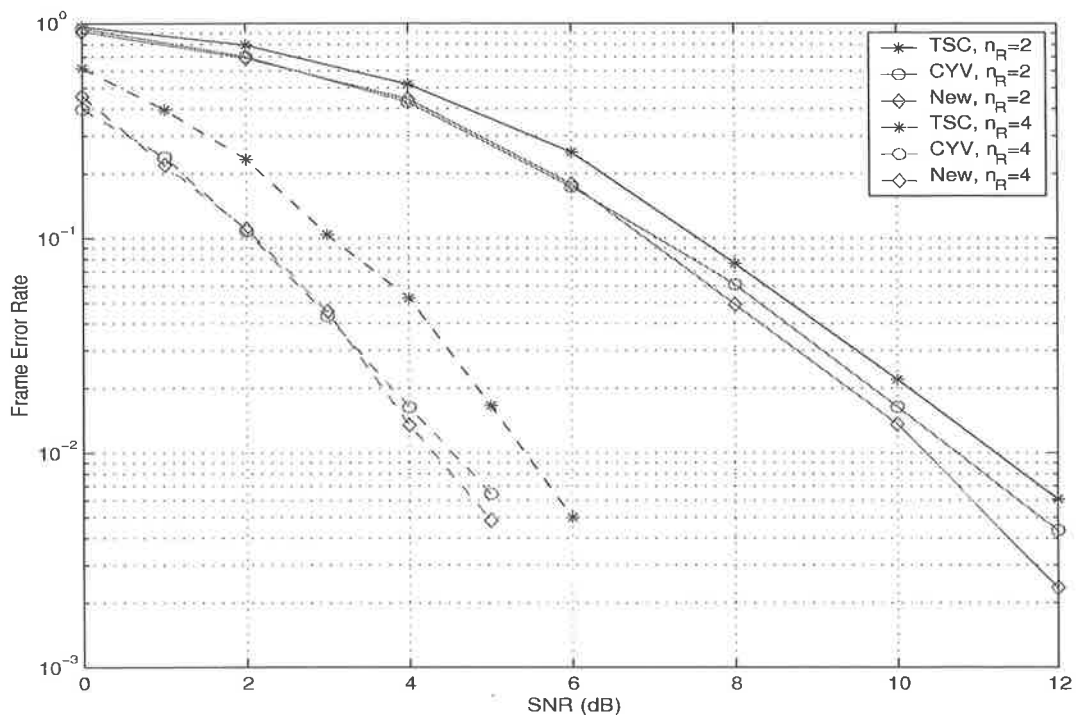
Figure 4.5: Performance of 16 states QPSK codes with 2 and 4 receive antennas

determinant and the trace metrics. For 4 receive antennas, the new codes do not outperform the CYV codes like for the 2 receive antenna case because the new codes are designed for 2 receive antennas. More importantly, when 4 receive antennas are employed, the system doesn't need to operate in the high SNR region to obtain a reasonable FER. Hence the metric $d(E)$ is no longer closely related to the code error performance because it is based on the asymptotic (high SNR) analysis.

The performance of these space time trellis codes in correlated fading environment is shown in Fig. 4.6. The results are obtained for 4 and 8 states codes with two receives antennas. The correlation factor between any two propagation paths is 0.5. As shown in Fig. 4.6, the new 4 and 8 states codes have the same performance as the corresponding CYV codes, while perform better than the 4 and 8 states TSC codes by 1.2 dB and 0.4 dB respectively.
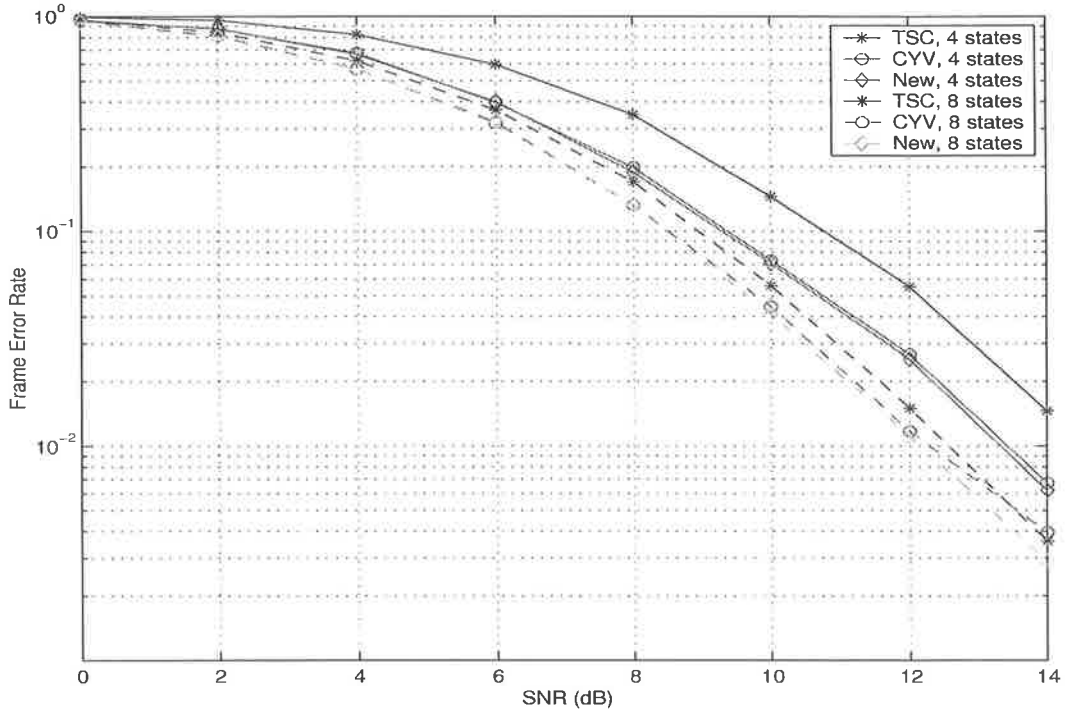
Figure 4.6: Performance of 4 and 8 states QPSK STTCM in correlated Rayleigh fading with 2 receive antennas

4.6 Conclusions

In this chapter, we analysed the performance of the space-time coded systems that employ multiple transmit and multiple receive antennas in spatially correlated Rayleigh fading channels. We presented an asymptotic tight expression and two upper bounds for the pairwise error probability. A number of new QPSK codes for two transmit antennas based on a new criterion of minimising the sum of the pairwise error probabilities were found. Simulation results showed that these new codes are superior to the other codes found by the rank determinant and the trace design criteria.

Chapter 5

Joint Space-Time Trellis Decoding and Channel Estimation

This chapter addresses the issues of joint space-time trellis decoding and channel estimation in time-varying fading channels that are spatially and temporally correlated. A recursive space-time receiver which incorporates per-survivor processing (PSP) and Kalman filtering into the Viterbi algorithm is proposed. The channel time-evolution is modeled by a multichannel autoregressive process, and a bank of Kalman filters is used to track the channel variations. Computer simulation results demonstrate that a performance close to the maximum likelihood receiver with perfect channel state information (CSI) can be obtained. The effects of the spatially independent fading channel assumption on the performance of a receiver operated in spatially correlated time-varying channels are examined.

5.1 Introduction

The works described in the previous chapter on space-time coding assume that the receiver has perfect channel state information of the wireless fading channels. This

assumption is only applicable to wireless applications that have limited mobility where the channels can be accurately estimated by using pilot symbols inserted periodically during transmissions. In this chapter, we consider the problem of decoding space-time trellis codes in time-varying fading channels. In recent years, there are a number of published works addressing this issue [23, 24, 60, 106]. In [23], an iterative space-time receiver that performs joint channel estimation and detection based on the expectation-maximization (EM) algorithm has been described. It requires the insertion of pilot symbols at the start of each frame. The performance of this receiver is reasonably good when the length of the frame is relatively short, but degrades significantly as the frame length increases. Thus for very fast time-varying channels, many pilot symbols would be needed and hence reduce the system throughput. In [60], Liu *et al.* proposes an iterative receiver based on Kalman filtering for decoding the Alamouti's space-time block code [7]. It models the time-varying channels as autoregressive (AR) processes and uses a Kalman filter for channel tracking. Although this receiver can accurately track the fading channels, it is very sensitive to channel estimation errors. Even small channel estimation errors can cause a considerable performance loss. In recent years, per-survivor processing (PSP) has been known as an effective approach for simultaneous estimation of the data sequence and the unknown channel parameters in single-input-single-output (SISO) systems [79]. In [106], a PSP based receiver for decoding space-time trellis codes is proposed which uses an accelerated self-tuning least mean square (LMS) algorithm for tracking the fading channels. It provides good performance for both slowly and moderately time-varying channels. Similarly to [106], a PSP based receiver which combines the Viterbi algorithm with data-aided channel estimation is proposed in [24]. However, the difference is that it uses a bank of Kalman filters to track the channel variations. Simulation results in [24, 106] show that these PSP based receivers generally offer superior performance than other competing receivers such as

the Wiener Interpolation Filter (WIF) [41], the iterative Kalman filter [24] and the iterative EM based receiver [23].

All the works mentioned previously are based on the assumption of independent and identically distributed (i.i.d) fading channels. However, in real propagation environment, spatial correlation in the fading channels can arise due to the insufficient spacing of the antennas at the mobile station and the lack of local scatterers around the base station antennas [2, 20, 86]. Such spatial correlation can considerably reduce the performance of the receiver if it is not properly incorporated into the design of the receiver. In [72, 73], we introduce a general state space model for general diversity communication systems and propose a number of receiver structures for joint channel estimation and decoding in time-varying correlated diversity channels. The application of these receivers to OFDM systems, where the diversity is obtained from the redundancy in the frequency domain, is discussed in [72, 73, 50]. The approximation of the correlated time-frequency channels by the multichannel AR processes is detailed in [72, 50]. In this chapter, we apply the PSP receiver derived in [72, 73] to the problem of joint space-time trellis decoding and channel estimation in a wireless communication system with multiple transmit and receive antennas. This PSP receiver incorporates the per-survivor processing and Kalman filtering into the Viterbi algorithm similar to that in [24]. However our proposed PSP receiver implicitly includes the channel estimation errors into the path metric of the Viterbi algorithm while the PSP receiver in [24] computes the path metric as if there is no channel estimation error. More importantly, we do not assume i.i.d fading channels and allow the channels to be spatially and temporally correlated. The proposed PSP receiver accounts for such spatio-temporal correlation in the channels when performing the channel estimation and the decoding steps. Simulations in section 5.5 demonstrate that in fast time-varying fading channels, the performance of a receiver can degrade to an unacceptable level if the spatial correlation in the

channels is ignored. To our knowledge, there is no other work on receiver structure design which performs joint space-time trellis decoding and channel estimation in spatially correlated time-varying fading channels.

This chapter is organized as follows. In section 5.2, the system model for transmission using multiple transmit and receive antennas is introduced. In section 5.3, the wireless fading channels and their spatio-temporal correlation is reviewed. The approximation of the physical wireless fading channels by the AR processes is presented. This approximation method is differed to one in [50, 72] for the OFDM systems. In section 5.4, the proposed PSP receiver for joint space-time trellis decoding and channel estimation is described. Some simulation results to compare the performance of proposed receiver to the maximum likelihood (ML) receiver with perfect CSI is presented in section 5.5. In addition, the effects of the i.i.d fading assumption on a receiver's performance are examined.

5.2   The System Model

Consider a wireless communication system with n_T transmit and n_R receive antennas. At each time t, the encoder produces n_T outputs $x_1(t), x_2(t), ..., x_{n_T}(t)$ where $x_i(t)$ is a signal from a certain signal constellation with unit average energy. The n_T coded symbols are chosen by the encoder so that a certain criterion is satisfied. The n_T coded symbols output from the encoder are then simultaneously transmitted by n_T antennas. At the receiver, the received signal at the jth receive antenna is

$$r_j(t) = \sum_{i=1}^{n_T} h_{i,j}(t) \, x_i(t) \, \sqrt{E_s} + n_j(t) \,,$$
$$= \sqrt{E_s} \, \mathbf{x}(t)\mathbf{h}_j(t) + n_j(t) \,, \tag{5.1}$$

where E_s is the transmitted power per symbol at each transmit antenna, $\mathbf{x}(t) = [x_1(t) \; x_2(t) \; \cdots \; x_{n_T}(t)]$, $\mathbf{h}_j(t) = [h_{1,j}(t) \; h_{2,j}(t) \; \cdots \; h_{n_T,j}(t)]^T$ and $n_j(t)$ are inde-

pendent samples of a zero-mean complex Gaussian process with variance N_0. It is further assumed that the additive noise is white in space and time

$$E[n_j(t)\, n_k^*(t+\tau)] = \begin{cases} N_0\delta(\tau), & j = k \\ 0, & j \neq k \end{cases} \tag{5.2}$$

where the superscript $*$ denotes the complex conjugate operation and $\delta(.)$ denotes the Dirac delta function. The coefficient $h_{i,j}(t)$ is the fading gain for the path from transmit antenna i to receive antenna j. The gain $h_{i,j}(t)$ is modeled as a complex Gaussian random process with zero-mean and variance one. We assume that the information is transmitted in a frame of T coded symbols per transmit antenna. It is further assumed that the fading gains are independent from frame to frame. However, within each frame the fading gains can be temporally and spatially correlated. The spatio-temporal correlation between any two fading gains is discussed in more details in the next section. Stacking the observations from the receive antennas yields the n_R sufficient statistics

$$\mathbf{r}(t) = [r_1(t)\, r_2(t)\, \cdots r_{n_R}(t)]^T = \sqrt{E_s}\, \chi(t)\, \mathbf{h}(t) + \mathbf{n}(t)\ , \tag{5.3}$$

where $\mathbf{h}(t) = [\mathbf{h}_1^T(t)\, \mathbf{h}_2^T(t)\, \cdots\, \mathbf{h}_{n_R}^T(t)]^T$, $\mathbf{n}(t) = [n_1(t)\, n_2(t)\, \cdots\, n_{n_R}(t)]^T$ and $\chi(t) = \mathbf{I}_{n_R} \otimes \mathbf{x}(t)$.

5.3   The Channel

5.3.1   The physical channel model

In this chapter, the channel model developed in [2] is used to characterize the frequency nonselective multiple-input-multiple-output (MIMO) wireless fading channels with multiple transmit and multiple receive antennas. We consider a wireless communication channel between the base station (BS) and a mobile unit (MU) as depicted in Fig. 5.1.
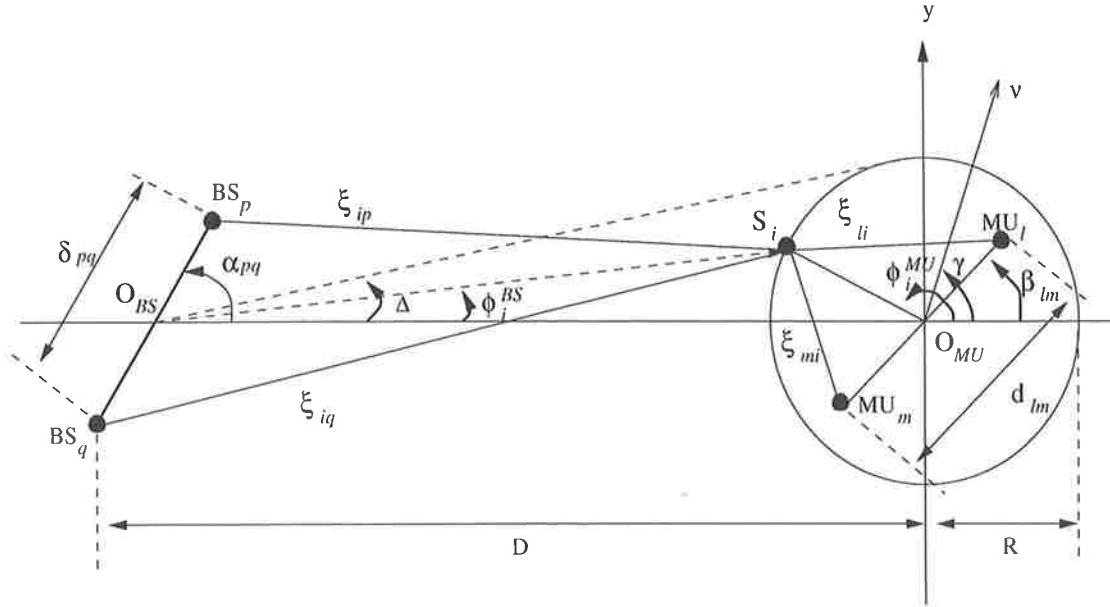
Figure 5.1: Channel model with local scatterers around MU

This multiple antennas configuration is initially proposed in [86] to investigate the effect of the spatial correlation on the capacity of multielement antenna systems. Subsequently, it has been extended in [2] to include the effect of the motion at the MU. We consider only the Rayleigh fading channels in this work. Hence, the fading channel model described in this section is a special case of that presented in [2]. The BS has no local scatterers as it is typically situated well above the city buildings, while the MU is immersed in a complex scattering environment. The BS receives signals through a narrow beamwidth $\Delta = \arctan(R/D)$ where D is the distance between the BS and the MU, and R is the radius of the scatterer ring. Let the BS be equipped with n_{BS} antenna elements and the MU with n_{MU} antenna elements. Let $1 \leq p, q \leq n_{BS}$ and $1 \leq l, m \leq n_{MU}$ denote the indices of the BS antennas and MU antennas respectively.

Suppose there are N independent scatterers, S_i, around the MU. The normalized complex path gain $h_{p,l}(t)$ connecting the BS antenna element p (BS_p) and the MU

antenna element l (MU_l) can be described mathematically by

$$h_{p,l}(t) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i \times \exp\left\{ j\psi_i - \frac{j2\pi}{\lambda}[\xi_{ip} + \xi_{li}] + j2\pi f_D[\cos(\phi_i^{MU} - \gamma)]t \right\} , \quad (5.4)$$

where g_i is the amplitude of the wave scattered by the ith scatterer toward the MU such that $N^{-1} \sum_{i=1}^{N} E[g_i^2] = 1$ as $N \to \infty$. Thus, in the limit case the fading gain $h_{p,l}(t)$ constructed according to the model represents a Rayleigh fading. The phase shift introduced by the ith scatterer is denoted by ψ_i. It is assumed that $\{\psi_i\}_{i=1}^{N}$ are i.i.d. random variable and is assumed to be uniformly distributed over $[-\pi, \pi)$. The maximum Doppler shift is denoted by $f_D = \nu/\lambda$ where ν is the speed of the MU. The direction of motion of the MU is γ. ξ_{ip} and ξ_{li} are the distances from the ith scatterer to BS_p and MU_l respectively. These distances depend on the angle of arrival (AOA) ϕ_i^{MU} of the traveling wave from the ith scatterer toward the MU . By assuming that $D \gg R \gg \max(\delta_{pq}, d_{lm})$, which is applicable for many practical cases of interest, it has been shown in [2] that the cross correlation between the gains of two arbitrary communication link $h_{p,l}(t)$ and $h_{q,m}(t)$ can be expressed as

$$\rho_{pl,qm}(\tau) = E[h_{p,l}(t) h_{q,m}^*(t+\tau)] = \exp\left\{ \frac{j2\pi \delta_{pq} \cos(\alpha_{pq})}{\lambda} \right\}$$
$$\times \int_{-\pi}^{\pi} \exp\{ \frac{j2\pi}{\lambda}[\delta_{pq}\Delta \sin(\alpha_{pq}) \ \sin(\phi^{MU}) + d_{lm} \cos(\phi^{MU} - \beta_{lm})]$$
$$- j2\pi f_D\tau[\cos(\phi^{MU} - \gamma)]\} f(\phi^{MU}) d\phi^{MU} , \quad (5.5)$$

where $f(\phi^{MU})$ is the pdf of the AOA seen by the MU. It is often assumed that the MU receives signals from all direction with equal probability, i.e. the AOA seen by the user is uniformly distributed over $[-\pi, \pi)$ with $f(\phi^{MU}) = 1/(2\pi)$ [20, 51, 86]. However, empirical measurements in [37] have shown that the AOA distribution are likely to be nonuniform. In [1], the von Mises angular distribution has been used for modeling the nonuniform AOA distribution at the MU and the validity for the use of the von Mises angular distribution has been verified with the measured data.

The von Mises angular distribution pdf is given by

$$f(\phi^{MU}) = \frac{\exp[\kappa \cos(\phi^{MU} - \mu)]}{2\pi I_0(\kappa)}, \quad \phi^{MU} \in [-\pi, \pi) \tag{5.6}$$

where $I_0(.)$ is the zeroth-order modified Bessel function, $\mu \in [-\pi, \pi)$ accounts for the mean direction of AOA seen by the user, and $\kappa \geq 0$ controls the spread of the AOA. For $\kappa \geq 3$, the spread of the AOA of the scatter components at MU is roughly equal to $2/\sqrt{\kappa}$. Using this distribution, a closed-form and mathematically tractable expression for the space-time cross-correlation between the two links, $h_{p,l}(t)$ and $h_{q,m}(t)$ of the MIMO wireless fading channel has been derived in [2]

$$\begin{aligned}
\rho_{pl,qm}(\tau) = {} & \frac{\exp[jc_{pq}\cos(\alpha_{pq})]}{I_0(\kappa)} I_0(\{\kappa^2 - a^2 - b_{lm}^2 - c_{pq}^2\Delta^2\sin^2(\alpha_{pq}) \\
& + 2ab_{lm}\cos(\beta_{lm} - \gamma) + 2c_{pq}\Delta sin(\alpha_{pq})[a\sin(\gamma) - b_{lm}\sin(\beta_{lm})] \\
& - j2\kappa[a\cos(\mu - \gamma) - b_{lm}\cos(\mu - \beta_{lm}) - c_{pq}\Delta\sin(\alpha_{pq})\sin(\mu)]\}^{1/2}) \,, \tag{5.7}
\end{aligned}$$

where $a = 2\pi f_D\tau$, $b_{lm} = 2\pi d_{lm}/\lambda$ and $c_{pq} = 2\pi\delta_{pq}/\lambda$.

5.3.2 Autoregressive Channel Model

It has been shown in [88, 103] that the time varying Rayleigh fading channels can be well described by a hidden Markov model. As has been done in [25, 95] for channel equalization problems, the time varying fading channels can be modeled by a multichannel autoregressive (AR) process of order P

$$\mathbf{h}(t) = \sum_{p=1}^{P} \mathbf{A}(p)\, \mathbf{h}(t - p) + \mathbf{v}(t) \,, \tag{5.8}$$

where the matrices $\mathbf{A}(p)$ are the time invariant matrices which are determined from the second order statistic of the fading channels and $\mathbf{v}(t)$ is a zero mean complex white Gaussian noise process with the correlation matrix $E[\mathbf{v}(t)\mathbf{v}^H(t+\tau)] = \mathbf{Q}\delta(\tau)$. By postmultiplying (5.8) with $\mathbf{h}^H(t - \tau)$ and taking the expectation on both sides

we arrive at [78]

$$\mathbf{R}(-\tau) = \sum_{p=1}^{P} \mathbf{A}(p)\mathbf{R}(p - \tau) + \mathbf{Q}\delta(\tau), \quad \tau = 0, ..., P \qquad (5.9)$$

where $\mathbf{R}(\tau) = E[\mathbf{h}(t)\mathbf{h}^H(t + \tau)]$ is the cross-correlation matrix. Provided that the correlation matrices $\mathbf{R}(\tau)$ is available, the matrices $\mathbf{A}(p)$ and $\mathbf{Q}$ can then be obtained by solving the above equations. This technique for determining the AR coefficients is similar to [11, 22] and those in [25, 95] for channel equalization problems. In this chapter, the proposed receiver uses Kalman filters for estimating the fading channel. Thus, in order to keep the complexity of the receiver at a reasonable level, we use the first order AR process to model the fading channels. For $P = 1$, the channel fading process has an AR representation:

$$\mathbf{h}(t) = \mathbf{A}\,\mathbf{h}(t - 1) + \mathbf{v}(t) \; . \qquad (5.10)$$

The matrices $\mathbf{A}$ and $\mathbf{Q}$ are evaluated by solving (5.9) for $\tau = 0, 1$

$$\mathbf{A} = \mathbf{R}(-1)\,\mathbf{R}^{-1}(0) \; ,$$

$$\mathbf{Q} = \mathbf{R}(0) - \mathbf{A}\mathbf{R}(1) \; . \qquad (5.11)$$

Let the BS and the MU assume the roles of the transmitter and receiver respectively. Thus, $n_T = n_{BS}$ and $n_R = n_{MU}$. With this role assignment, $[\mathbf{R}(\tau)]_{e,f} = \rho_{pl,qm}(\tau)$ is the cross-correlation between two fading gains $h_{p,l}(t)$ and $h_{q,m}(t + \tau)$ where $1 \leq p, q \leq n_T$ and $1 \leq l, m \leq n_R$. Due to the arrangement of $\mathbf{h}(t)$ in (5.3), the indices $e = p+(l-1)n_T$ and $f = q+(m-1)n_T$ where $1 \leq e, f \leq n_T n_R$. The cross-correlation term $\rho_{pl,qm}(\tau)$ is computed using (5.7).

For the spatially independent channel, the channel autocorrelation is then given by the Jake's formula $\rho_{pl,pl}(\tau) = E[h_{pl}(t)h_{pl}^*(t + \tau)] = J_0(2\pi f_D T_s \tau)$ where $J_0(.)$ is the zeroth-order Bessel function of the first kind as in [23, 24, 57, 60, 106]. The

matrices $\mathbf{A}$ and $\mathbf{Q}$ are then given by

$$\mathbf{A} = J_0(2\,\pi\,f_D\,T_s)\,\mathbf{I} \tag{5.12}$$

$$\mathbf{Q} = \left(1 - J_0^2(2\,\pi\,f_D\,T_s)\right)\mathbf{I}\,. \tag{5.13}$$

This turns out to be the same approximation method as used in [57, 60] for MIMO systems.

5.4 Receiver Structures

Equations (5.3) and (5.10) can be seen as the state-space equations which describe a linear time varying system. The state of the system is the gain of the fading channels and $\mathbf{A}$ is the state transition matrix. Given $\mathbf{A}$, $\mathbf{Q}$ and $\chi(t)$, a minimum variance unbiased estimate of the channel gains can be obtained recursively via a Kalman filter [9]. However, the matrix $\chi(t)$ is the transmitted coded symbols by n_T transmit antennas at time t and is therefore unknown to the receiver. Thus, this problem can be posed as the channel estimation with unknown signal model, where the unknown model parameter $\chi(t)$ belongs to a known discrete set. This problem can also be viewed as joint channel estimation and decoding of the space time trellis code in time-varying fading channels. At the receiver, the matrices $\mathbf{A}$ and $\mathbf{Q}$ can be estimated from the received signals using the technique described in [95] for the SISO case. In this work, we assume that these matrices are known at the receiver and focus on the problem of joint decoding and channel tracking.

5.4.1 Maximum Likelihood Sequence Estimation (MLSE)

When the ideal channel state information CSI $\mathbf{h}(t), t = 1, \ldots, T$ are available, the problem of decoding the space time trellis codes in the time-varying fading channels can be easily achieved by using the Viterbi algorithm which minimizes the following

metric

$$J_m^{ML} = \sum_{t=1}^{T} ||\mathbf{r}(t) - \sqrt{E_s}\, \chi_m(t)\, \mathbf{h}(t)||^2 \,, \tag{5.14}$$

where the subscript m denotes the mth hypothesized sequence $\{\mathbf{x}_m(t)\}_{t=1}^{T}$ being transmitted. Since the perfect CSI is not available at the receiver, we need to perform joint channel estimation and decoding. Assuming each sequence is transmitted with equal probability, this problem can be solved by finding the sequence $\{\mathbf{x}_m(t)\}_{t=1}^{T}$ which maximizes the likelihood function $p(\mathbf{r}(1), \cdots, \mathbf{r}(T) \mid \{\mathbf{x}_m(t)\}_{t=1}^{T})$. This can be achieved by using a bank of Kalman filters to track the time-varying channels, one for each of the possible sequence. Given a sequence of the received signals $\{\mathbf{r}(t)\}_{t=1}^{T}$, the channel estimate and its associated error covariance for the sequence $\{\mathbf{x}_m(t)\}_{t=1}^{T}$ can be computed recursively via the measurement and time update equations of the Kalman filter modeled to that sequence:

Measurement Update Equations:

$$\hat{\mathbf{h}}_m(t|t) = \hat{\mathbf{h}}_m(t|t-1) + \mathbf{G}_m(t)\left(\mathbf{r}(t) - \sqrt{E_s}\chi_m(t)\, \hat{\mathbf{h}}_m(t|t-1)\right)$$

$$\mathbf{G}_m(t) = \sqrt{E_s}\mathbf{\Sigma}_m(t|t-1)\chi_m^H(t)\left(E_s\chi_m(t)\, \mathbf{\Sigma}_m(t|t-1)\, \chi_m^H(t) + N_0 I_{n_R}\right)^{-1}$$

$$\mathbf{\Sigma}_m(t|t) = \mathbf{\Sigma}_m(t|t-1) - \sqrt{E_s}\mathbf{G}_m(t)\chi_m(t)\, \mathbf{\Sigma}_m(t|t-1) \,. \tag{5.15}$$

Time Update Equations:

$$\hat{\mathbf{h}}_m(t+1|t) = \mathbf{A}\, \hat{\mathbf{h}}_m(t|t)$$

$$\mathbf{\Sigma}_m(t+1|t) = \mathbf{A}\, \mathbf{\Sigma}_m(t|t)\mathbf{A}^H + \mathbf{Q} \,. \tag{5.16}$$

Using these channel estimates, the receiver recursively computes the log likelihood metric for the hypothesized sequence m according to

$$J_m^{MLSE} = -\sum_{t=1}^{T} \epsilon_m^H(t)\, \mathbf{\Omega}_m^{-1}(t)\epsilon_m(t) + \log\, \det \mathbf{\Omega}_m(t) \,, \tag{5.17}$$

where $\epsilon_m(t)$ and $\boldsymbol{\Omega}_m(t)$ are the innovation sequence and its associated covariance for the Kalman filter tuned to the hypothesized sequence m:

$$\epsilon_m(t) = \mathbf{r}(t) - \sqrt{E_s}\chi_m(t)\ \hat{\mathbf{h}}_m(t|t-1)\ , \tag{5.18}$$

$$\boldsymbol{\Omega}_m(t) = E_s\chi_m(t)\ \boldsymbol{\Sigma}_m(t|t-1)\ \chi_m^H(t) + N_0 I_{n_R}\ . \tag{5.19}$$

The receiver then chooses the sequence that maximises the log likelihood metric J_m^{MLSE}.

5.4.2   A Per-Survivor Processing (PSP) Approach

For a space time trellis code with the rate of R bits/sec/Hz, there are 2^{RT} possible sequences $\{\mathbf{x}_m(t)\}_{t=1}^T$. Hence we would need a bank of 2^{RT} Kalman filters to implement the MLSE receiver. This direct implementation of this optimal receiver is prohibited to be used in practical applications as the frame length T is generally large. In recent years, per-survivor processing [79] with Kalman filtering has been seen as an attractive approach to performing maximum likelihood sequence estimation over mobile radio channels that are rapidly time-varying [25, 74] (with $n_T = n_R = 1$). In [72, 73], we have extended these methods to perform joint channel estimation and decoding of block codes in general diversity time-varying correlated fading channels. In this chapter, we extends this work further to perform joint channel estimation and space-time trellis decoding by incorporating the per-survivor processing and Kalman filtering into the Viterbi algorithm.

Let S denotes the total number of states in the trellis. This PSP receiver works as follow: At the symbol interval $t-1$, we assume that each state, $\mu_s(t-1)$ for $s = 1, ..., S$, retains a survivor path. Associated with this survivor path are the channel estimate $\hat{\mathbf{h}}_{\mu_s}(t-1|t-1)$ and the cumulative log likelihood path metric $\Gamma_{\mu_s}(t-1)$. At the tth symbol interval, for all valid transitions to state $\mu_y(t)$, $\mu_s(t-1) \to \mu_y(t)$,

Time = t-1   Time = t
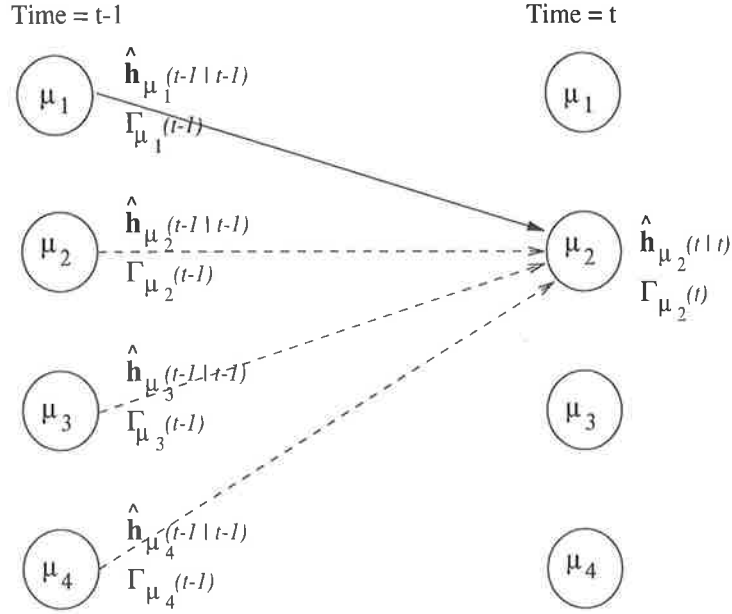
Figure 5.2: PSP receiver with Viterbi algorithm

the following branch metric is computed:

$$\Lambda\left(\mu_s(t-1) \to \mu_y(t)\right) = -\epsilon_{s,y}^H(t)\,\Omega_{s,y}^{-1}(t)\,\epsilon_{s,y}(t) - \log\det\Omega_{s,y}(t) \tag{5.20}$$

where

$$\epsilon_{s,y}(t) = \mathbf{r}(t) - \sqrt{E_s}\chi_{s,y}(t)\,\hat{\mathbf{h}}_{\mu_s}(t|t-1) \tag{5.21}$$

$$\Omega_{s,y}(t) = E_s\chi_{s,y}(t)\,\Sigma_{\mu_s}(t|t-1)\,\chi_{s,y}^H(t) + N_0 I_{n_R} \tag{5.22}$$

and $\chi_{s,y}(t)$ are the coded symbols that are output by the encoder during the transition $\mu_s(t-1) \to \mu_y(t)$. The predicted channel estimate $\hat{\mathbf{h}}_{\mu_s}(t|t-1)$ and its error covariance is computed using the time update equations in (5.16).

Once the branch metric of all valid transitions $\mu_s(t-1) \to \mu_y(t)$ is computed, the survivor path sequence to the state $\mu_y(t)$ and its cumulative path metric $\Gamma_{\mu_y}(t)$ are determined by performing a maximization over the previous valid states $\mu_s(t-1)$

$$\Gamma_{\mu_y}(t) = \max_{\mu_s(t-1)}\left[\,\Gamma_{\mu_s}(t-1) + \Lambda\left(\mu_s(t-1) \to \mu_y(t)\right)\,\right] \tag{5.23}$$

The channel estimate associate with this survivor path is then updated by a Kalman filter tuned to the coded symbols for the survivor transition. Let the predecessor state of $\mu_y(t)$ which maximizes the cumulative path metric $\Gamma_{\mu_y}(t)$ be denoted by $\mu_{opt}(t-1)$. The channel estimate at time t of the survivor path at state $\mu_y(t)$ can be estimated by the Kalman filter tuned to $\chi_{opt,y}(t)$ (i.e. the coded symbols for the survivor transition $\mu_{opt}(t-1) \rightarrow \mu_y(t)$)

$$\hat{\mathbf{h}}_y(t|t) = \hat{\mathbf{h}}_{opt}(t|t-1) + \mathbf{G}_{opt,y}(t) \left(\mathbf{r}(t) - \sqrt{E_s}\chi_{opt,y}(t) \, \hat{\mathbf{h}}_{opt}(t|t-1) \right)$$

$$\mathbf{G}_{opt,y}(t) = \sqrt{E_s}\mathbf{\Sigma}_{opt}(t|t-1)\chi_{opt,y}^H(t) \left(E_s\chi_{opt,y}(t) \, \mathbf{\Sigma}_{opt}(t|t-1) \, \chi_{opt,y}^H(t) + N_0 I_{n_R} \right)^{-1}$$

$$\mathbf{\Sigma}_y(t|t) = \mathbf{\Sigma}_{opt}(t|t-1) - \sqrt{E_s}\mathbf{G}_{opt,y}(t)\chi_{opt,y}(t) \, \mathbf{\Sigma}_{opt}(t|t-1) \tag{5.24}$$

Thus, this receiver requires only one Kalman filter for each state in the trellis. A pictorial description of the per-survivor processing with Viterbi algorithm is given in Fig. 5.2, assuming the 4 states space-time trellis codes with the rate of 2 bits/sec/Hz.

5.5 Performance Evaluation

The performance of the proposed PSP receiver is evaluated by using computer simulations. We consider the downlink of a single user system where the BS is equipped with two transmit antennas and the MU has two receive antennas. The four-states QPSK space time trellis codes in [89] is adopted. This space time code achieves a bandwidth efficiency of 2 bits/sec/Hz. We assume information is transmitted at the rate of 20 Ksps at the carrier frequency of 2.4 GHz. The information is transmitted in frame of 130 coded symbols per transmit antenna. At the beginning of each frame, each antenna transmits an orthogonal sequence (with length of 2 symbols). These orthogonal pilot sequences are used by the receiver to obtain the initial fading channel estimate to initialize the Kalman filters. We assume perfect time and frequency synchronization, and that the received signals are sampled at the symbol rate T_s. The fading channels used for the simulation will be generated according to (5.10).

The matrix $\mathbf{A}$ and $\mathbf{Q}$ are determined by the space-time correlation (5.7) with the following physical parameters: The transmit antennas at the BS are separated by a distance of 10λ while the receive antennas of the MU are separated by 0.5λ. With the assumed carrier frequency, these antenna spacings correspond to 1.25 metres and 6.25 centimetres respectively. The MU is moving in the direction $\gamma = 45°$. For the macrocells in urban, suburban and rural areas, the angular spread Δ at the BS is often less than 15° and in some cases less than 5°. We choose $\Delta = 10°$ in this simulation. The angle spread at the MU is set to $2/\sqrt{\kappa} \approx 66°$ (i.e. $\kappa = 3$) with the mean AOA $\mu = 180°$. The antenna array at the BS is assumed to be in parallel position to the MU antenna array (i.e. $\alpha_{1,2} = 90°$, $\beta_{1,2} = 90°$).

We examine the performance of the PSP receiver for two scenarios. The first scenario corresponds to moderate time-varying channels where the MU is moving at the speed of 9 km/h. With this speed, the fading rate $f_D T_s = 0.001$. The second scenario is when the speed of the MU is 90 km/h. This scenario corresponds to fast time-varying channels with the fading rate $f_D T_s = 0.01$. We plot the frame error rate (FER) of the PSP receiver for different values of signal to noise ratio (SNR). We define the SNR as the ratio between the total power transmitted and the total noise power per receive antenna, i.e. $SNR = n_T E_s/N_0$. In addition, we compare the performance of the PSP receiver with that of the maximum likelihood (ML) receiver which has perfect knowledge of the channels. For a fairer performance comparison with the double differential space-time coding scheme proposed in [61] which offers robust means of handling time-selectivity of the channels, we also examine the bit error rate (BER) of the proposed PSP receiver. Fig. 5.3 shows the FER and BER of the PSP and the ML receivers for $f_D T_s = 0.001$. At the $FER = 10^{-2}$, the FER performance of the PSP receiver is within 0.1 dB of the ML receiver with ideal CSI. Simulation results of the BER for this system however show a larger gap between the performance of the PSP receiver and the ML receiver with perfect CSI. This effect
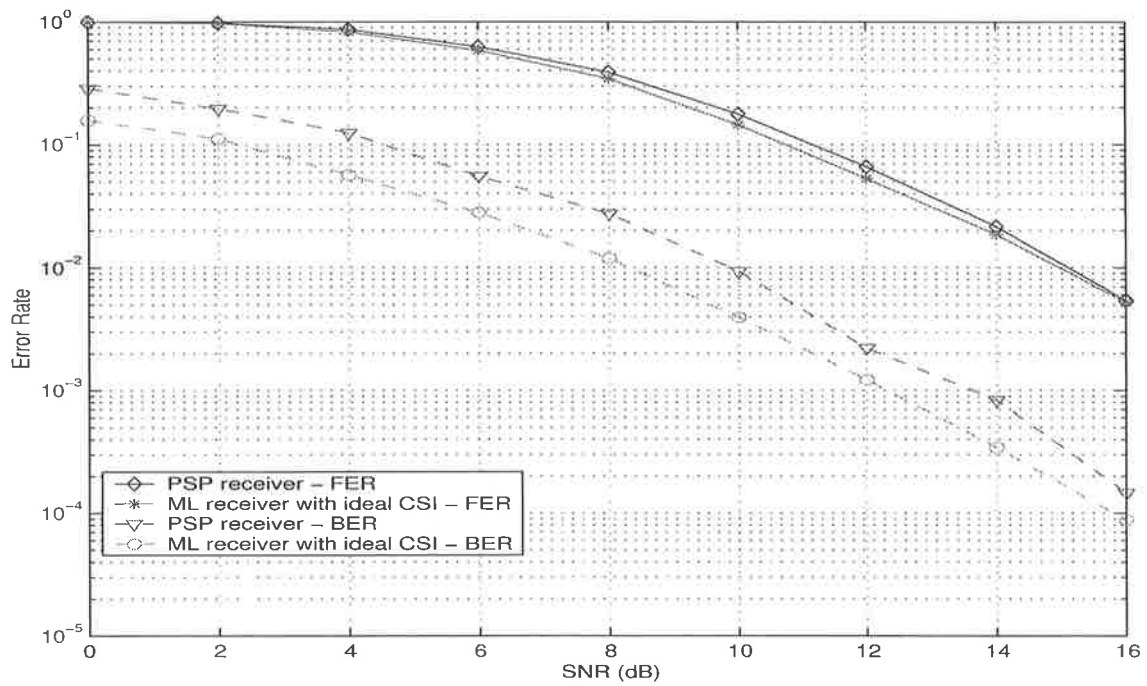
Figure 5.3: Performance of the PSP receiver for $f_D T_s = 0.001$
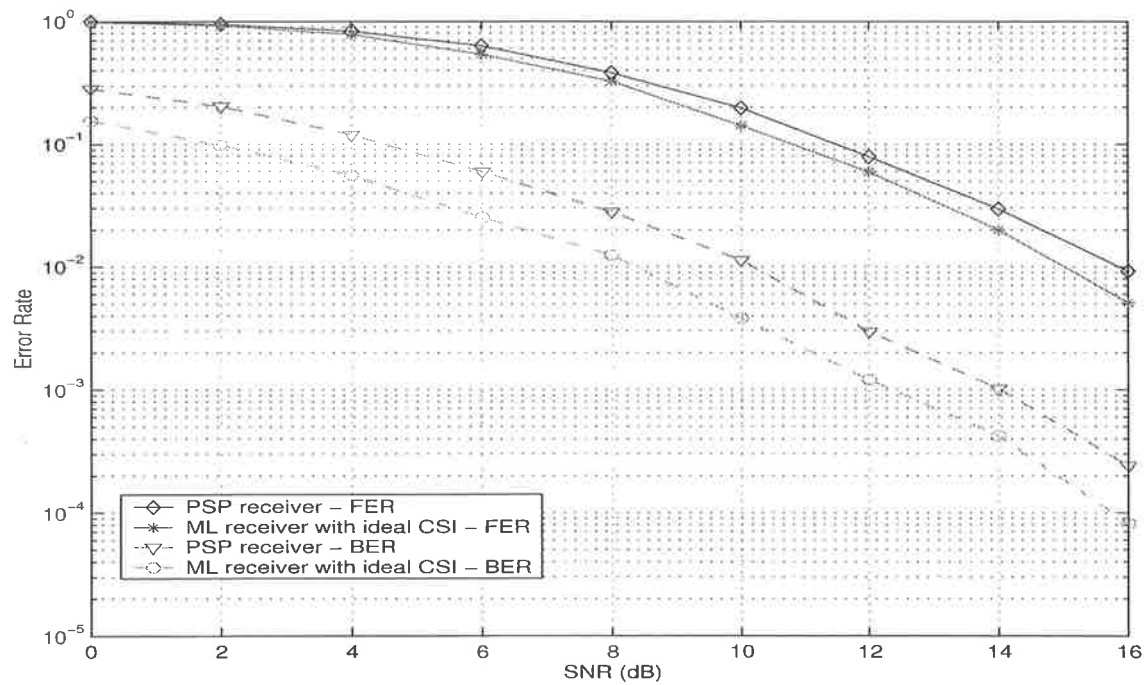
Figure 5.4: Performance of the PSP receiver for $f_D T_s = 0.01$

has been also observed in [23]. An explanation for this is that when there is an error occurred, it causes error propagation which results in more errors in the subsequence symbols of that frame. However this gap is still small when compared to the double differential scheme which has a 6 dB performance loss. At the $BER = 10^{-3}$, the proposed PSP receiver is only 1.2 dB worse than the ML receiver with perfect CSI. Fig. 5.4 shows the FER and BER of the PSP and ML receivers for the fast time-varying channels with fading rate $f_D T_s = 0.01$. As expected, the performance gap between the PSP receiver and the ML receiver is larger for this case as compared to the moderate time-varying channels case. This is due to the increase in the amount of uncertainty in the predicted channel estimate. However, the increase in performance gap is not significant. At the $FER = 10^{-2}$ and $BER = 10^{-3}$, the respective FER and BER performance gaps between the proposed PSP receiver and the ML receiver with perfect CSI are 0.8 dB and 1.6 dB.

We also examine the effect of the spatial correlation in the channels on the performance of the PSP receiver which assumes an i.i.d fading channel model. The space-time correlated fading channels used for the simulations are generated using the setup described previously. We compare the performance of the PSP receiver which adopts the i.i.d fading channel model to one which adopts the true channel model. Fig. 5.5 and Fig. 5.6 show the error performance degradation that results when the i.i.d. channel model is adopted by the PSP receiver for moderately and fast time-varying cases, respectively. For moderately time-varying channels, the performance of the PSP receiver which assumes the i.i.d fading channel model degrades by 1 dB as compared to the one with true channel model. However under fast time-varying channels, this channel model mismatch due to the spatial correlation causes large channel estimation errors at the receiver and results in large performance degradation. This result indicates that under fast time-varying channels, it is very important for the receiver to account for such spatial correlation in the
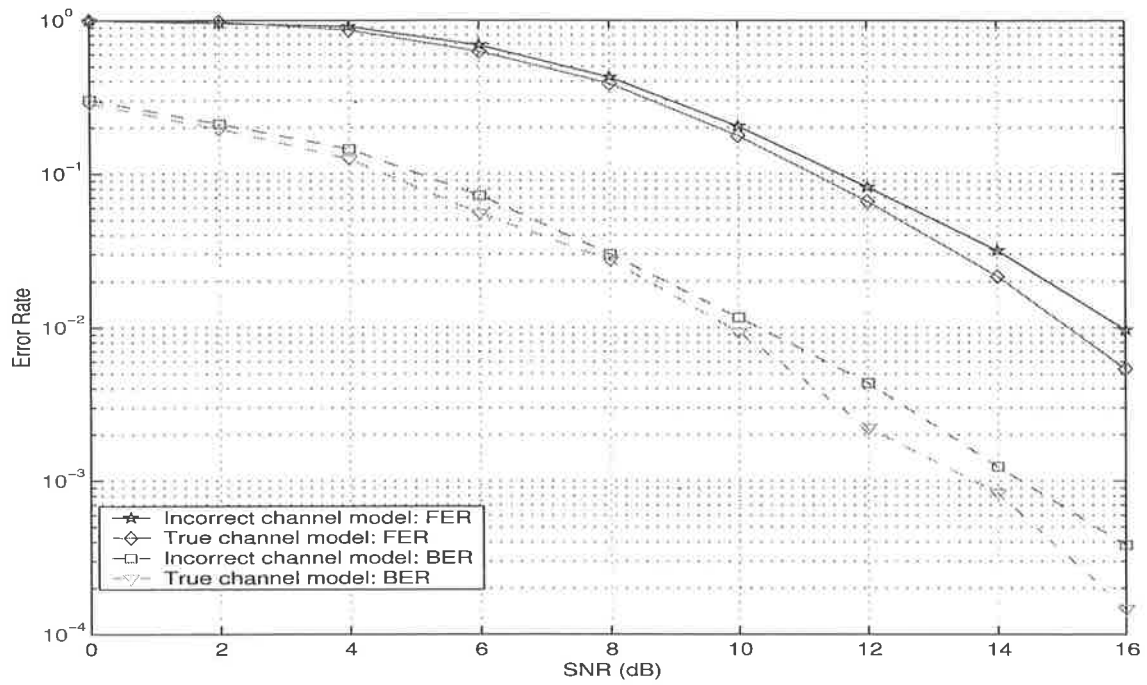
Figure 5.5: Receiver's performance with incorrect channel model $f_D T_s = 0.001$
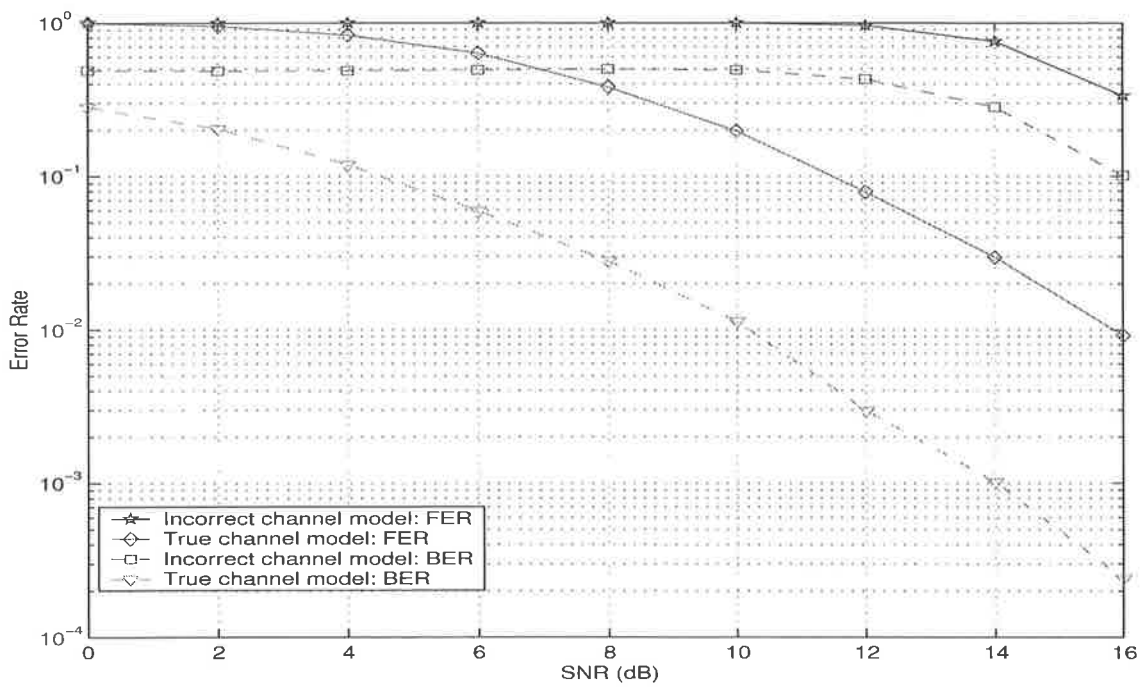
Figure 5.6: Receiver's performance with incorrect channel model $f_D T_s = 0.01$

channels.

In the previous simulations, we use the AR channel model to generate the fading channels. We now examine the performance of this PSP receiver under realistic situations. Let's define the following $n_T \, n_R T_T \times n_T \, n_R T_T$ matrix

$$\mathbf{R}_{T_T} = E[\mathbf{h h}^H] = \begin{bmatrix} \mathbf{R}(0) & \mathbf{R}(1) & \cdots & \mathbf{R}(T_T - 1) \\ \mathbf{R}(-1) & \mathbf{R}(0) & \cdots & \mathbf{R}(T_T - 2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}(-T_T + 1) & \mathbf{R}(-T_T + 2) & \cdots & \mathbf{R}(0) \end{bmatrix} \tag{5.25}$$

$$\mathbf{R}(\tau) = E[\mathbf{h}(t)\mathbf{h}^H(t + \tau)] \tag{5.26}$$

where $\mathbf{h} = [\mathbf{h}^T(1) \; \mathbf{h}^T(2) \; \cdots \mathbf{h}^T(T_T)]^T$, $T_T = T + T_p$ is the total frame length, T_p is the length of the pilot sequence, and $[\mathbf{R}(\tau)]_{e,f} = \rho_{pl,qm}(\tau)$ is the cross-correlation between two fading gains $h_{p,l}(t)$ and $h_{q,m}(t + \tau)$ given by (5.7). The fading channels with the cross-correlation (5.7) can then be generated by having

$$\mathbf{h} = \mathbf{R}_{T_T}^{1/2} \, \mathbf{h}_w \tag{5.27}$$

where $\mathbf{h}_w$ is an $n_T \, n_R \, T_T$ zero mean i.i.d complex vector with covariance $E[\mathbf{h}_w \, \mathbf{h}_w^H] = \mathbf{I}$. For a 2 transmit and 2 receive antennas system, if a frame contains 130 coded symbols per transmit antennas, the size of the matrix $\mathbf{R}$ will then be $4(130 + T_p) \times 4(130 + T_p)$. Due to the complexity involved in finding the square root of such large matrix, the size of the frame is now reduced so that it contains only 25 coded symbols per transmit antenna. An orthogonal pilot sequence (with length of 2 symbols) is still used so that the receiver can obtain the initial channel estimate. At the receiver, we approximate the fading channel with a first order AR process (5.10). The matrices $\mathbf{A}$ and $\mathbf{Q}$ required by the Kalman filters for estimating the fading channel are given by (5.11).

Fig. 5.7 compares the performance of the PSP receiver with that of the ML receiver which has ideal CSI at the fading rate $f_D T_s = 0.001$. At the $FER = 10^{-2}$,

the PSP receiver is about 0.8 dB worse than the ML receiver with perfect CSI. This performance gap is slightly larger than the previous case when the AR model is used for generating the fading channel. The BER performance gap, however, is about the same as the previous case. At the $BER = 10^{-3}$, the performance gap between the two receivers is about 1.2dB. The performance comparison of these two receivers under fast time-varying channels with fading rate $f_D T_s = 0.01$ is shown in Fig. 5.8. At the $FER = 10^{-2}$ and $BER = 10^{-3}$, the performance gaps between these two receivers are 2.0 dB and 2.7 dB, respectively. These gaps are again only slightly larger than the corresponding case when the fading channel is generated using the AR model. This suggests that the PSP receiver is quite robust to modelling error and that the first order AR model can provide a reasonably accurate approximation of the time-varying channel at these fading rate. Using higher order AR model therefore can only increase the performance of the receiver slightly, albeit at the cost of increased receiver complexity. For this reason, we will not investigate the performance of the PSP receiver with higher order AR model.

5.6   Conclusions

This chapter proposes a PSP based receiver for joint space-time trellis decoding and channel estimation in spatially correlated time-varying fading channels. The physical channel model of the MIMO wireless fading channel with multiple transmit multiple receive antennas is approximated by a statistical channel model. This statistical channel model permits Kalman filtering to be used for tracking the channel variations. Simulation results have suggested that the proposed PSP receiver with first order AR channel approximation performs very well. In addition, these simulation results also confirm the importance for the receiver to account for the spatial correlation in the channel.
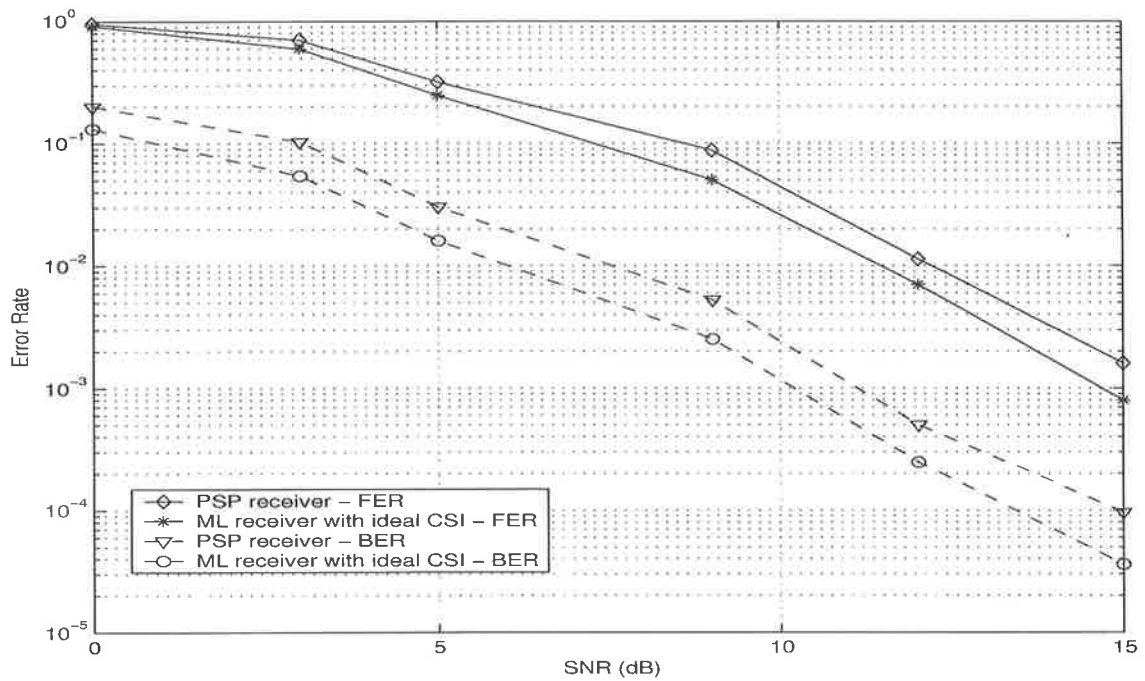
Figure 5.7: Receiver's performance under realistic fading channel $f_D T_s = 0.001$
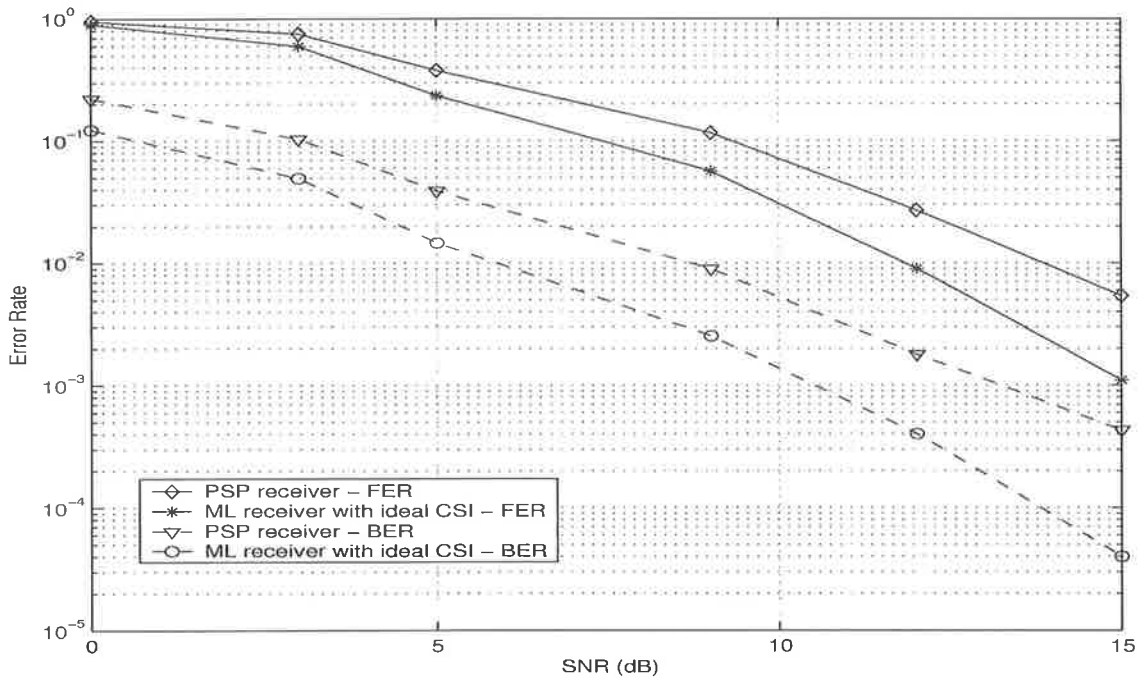
Figure 5.8: Receiver's performance under realistic fading channel $f_D T_s = 0.01$

Chapter 6

Recursive Receivers for General Diversity Channels

This chapter addresses the design and performance of time recursive receivers for diversity based communication systems with flat Rayleigh or Ricean fading. It introduces a general state-space model for such systems where there is temporal correlation in the channel gains. Such an approach encompasses a wide range of diversity systems such as spatial diversity, frequency diversity and code diversity systems which are used in practice. The chapter describes a number of noncoherent receiver structures derived from both sequence and *a posteriori* probability based cost functions and compare their performance using an orthogonal frequency division multiplex example. In this example, the chapter shows how a standard physical delay-doppler scattering channel model can be approximated by the proposed state-space model. Simulations are used to show that significant performance gains can be obtained by exploiting temporal as well as diversity channel correlations. The chapter argues that such time-recursive receivers offer some advantages over block processing schemes such as computational and memory requirement reductions and the easier incorporation of adaptivity in the receiver structures.

6.1   Introduction

In recent years there has been an increasing emphasis on exploitation of different types of diversity in communications systems, particularly those which utilise a wireless transmission medium. Examples of types of diversity which have been exploited to improve the performance of digital communications systems are time diversity (also known as fractional sampling) [52], spatial diversity via the use of several transmitter and/or receiver antennae [93, 36], code diversity (as used in spread spectrum communications), and frequency diversity such as used in orthogonal frequency division multiplexing (OFDM) systems. Also, these different types of diversity can also be combined, for example, a multicarrier code-division multiple access (CDMA) system. In these types of systems, the presence of a number of different diversity channels can be used to improve the system performance by transmitting and/or receiving the same information symbols simultaneously across a number of channels. This goal can be achieved by introducing coding in both the temporal and diversity dimensions. This chapter is concerned with the design of receivers for such *diversity systems.*

Given that there are many different types of diversity which may be exploited, it is desirable to develop a generalised and unified taxonomy for system modelling and signal processing for such systems. This framework can be used to derive various types of receiver structures which are applicable to any type of diversity system. In particular, we will consider systems where the channel is modelled by vector time series where each element of the vector represents the complex gain of a different diversity channel. We will assume that these gains are correlated in both the diversity dimension (ie the gains of different channels are statistically dependent at each time), and the temporal direction. Temporal correlation can arise from doppler effects [51] and from oversampling as examples. We particularly focus on state-space channel models which lead naturally to time recursive receiver structures. These

structures generally tend to offer reduced computational complexity and memory requirements when compared to block processing systems. Another potential advantage of time recursive receivers is the ease of development of adaptive processing algorithms, however we don't address this issue in this chapter.

Since the computational complexity of the optimal receiver is generally too large to be feasible, we focus on the derivation and comparison of a number of sub-optimal receivers which have computational complexity which is linear in the size of the transmission codebook. In this chapter, we will tend to focus on single-user systems where the general aim of the receiver is to exploit the temporal and diversity correlation to improve performance. Much of what is considered here also applies to the multiuser case, where the diversity correlation leads to undesirable interference between users. The task of the receiver is then to mitigate this interference.

In this chapter, we address only flat fading channels. Related work which addresses, in a different way, the frequency selective space-time case, can be found in [85]. Recent work in the space-time coding context [14] utilises fading models where correlation between the diversity channels is included. Of course, such correlation is an inherent feature in CDMA systems, and the main focus of various multiuser detection schemes (see [33, 71] for an overview) is the mitigation of the detrimental effects of such correlation. However, there has been significantly less published concerning the presence of, and exploitation of temporal correlation in the channel. We believe that this work makes the following contributions :

- A generalised state-space model for diversity communication systems,

- Development of a number of time-recursive receiver structures which exploit temporal correlations in the channel, and

- A performance comparison of these proposed receivers.

In addition, we examine the issue of matching our proposed model to a physical cor-

relation model for delay-doppler channels [49], [50]. Some of the proposed receivers are well-known, but others are novel, especially in the context of their application in this problem.

There has been considerable interest also in the code design issue for diversity systems, particularly space-time diversity systems (see eg [14, 89, 44, 4, 104]). We don't directly address the design issue in this chapter, but we will make some pertinent observations regarding code design in the context of the OFDM example presented in section 6.4.

The layout of the chapter is as follows : In section 6.2, we introduce our model of a general diversity system in complex number form, and then in a real quadrature signal form. The latter will form the basis for the receiver design. The model for the received signal constitutes the measurement equation for a time varying state space system. Our model for the diversity path gains is cast as a quadrature Gauss-Markov process which constitutes the state equation for the state space system. We also give some examples, which are by no means exhaustive, of how some well-known examples of diversity systems may be cast within our framework. Some issues relating to partial diversity and an associated model reduction method are discussed. In section 6.3, we introduce a number of candidate noncoherent receiver structures which are based on the state-space diversity model from 6.2. In section 6.4, we detail a specific application of our techniques to the OFDM system. The diversity dimension here is frequency. We examine a general physical delay-doppler scattering model for path correlations as a function of time and frequency, and detail a technique which permits approximation of this correlation function by the signal model introduced in 6.2. Finally we present some simulation results for the OFDM case, which compare the performance of our candidate receivers to the coherent case (path gains known by the receiver), and to standard approaches which do not exploit the time correlation.

6.2 The Communication System Model

We consider a communication system consisting of $L \geq 1$ diversity paths. These diversity paths are assumed to undergo Rayleigh/Ricean flat fading. The path gains are correlated with each other and are also temporally correlated. At each time t, we map p bits of the input data stream onto one of the codewords of the set $\tilde{\Theta} = \{\tilde{\mathbf{X}}^{(1)}, \tilde{\mathbf{X}}^{(2)}, ..., \tilde{\mathbf{X}}^{(K)}\}$ where $\tilde{\mathbf{X}}^{(k)} \in \mathbb{C}^{M \times L}$ and $K = 2^p$. The codeword is then transmitted serially across the L channels, ie column ℓ of the selected code is transmitted serially across diversity path ℓ. We assume that the receiver measures a linear superposition of the L diversity paths. Thus the received signal in block of M complex samples (M is the temporal dimension) can be written as

$$\tilde{\mathbf{y}}_t = \tilde{\mathbf{X}}_t \tilde{\mathbf{h}}_t + \tilde{\mathbf{n}}_t \, , \tag{6.1}$$

where $\tilde{\mathbf{h}}_t \in \mathbb{C}^L$ is the channel gains and $\tilde{\mathbf{n}}_t \in \mathbb{C}^M$ is a zero mean, complex, circular, white Gaussian noise sequence with covariance $2\sigma^2 \mathbf{I}_M$.

6.2.1 Examples

We now give several examples of how this model may be applied to represent different diversity systems.

Space-time coded systems

By examining (4.4) and (5.3), it is evident that space-time coded systems in quasi-static fading channels or in time-varying fading channels can be easily cast into this general diversity system model (6.1).

Orthogonal frequency division multiplexing (OFDM)

Here the diversity dimension is frequency, so we can regard OFDM as a frequency-time coded system. Suppose we seek to transmit n complex symbols in a vector

$\mathbf{s}_t$ for time period t. We use a pre-coder matrix $\mathbf{C} \in \mathbb{C}^{L \times n}$ to map the n symbols onto the L orthogonal frequency channels. Here $M = L$. The signal transmitted on channel ℓ will be the ℓth element of the vector $\mathbf{C}\mathbf{s}_t$. The cyclic prefix which is added at the transmitter and stripped off at the receiver turns the time and frequency selective fading into a time fading on each carrier, see [49]. Thus the received signal for symbol period t is the superposition

$$[\tilde{\mathbf{y}}_t]_m = \sum_{\ell=0}^{L-1} [\tilde{\mathbf{h}}_t]_\ell \, [\mathbf{C}\,\mathbf{s}_t]_\ell \, e^{2\pi i m \ell / L} + [\tilde{\mathbf{n}}_t]_m \ , \tag{6.2}$$

where $[\tilde{\mathbf{h}}_t]_\ell$ denotes the complex gain of channel ℓ, and $\tilde{\mathbf{n}}_t$ is additive Gaussian white noise. Thus we have the form of (6.1) where the code matrix has the form $\tilde{\mathbf{X}}_t = \mathbf{F}^H \mathbf{S}_t$ where $\mathbf{S}_t$ is a $L \times L$ diagonal matrix with its diagonal elements being the entries of the vector $\mathbf{C}\,\mathbf{s}_t$, and $\mathbf{F}$ is the Fourier matrix $[\mathbf{F}]_{ml} = e^{-2\pi i m \ell / L}$. More details on OFDM as a diversity system are given in section 6.4.

Code division multiple access (CDMA) - uplink

This is an example of a multiuser diversity system. In this case, column ℓ of the code matrix is the spreading code for user ℓ multiplied by the complex modulation symbol for user ℓ at symbol time t. Thus if $\mathbf{W} \in \mathbb{C}^{M \times L}$ denotes a matrix of the length M spreading codes for active users, and $\mathbf{S}_t$ is a diagonal matrix of size $L \times L$ with the diagonal elements being the modulation symbols of the active users for symbol period t, then the associated diversity code is $\tilde{\mathbf{X}}_t = \mathbf{W}\,\mathbf{S}_t$. The path gains in this case are the propagation gains from each user mobile station to the base station. In this case, our receiver designs will constitute noncoherent multiuser detectors.

6.2.2   Real Quadrature Model Form

It will be convenient for the resulting receiver derivations, to utilise a real quadrature model for the signal. Let denote the real and imaginary part of an arbitrary complex

matrix (or vector) $\tilde{\theta}_t$ by θ_t^r and θ_t^i, respectively. The complex received signal (6.1) can be modelled into quadrature form by

$$y_t = \begin{bmatrix} \mathbf{y}_t^r \\ \mathbf{y}_t^i \end{bmatrix} = \mathbf{X}_t \ \mathbf{h}_t + \mathbf{n}_t \ , \tag{6.3}$$

where $\mathbf{y}_t \in \mathbb{R}^{2M}$, $\mathbf{h}_t \in \mathbb{R}^{2L}$ and $\mathbf{X}_t \in \mathbb{R}^{2M \times 2L}$ are given by

$$\mathbf{h}_t = \begin{bmatrix} \mathbf{h}_t^r \\ \mathbf{h}_t^i \end{bmatrix} \ , \quad \mathbf{n}_t = \begin{bmatrix} \mathbf{n}_t^r \\ \mathbf{n}_t^i \end{bmatrix} \ , \quad \mathbf{X}_t = \begin{bmatrix} \mathbf{X}_t^r & -\mathbf{X}_t^i \\ \mathbf{X}_t^i & \mathbf{X}_t^r \end{bmatrix} \ . \tag{6.4}$$

The transmitted codeword $\mathbf{X}_t$ is now selected from the corresponding set $\Theta = \{\mathbf{X}^{(1)}, \mathbf{X}^{(2)}, ..., \mathbf{X}^{(K)}\}$. Here $\mathbf{n}_t \in \mathbb{R}^{2M}$ is a zero mean, Gaussian white noise vector with covariance $\sigma^2 \mathbf{I}_{2M}$.

6.2.3 Time-Correlated Rayleigh/Ricean Flat Fading Channel Model

In this section, we define the model for the diversity path gains $\tilde{\mathbf{h}}_t$ in both complex, and real quadrature form. The channel gains are modelled as a first-order vector Gauss-Markov process expressed in complex form by

$$\tilde{\mathbf{h}}_{t+1} = \tilde{\mathbf{A}} \ \tilde{\mathbf{h}}_t + \tilde{\mathbf{v}}_t \ , \tag{6.5}$$

where $\tilde{\mathbf{A}} \in \mathbb{C}^{L \times L}$ is a known, stable matrix and $\tilde{\mathbf{v}}_t$ is a complex Gaussian white noise process with mean $\tilde{\mu}$ and covariance $\tilde{\mathbf{Q}}$. We assume that at the initial time $t = 0$ that $\tilde{\mathbf{h}}_0$ is chosen to be Gaussian with the steady state statistics

$$E\{\tilde{\mathbf{h}}_0\} = (\mathbf{I} - \tilde{\mathbf{A}})^{-1} \ \tilde{\mu} \ , \tag{6.6}$$

and $\text{Cov}\{\tilde{\mathbf{h}}_0\} = \tilde{\mathbf{P}}$ where $\tilde{\mathbf{P}}$ satisfies the Lyapunov equation $\tilde{\mathbf{A}}\tilde{\mathbf{P}}\tilde{\mathbf{A}}^H + \tilde{\mathbf{Q}} = \tilde{\mathbf{P}}$. Thus $\tilde{\mathbf{h}}_t$ will be a stationary process with mean given by (6.6) and with autocorrelation matrix

$$E\{\tilde{\mathbf{h}}_t \ \tilde{\mathbf{h}}_{t-\tau}^H\} = \begin{cases} \tilde{\mathbf{A}}^\tau \ \tilde{\mathbf{P}} & \tau \geq 0 \\ \tilde{\mathbf{P}} \ \left(\tilde{\mathbf{A}}^H\right)^{-\tau} & \tau < 0 \ . \end{cases} \tag{6.7}$$

This channel model is general and permits both temporal and diversity correlations in the gains. Several channel models considered in [14, 44, 60] and previous chapters for the space-time coded systems can be considered as special cases of this model. The quasi-static independent and identical distributed (i.i.d) fading model used in [44] can be represented by this general model with $\tilde{\mathbf{A}} = \mathbf{0}$ and $\tilde{\mathbf{P}} = \mathbf{I}$ where $\mathbf{0}$ is the matrix of all zeros. Setting $\tilde{\mathbf{A}} = \mathbf{0}$ and $\tilde{\mathbf{P}} = E\{\tilde{\mathbf{h}}_t \, \tilde{\mathbf{h}}_t^H\} \neq \mathbf{I}$, this model corresponds to the quasi-static correlated fading channel model in [14] and in chapter 4. The time-varying i.i.d fading channel model in [60] can be cast into this general model by having $\tilde{\mathbf{P}} = \mathbf{I}$ and $\tilde{\mathbf{A}} = \alpha\mathbf{I}$ where $\alpha < 1$. By having $\tilde{\mathbf{A}}$ and/or $\tilde{\mathbf{P}}$ not proportional to an identity matrix and $\tilde{\mathbf{A}} \neq \mathbf{0}$, we then have a time-varying correlated fading channel similar to that in chapter 5.

The analogous real quadrature form for (6.5) is

$$\mathbf{h}_{t+1} = \mathbf{A} \, \mathbf{h}_t + \mathbf{v}_t , \tag{6.8}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^r & -\mathbf{A}^i \\ \mathbf{A}^i & \mathbf{A}^r \end{bmatrix} , \qquad \mathbf{v}_t = \begin{bmatrix} \mathbf{v}_t^r \\ \mathbf{v}_t^i \end{bmatrix} . \tag{6.9}$$

The process $\mathbf{v}_t \in \mathbb{R}^{2L}$ is a Gaussian white noise process with mean μ and covariance matrix $\mathbf{Q}$ given by

$$\mu = \begin{bmatrix} \mu^r \\ \mu^i \end{bmatrix} , \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{rr} & \mathbf{Q}_{ri} \\ \mathbf{Q}_{ri}^T & \mathbf{Q}_{ii} \end{bmatrix} , \tag{6.10}$$

with $\mathbf{Q}_{rr} + \mathbf{Q}_{ii} = \mathbf{Q}^r$ and $\mathbf{Q}_{ri}^T - \mathbf{Q}_{ri} = \mathbf{Q}^i$. The mean and autocorrelation sequence for $\mathbf{h}_t$ are then given by

$$\nu = E\{\mathbf{h}_t\} = (\mathbf{I} - \mathbf{A})^{-1} \, \mu ,$$

$$E\{\mathbf{h}_t \, \mathbf{h}_{t-\tau}^T\} = \begin{cases} \mathbf{A}^\tau \, \mathbf{P} & \tau \geq 0 \\ \mathbf{P} \, (\mathbf{A}^T)^{-\tau} & \tau < 0 , \end{cases} \tag{6.11}$$

where

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{rr} & \mathbf{P}_{ri} \\ \mathbf{P}_{ri}^T & \mathbf{P}_{ii} \end{bmatrix}, \tag{6.12}$$

with $\mathbf{P}_{rr} + \mathbf{P}_{ii} = \mathbf{P}^r$ and $\mathbf{P}_{ri}^T - \mathbf{P}_{ri} = \mathbf{P}^i$. Equations (6.3) and (6.8) are respectively the observation and state equations for a state space model of our received signals. The mean and autocorrelations for h_t can be shown to be identical to the analogous means and autocorrelations of the complex sequence $\tilde{\mathbf{h}}_t$. Thus the magnitude gain of the channels is Ricean (if $\mu \neq \mathbf{0}$) or Rayleigh (if $\mu = \mathbf{0}$) with joint correlations in the diversity and time dimensions given by (6.11). Equation (6.11) shows that the autocorrelation sequence has the form of a mixture of decaying complex exponentials (determined by the eigenvalues of the stable matrix $\mathbf{A}$).

6.2.4 Model Reduction

In the general diversity model (6.1), we made the assumption that the covariance matrix $\tilde{\mathbf{P}}$ of the channel gains $\tilde{\mathbf{h}}_t$ was strictly positive definite. This is equivalent to the assumption that full diversity gain L is achieved. As observed in [49], for some systems the effective amount of diversity can be less than L. This is manifested by the resulting covariance matrix $\tilde{\mathbf{P}}$ not being strictly positive definite. In this case, we can reduce the dimension of the problem to $\overline{L}$ ($< L$), the *effective* diversity of the channel, which is given by the number of non-zero eigenvalues of $\tilde{\mathbf{P}}$. This permits a reduction in the computational complexity of the various receivers presented later.

We proceed as follows : Let $\tilde{\mathbf{P}} = \tilde{\mathbf{\Phi}} \, \mathbf{\Lambda} \, \tilde{\mathbf{\Phi}}^H$ denotes an eigenvalue decomposition, where $\tilde{\mathbf{\Phi}}$ is unitary and $\mathbf{\Lambda}$ is a diagonal matrix of the form

$$\mathbf{\Lambda} = \begin{bmatrix} \overline{\mathbf{\Lambda}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{6.13}$$

where $\overline{\mathbf{\Lambda}}$ is $\overline{L} \times \overline{L}$ and is strictly positive definite. Here $\overline{L}$ represents the true diversity of the model. Partition $\tilde{\mathbf{\Phi}} = [\overline{\mathbf{\Phi}} \; \hat{\mathbf{\Phi}}]$ conformally with $\mathbf{\Lambda}$, ie $\overline{\mathbf{\Phi}}$ is of size $L \times \overline{L}$. Let

$\overline{\mathbf{h}}_t = \overline{\mathbf{\Phi}}^H \tilde{\mathbf{h}}_t$, then if $E\{\tilde{\mathbf{h}}_t\} = 0$, the following model is equivalent to one described previously by (6.5) and (6.1), however, with a reduced state dimension as shown in the Appendix:

$$\begin{aligned}
\overline{\mathbf{h}}_{t+1} &= \overline{\mathbf{A}}\,\overline{\mathbf{h}}_t + \overline{\mathbf{v}}_t \\
\tilde{\mathbf{y}}_t &= \overline{\mathbf{X}}_t\,\overline{\mathbf{h}}_t + \tilde{\mathbf{n}}_t \,,
\end{aligned} \qquad (6.14)$$

where

$$\begin{aligned}
\overline{\mathbf{A}} &= \overline{\mathbf{\Phi}}^H \mathbf{A}\,\overline{\mathbf{\Phi}} \\
\overline{\mathbf{v}}_t &= \overline{\mathbf{\Phi}}^H \tilde{\mathbf{v}}_t \\
\overline{\mathbf{X}}_t &= \mathbf{X}_t\,\overline{\mathbf{\Phi}} \,.
\end{aligned} \qquad (6.15)$$

6.3 The Receivers

In this section, we describe a number of noncoherent receiver structures for estimating the code sequence $\mathbf{X}_t$ in the absence of knowledge of the channel gains. The first class of receivers, which includes the Generalised Likelihood Ratio Test (GLRT)[14, 104, 101], Noncoherent Maximum Likelihood (NCML) [14, 44] and MAP receivers, are well-known and do not exploit the temporal correlation in the signal. The NCML and MAP receivers however do exploit the correlation in the diversity dimension. The second class are Kalman filter based receivers which includes the Per-survivor processing (PSP), M-algorithm, A posteriori probability (APP) and iterative Expected-Maximisation (EM) receivers. These receivers are based on the state space model from section 6.2. The PSP, APP, M-algorithm and EM algorithms are all well-known. However, the novelty of these receivers is that we incorporate Kalman filtering into these standard algorithms to exploit the temporal correlation in the signals, and track the channel variations.

6.3.1 Generalised Likelihood Ratio Test (GLRT) Receiver

When the probability distribution of the diversity path gains is unknown, a GLRT receiver can be used. The GLRT receiver computes the joint maximum likelihood estimate of the channel and the transmitted codeword. Thus, the transmitted codeword estimated by the GLRT receiver is

$$\hat{\mathbf{X}}_t^{GLRT} = \arg\max_{\mathbf{X}\in\Theta} \sup_{\mathbf{h}_t} p(\mathbf{y}_t \,|\, \mathbf{X}_t = \mathbf{X}, \,\mathbf{h}_t) \;, \tag{6.16}$$

where $p(\mathbf{y}_t \,|\, \mathbf{X}_t, \,\mathbf{h}_t) = \mathcal{N}(\mathbf{y}_t - \mathbf{X}_t\mathbf{h}_t, \sigma^2\mathbf{I}_M)$ is the data likelihood function, which is a Gaussian with mean $\mathbf{X}_t\,\mathbf{h}_t$ and covariance $\sigma^2\mathbf{I}$. Here, $\mathcal{N}(\mathbf{x}, \mathbf{P})$ denotes the Gaussian density

$$\mathcal{N}(\mathbf{x}, \mathbf{P}) = \frac{1}{(2\pi)^{N/2}\,(\det\mathbf{P})^{1/2}}\, e^{-\mathbf{x}^T\mathbf{P}^{-1}\mathbf{x}/2} \;, \tag{6.17}$$

where N is the dimension of the vector $\mathbf{x}$. Now, maximising $p(\mathbf{y}_t \,|\, \mathbf{X}_t, \,\mathbf{h}_t)$ is equivalent to minimising

$$J_t^{GLRT}(\mathbf{y}_t | \mathbf{X}_t, \mathbf{h}_t) = ||\mathbf{y}_t - \mathbf{X}_t\mathbf{h}_t||^2 \;. \tag{6.18}$$

So maximising J_t^{GLRT} over $\mathbf{h}_t$ yields the estimate

$$\hat{\mathbf{h}}_t = (\mathbf{X}_t^T\,\mathbf{X}_t)^{-1}\,\mathbf{X}_t^T\,\mathbf{y}_t \;. \tag{6.19}$$

Substitute this estimate into (6.18) gives

$$J_t^{GLRT}(\mathbf{y}_t | \mathbf{X}_t, \hat{\mathbf{h}}_t) = ||\mathbf{y}_t - \mathbf{X}_t(\mathbf{X}_t^T\,\mathbf{X}_t)^{-1}\mathbf{X}_t^T\,\mathbf{y}_t||^2 \;. \tag{6.20}$$

Thus, the GLRT receiver estimates the transmitted codeword according to

$$\hat{\mathbf{X}}_t^{GLRT} = \arg\min_k ||\mathbf{P}_k^{\perp}\,\mathbf{y}_t||^2 \;, \tag{6.21}$$

where $\mathbf{P}_k^{\perp} = \mathbf{I} - \mathbf{X}^{(k)}\,(\mathbf{X}^{(k)^T}\mathbf{X}^{(k)})^{-1}\,(\mathbf{X}^{(k)})^T$ is the projector orthogonal to the subspace spanned by the columns of the codeword $\mathbf{X}^{(k)}$.

6.3.2 Noncoherent Maximum Likelihood (NCML) Receiver

In the situation where the probability distribution of the fading channel is known, one can use the NCML receiver. The NCML receiver maximises the marginal density $p(\mathbf{y}_t|\mathbf{X}_t)$ over $\mathbf{X}_t$

$$\hat{\mathbf{X}}_t^{NCML} = \arg \max_{\mathbf{X} \in \Theta} p(\mathbf{y}_t \,|\, \mathbf{X}_t = \mathbf{X})$$

$$= \arg \max_{\mathbf{X} \in \Theta} E_{\mathbf{h}_t} \{ p(\mathbf{y}_t, \, \mathbf{h}_t | \, \mathbf{X}_t = \mathbf{X}) \} , \tag{6.22}$$

where the expectation is with respect to the channel $\mathbf{h}_t$. This is equivalent to choosing the codeword which minimises the following cost function:

$$J_t^{NCML}(\mathbf{X}) = (\mathbf{y}_t - \mathbf{X}\,\nu)^T (\mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2\mathbf{I})^{-1}(\mathbf{y}_t - \mathbf{X}\,\nu) + \log \det(\mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2\mathbf{I}) .$$

$$\tag{6.23}$$

where ν and $\mathbf{P}$ are the mean and covariance of $\mathbf{h}_t$. This receiver does not exploit the temporal correlation as it operates independently from codeword period to codeword period. However, it can be extended to account for the temporal correlation by incorporating multiple codewords into the decision metric (6.23), albeit at the cost of exponential increase in processing complexity.

6.3.3 Maximum A Posteriori Probability (MAP) Receiver

Another receiver which can be used when the statistics of the fading channel is known is the MAP receiver. The MAP receiver works by maximising over $\mathbf{h}_t$ and $\mathbf{X}_t$, the *a posteriori* probability $p(\mathbf{X}_t, \mathbf{h}_t| \mathbf{y}_t) \propto p(\mathbf{y}_t|\mathbf{h}_t, \mathbf{X}_t)\, p(\mathbf{h}_t)$, assuming the codes are chosen from the codebook with uniform probability. That is

$$\hat{\mathbf{X}}_t^{MAP} = \arg \max_{\mathbf{X} \in \Theta} \sup_{\mathbf{h}_t} p(\mathbf{y}_t, \, \mathbf{h}_t \,|\, \mathbf{X}_t = \mathbf{X}) ,$$

$$= \arg \max_{\mathbf{X} \in \Theta} \sup_{\mathbf{h}_t} p(\mathbf{y}_t \,|\, \mathbf{h}_t, \, \mathbf{X})\, p(\mathbf{h}_t) , \tag{6.24}$$

where

$$p(\mathbf{y}_t \,|\, \mathbf{X}, \mathbf{h}_t)\, p(\mathbf{h}_t) = \mathcal{N}(\mathbf{y}_t - \mathbf{X}\,\mathbf{h}_t;\, \sigma^2 \mathbf{I})\, \mathcal{N}(\mathbf{h}_t - \nu;\, \mathbf{P})\,. \tag{6.25}$$

Using the Gaussian Product Lemma, $p(\mathbf{y}_t \,|\, \mathbf{X}, \mathbf{h}_t)\, p(\mathbf{h}_t)$ can be expressed as

$$p(\mathbf{y}_t \,|\, \mathbf{X}, \mathbf{h}_t)\, p(\mathbf{h}_t) = \mathcal{N}(\mathbf{y}_t - \mathbf{X}\,\nu;\, \mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2 \mathbf{I})\, \mathcal{N}(\mathbf{h}_t - \hat{\mathbf{h}}_t(\mathbf{X});\, \hat{\mathbf{P}}_t(\mathbf{X}))\,, \tag{6.26}$$

where

$$\hat{\mathbf{h}}_t(\mathbf{X}) = \nu + \mathbf{P}\,\mathbf{X}^T\, (\mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2 \mathbf{I})^{-1}\, (\mathbf{y}_t - \mathbf{X}\,\nu)$$

$$\hat{\mathbf{P}}_t(\mathbf{X}) = \mathbf{P} - \mathbf{P}\,\mathbf{X}^T\, (\mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2 \mathbf{I})^{-1}\, \mathbf{X}\,\mathbf{P}\,. \tag{6.27}$$

Clearly $\hat{\mathbf{h}}_t(\mathbf{X})$ maximises $p(\mathbf{y}_t \,|\, \mathbf{X}, \mathbf{h}_t)\, p(\mathbf{h}_t)$ over $\mathbf{h}_t$ and the maximised value is

$$p(\mathbf{y}_t \,|\, \mathbf{X}, \hat{\mathbf{h}}_t(\mathbf{X}))\, p(\hat{\mathbf{h}}_t(\mathbf{X})) = \mathcal{N}(\mathbf{y}_t - \mathbf{X}\,\nu;\, \mathbf{X}\,\mathbf{P}\,\mathbf{X}^T + \sigma^2 \mathbf{I})\, \mathcal{N}(0;\, \hat{\mathbf{P}}_t(\mathbf{X}))\,. \tag{6.28}$$

Thus, the cost function that the MAP receiver needs to minimise is the same as that of the noncoherent ML receiver plus an extra term $\log \det \hat{\mathbf{P}}_t$ (neglecting constant terms):

$$J_t^{MAP}(\mathbf{X}) = J_t^{NCML}(\mathbf{X}) + \log \det \hat{\mathbf{P}}_t(\mathbf{X})\,. \tag{6.29}$$

More detailed description of the GLRT, NCML and MAP approaches can be found in [55].

The receivers described above do not exploit the temporal correlation of the fading channels and operate on each symbol period independently. In the following sections, we describe several receivers which take this temporal correlation into account when performing the detection.

6.3.4 Sequence Estimation Approaches

Consider a sequence of codewords $\{\mathbf{X}_0, \mathbf{X}_1, ..., \mathbf{X}_{T-1}\}$ being transmitted during T codeword intervals. In order to obtain an optimal solution which maximises the likelihood $p(\mathbf{y}_0, ..., \mathbf{y}_{T-1} | \mathbf{X}_0, ..., \mathbf{X}_{T-1})$, there must be a Kalman filter for each of the

possible sequence (model). This direct implementation of the optimal receiver has a complexity of $O(K^T)$ which grows exponentially with the length of the sequence. Thus, even for small T, it may not be practical to use this method. Therefore, we need to examine other suboptimal methods. In this section, we describe two receivers which approximately optimise the likelihood of the transmitted sequence, i.e $p(\mathbf{y}_0, ..., \mathbf{y}_{T-1} \mid \mathbf{X}_0, ..., \mathbf{X}_{T-1})$.

Per-Survivor Processing (PSP) Receiver

In this section, we describe a receiver which approximates the optimal receiver by using the per-survivor processing method [79]. This method can be implemented via the Viterbi Algorithm. In this method, there will be K Kalman filters (KFs), each tuned to one of the K codewords. Fig. 6.1 illustrates the structure of this receiver.
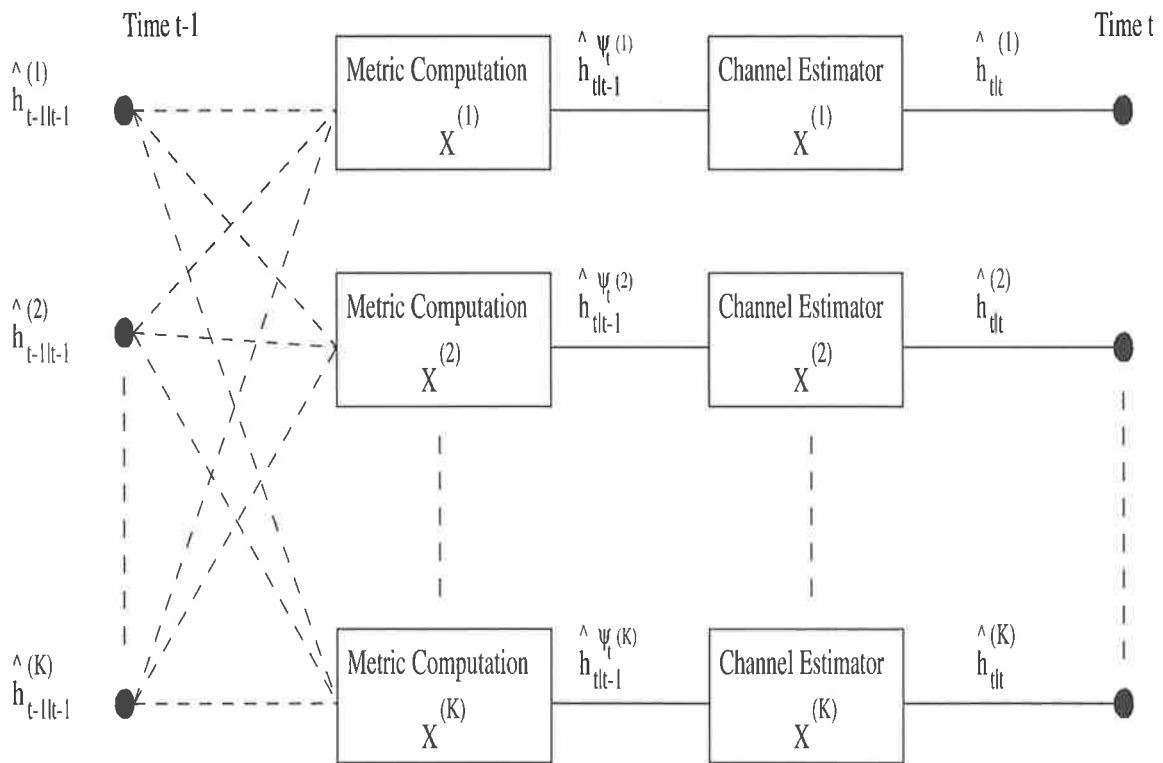

Figure 6.1: Structure of PSP Receiver

At time t, for each codeword the receiver calculates the log likelihood from the possible paths. After computing the path metric, it retains the path with the largest log likelihood and computes the channel estimate for this path. The approximately optimal log likelihood at time t of the kth KF tuned to $\mathbf{X}_t = \mathbf{X}^{(k)}$ can be evaluated according to

$$\delta_t^{(k)} = \max_{j=1,...,K} \left(\delta_{t-1}^{(j)} - \epsilon_t^{(j,k)^T} \Omega_t^{(j,k)-1} \epsilon_t^{(j,k)} - \log \det \Omega_t^{(j,k)} \right) , \qquad (6.30)$$

where $\delta_{t-1}^{(j)}$ is the cummulative path metric at time $t-1$ of the survivor path ending at $\mathbf{X}_{t-1} = \mathbf{X}^{(j)}$, $\epsilon_t^{(j,k)}$ is the innovation sequence of the path from $\mathbf{X}_{t-1} = \mathbf{X}^{(j)}$ to $\mathbf{X}_t = \mathbf{X}^{(k)}$, and $\Omega_t^{(j,k)}$ is its associated covariance

$$\epsilon_t^{(j,k)} = \mathbf{y}_t - \mathbf{X}^{(k)} \hat{\mathbf{h}}_{t|t-1}^{(j)} \qquad (6.31)$$

$$\Omega_t^{(j,k)} = \mathbf{X}^{(k)} \Sigma_{t|t-1}^{(j)} \mathbf{X}^{(k)^T} + \sigma^2 \mathbf{I}$$

The predicted channel estimate $\hat{\mathbf{h}}_{t|t-1}^{(j)}$ and its covariance $\Sigma_{t|t-1}^{(j)}$ of the path from $\mathbf{X}_{t-1} = \mathbf{X}^{(j)}$ are determined from the filtered channel estimate $\hat{\mathbf{h}}_{t-1|t-1}^{(j)}$ and its covariance $\Sigma_{t-1|t-1}^{(j)}$ via the KF time update equations

$$\hat{\mathbf{h}}_{t|t-1}^{(j)} = \mathbf{A}\hat{\mathbf{h}}_{t-1|t-1}^{(j)} + \mu , \qquad (6.32)$$

$$\Sigma_{t|t-1}^{(j)} = \mathbf{A}\Sigma_{t-1|t-1}^{(j)}\mathbf{A}^T + \mathbf{Q} .$$

Once the survivor path at $\mathbf{X}_t = \mathbf{X}^{(k)}$ is determined, the KF then estimates the channel associated with this path using the received signal at time t. Let the predecessor of $\mathbf{X}_t = \mathbf{X}^{(k)}$ which maximises the log likelihood $\delta_t^{(k)}$ be denoted by $\psi_t(k)$. The filtered channel estimate of the survivor path at time t ending with $\mathbf{X}_t = \mathbf{X}^{(k)}$ can be evaluated by the KF measurement update equations

$$\hat{\mathbf{h}}_{t|t}^{(k)} = \hat{\mathbf{h}}_{t|t-1}^{(\psi_t(k))} + \mathbf{G}_t^{(\psi_t(k))} \left(\mathbf{y}_t - \mathbf{X}^{(k)} \hat{\mathbf{h}}_{t|t-1}^{(\psi_t(k))} \right) ,$$

$$\mathbf{G}_t^{(\psi_t(k))} = \Sigma_{t|t-1}^{(\psi_t(k))} \mathbf{X}^{(k)^T} \left(\mathbf{X}^{(k)} \Sigma_{t|t-1}^{(\psi_t(k))} \mathbf{X}^{(k)^T} + \sigma^2 \mathbf{I} \right)^{-1} ,$$

$$\Sigma_{t|t}^{(k)} = \Sigma_{t|t-1}^{(\psi_t(k))} - \mathbf{G}_t^{(\psi_t(k))} \mathbf{X}^{(k)} \Sigma_{t|t-1}^{(\psi_t(k))} . \qquad (6.33)$$

This method has a complexity of O(K) which is a huge reduction as compare to the optimal detector. It is possible to develop an $O(K^2)$ version of this approach, in which each Kalman filter which tunes to $\mathbf{X}_t$, produces K filtered channel estimates, one for each survivor paths from time $t - 1$. Once the channel estimates are made, it computes the path metric and selects the one which maximises the log likelihood. The computation of the path metric is the same as in (6.30), with the predicted channel estimate and its covariance replaced by the filtered channel estimates and its associated covariance.

The space-time coded system in time varying fading channels from (5.3) can be cast into this general diversity form (6.1) by letting $\tilde{\mathbf{X}}_t = \sqrt{E_s}\chi(t)$. Since $\chi(t) = \mathbf{I}_{N_R} \otimes [x_1(t) \cdots x_{n_T}](t)$, the total number of possible $\tilde{\mathbf{X}}_t$ is $K = \beta^{N_T}$ where β is the size of the modulation constellation in which $x_i(t)$ is belonged to and n_T is the number of transmit antennas. Thus, this PSP receiver can be used for joint channel estimation and space-time trellis decoding similar to that in chapter 5. However, it should be noticed that this PSP receiver is not the same as that presented in chapter 5. Firstly, the number of KFs in this PSP receiver is K (which is dependent on β and n_T) while the number of KFs in the PSP receiver in the previous chapter equals to the number of states of the space-time trellis code (which is independent of β and n_T). Secondly, this PSP receiver does not exploit the trellis structure of the space-time trellis code and hence it would lead to a lower performance than the PSP receiver in the previous chapter. Nevertheless, this PSP receiver might be a preferable receiver in a situation when K is much smaller than the number of states of the space-time trellis code. Using this PSP receiver will then significantly reduce the processing complexity.

M-Algorithm Receiver

The PSP method offers a much lower complexity than the optimal receiver. However, there are situations in which the codebook size, K, is too large which may prevent the PSP method from being used. Also the PSP receiver retains only one survivor path for each codeword at time t. This could be a disadvantage as the discarded paths at some codewords could have a much larger likelihood than the survivor paths at other codewords. Motivated by this observation, we developed a receiver which is based on the M-algorithm [42] for detecting the transmitted codeword. The complexity of this receiver can vary from $O(1)$ to $O(K)$ depending on the computation complexity that can be afforded at the receiver terminal.
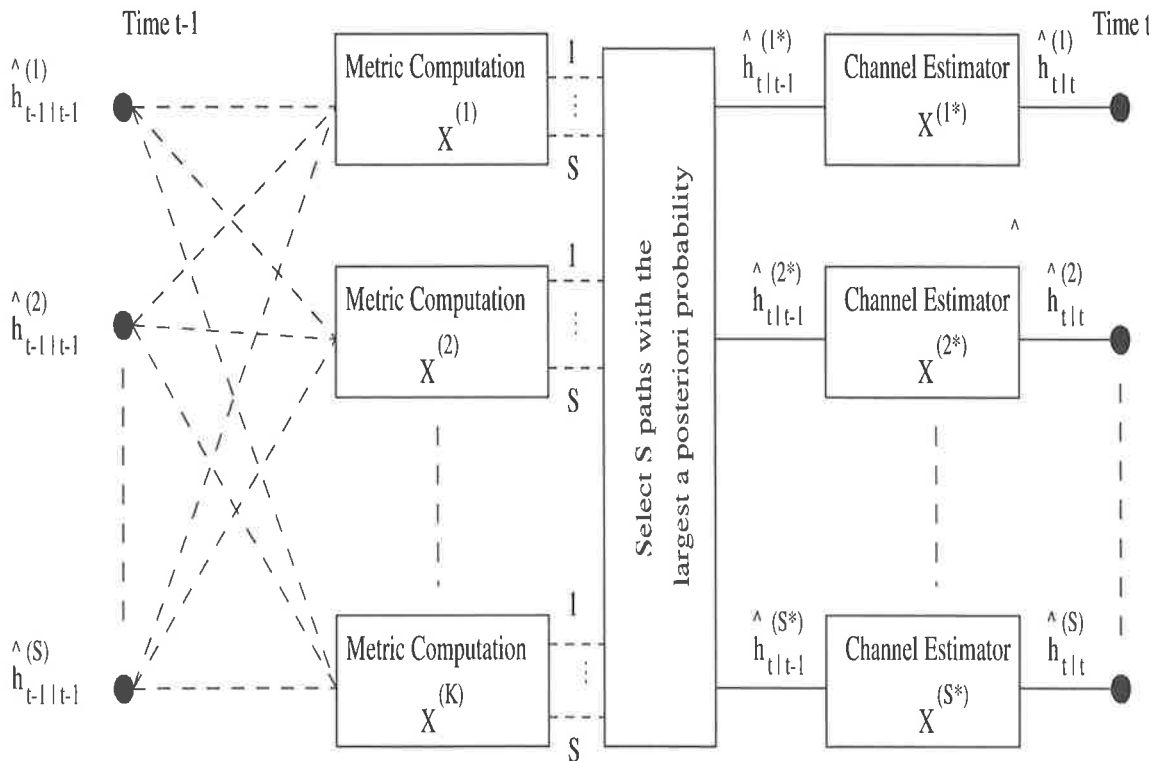

Figure 6.2: Structure of M-algorithm Receiver

The M-algorithm receiver is implemented using a bank of S Kalman filters where

$1 \leq S \leq K$. A pictorial description of the M-algorithm receiver is depicted in Fig. 6.2. This receiver works as follow: At the end of each symbol period, the receiver retains S channel estimates with the largest log likelihood. Suppose at the end of time $t-1$, the S survivor channel estimates are $\hat{\mathbf{h}}^{(m)}_{t-1|t-1}$ for $m = 1, ..., S$. At time t, for each codeword $\mathbf{X}_t$, the receiver computes S path metrics using the predicted channel estimates $\hat{\mathbf{h}}^{(m)}_{t|t-1} = \mathbf{A}\,\hat{\mathbf{h}}^{(m)}_{t-1|t-1} + \mu$. The log likelihood of the path from the mth survivor at time $t-1$ and $\mathbf{X}_t = \mathbf{X}^{(k)}$ is calculated by

$$\tilde{\delta}^{(m,k)}_t = \delta^{(m)}_{t-1} - \epsilon^{(m,k)^T}_t \, \Omega^{(m,k)^{-1}}_t \, \epsilon^{(m,k)}_t \; - \log \det \Omega^{(m,k)}_t \,. \tag{6.34}$$

where $\delta^{(m)}_{t-1}$ is the log likelihood of the survivor path m at time $t-1$ and

$$\epsilon^{(m,k)}_t = \mathbf{y}_t - \mathbf{X}^{(k)}\,\hat{\mathbf{h}}^{(m)}_{t|t-1} \,,$$
$$\Omega^{(m,k)}_t = \mathbf{X}^{(k)}\,\Sigma^{(m)}_{t|t-1}\,\mathbf{X}^{(k)^T} + \sigma^2 \mathbf{I} \,, \tag{6.35}$$

are the innovation sequence and its covariance, respectively. Thus, there are SK candidate paths in total and the M-algorithm selects S paths with the largest log likelihood. Let the path from the mth survivor at time $t-1$ and $\mathbf{X}_t = \mathbf{X}^{(k)}$ being selected by the M-algorithm as the sth survivor path at time t. The filtered channel estimate of this s^{th} survivor path can then be updated by using a Kalman filter tuned to $\mathbf{X}_t = \mathbf{X}^{(k)}$. Thus,

$$\hat{\mathbf{h}}^{(s)}_{t|t} = \hat{\mathbf{h}}^{(m)}_{t|t-1} + \mathbf{G}^{(m,k)}_t \left(\mathbf{y}_t - \mathbf{X}^{(k)}\,\mathbf{h}^{(m)}_{t|t-1} \right)$$
$$\mathbf{G}^{(m,k)}_t = \Sigma^{(m)}_{t|t-1}\,\mathbf{X}^{(k)^T} \left(\mathbf{X}^{(k)}\,\Sigma^{(m)}_{t|t-1}\,\mathbf{X}^{(k)^T} + \sigma^2 \mathbf{I} \right)^{-1}$$
$$\Sigma^{(s)}_{t|t} = \Sigma^{(m)}_{t|t-1} - \mathbf{G}^{(m,k)}_t\,\mathbf{X}^{(k)}\,\Sigma^{(m)}_{t|t-1} \,. \tag{6.36}$$

In both the PSP and M-algorithm, the optimal sequence is extracted by backtracking through the survivor paths. On-line versions can also be derived.

6.3.5   Codeword by Codeword Estimation

The receivers described previously are used for sequence estimation. In this section, we describe two other techniques which detect the transmitted codeword at each codeword interval. Thus, these receivers can be used in applications which can't tolerate the delays associated with the Viterbi algorithm backtracking, or when we desire to incorporate adaptivity to the channel statistics.

The *A Posteriori* Probability (APP) Receiver

Unlike the receivers described previously which retain several estimates of the fading channels at the end of each codeword interval, this receiver produces a single channel estimate which is the weighted sum of all the estimates. This receiver is implemented by using a bank of K Kalman filters where each Kalman filter is tuned to one of the codewords. Fig. 6.3 illustrates the structure of this receiver.

Based on the assumed signal model, each Kalman filter gives a conditional channel estimate.

$$\hat{\mathbf{h}}_{t|t}^{(k)} = \hat{\mathbf{h}}_{t|t-1} + \mathbf{G}_t^{(k)} \left(\mathbf{y}_t - \mathbf{X}^{(k)} \hat{\mathbf{h}}_{t|t-1} \right) \ ,$$

$$\mathbf{G}_t^{(k)} = \Sigma_{t|t-1} \mathbf{X}^{(k)T} \left(\mathbf{X}^{(k)} \Sigma_{t|t-1} \mathbf{X}^{(k)T} + \sigma^2 \mathbf{I} \right)^{-1} \ ,$$

$$\Sigma_{t|t}^{(k)} = \Sigma_{t|t-1} - \mathbf{G}_t^{(k)} \mathbf{X}^{(k)} \Sigma_{t|t-1} \ . \tag{6.37}$$

where $\hat{\mathbf{h}}_{t|t-1} = \mathbf{A}\hat{\mathbf{h}}_{t-1|t-1} + \mu$ is the predicted channel estimate. These channel estimates will then be weighted and summed to yield the mean channel estimate. The weighting coefficient of the channel estimate $\hat{\mathbf{h}}_{t|t}^{(k)}$ is the *a posteriori* probability that $\mathbf{X}_t = \mathbf{X}^{(k)}$, i.e. $p(\mathbf{X}_t = \mathbf{X}^{(k)} \,|\, \mathbf{y}_0, ..., \mathbf{y}_t)$. Using the Bayes rule, the *a posteriori* probability that $\mathbf{X}_t$ take on each value in Θ is computed according to

$$\Pi_t(k) = p(\mathbf{X}_t = \mathbf{X}^{(k)} \,|\, \mathbf{y}_0, ..., \mathbf{y}_t) = c_t \, p(\mathbf{y}_t \,|\, \mathbf{y}_0, ..., \mathbf{y}_{t-1}, \mathbf{X}_t = \mathbf{X}^{(k)}) \ , \tag{6.38}$$

where

$$p(\mathbf{y}_t \,|\, \mathbf{y}_0, ..., \mathbf{y}_{t-1}, \mathbf{X}_t = \mathbf{X}^{(k)}) = \mathcal{N}(\epsilon_t^{(k)}; \Omega_t^{(k)}) \ , \tag{6.39}$$
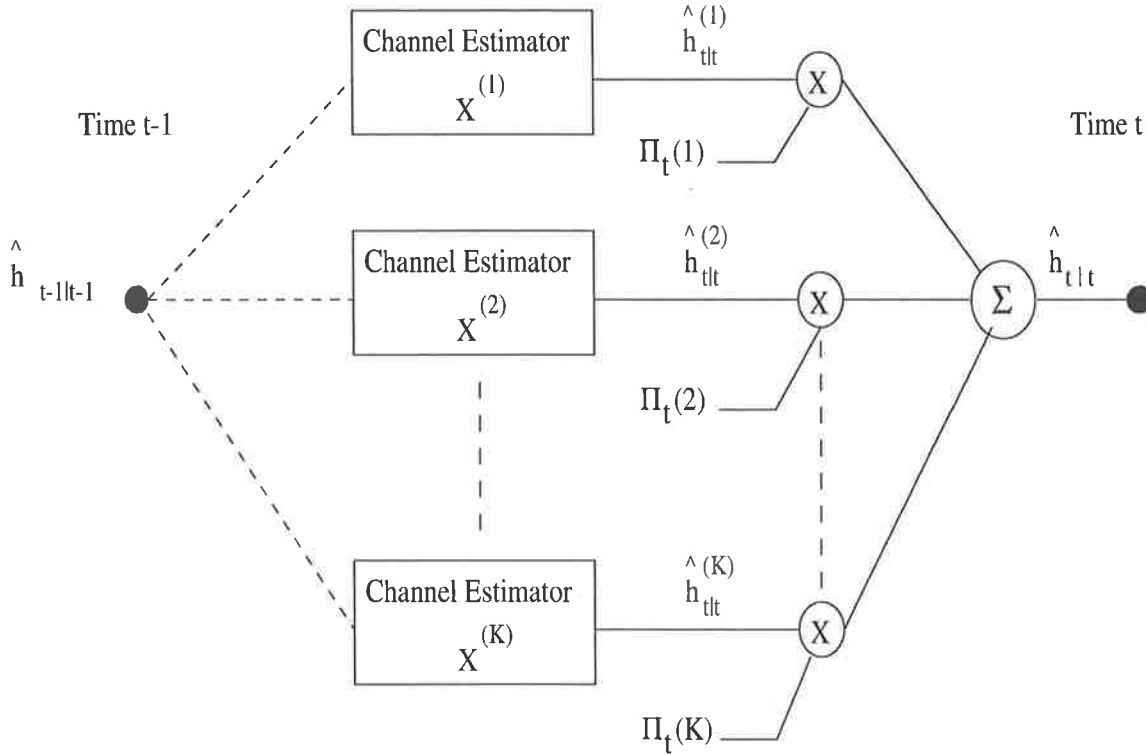
Figure 6.3: Structure of APP Receiver

c_t is the normalised factor such that $\sum_k \Pi_t(k) = 1$, $\epsilon_t^{(k)} = \mathbf{y}_t - \mathbf{X}^{(k)} \hat{\mathbf{h}}_{t|t-1}$ is the innovations sequence from the Kalman filter tuned to $X_t = X^{(k)}$ and $\Omega_t^{(k)} = \mathbf{X}^{(k)} \Sigma_{t|t-1} \mathbf{X}^{(k)^T} + \sigma^2 \mathbf{I}$ is its associated covariance. Thus, the mean channel estimate and its associated covariance at time t are

$$\hat{\mathbf{h}}_{t|t} = \sum_k \Pi_t(k) \, \hat{\mathbf{h}}_{t|t}^{(k)} \, , \tag{6.40}$$

$$\Sigma_{t|t} = \sum_k \Pi_t(k) \left\{ \Sigma_{t|t}^{(k)} + \hat{\mathbf{h}}_{t|t}^{(k)} (\hat{\mathbf{h}}_{t|t}^{(k)})^T \right\} - \hat{\mathbf{h}}_{t|t} \, \hat{\mathbf{h}}_{t|t}^T \, .$$

In effect we collapse the K Kalman filters' estimates back to one each time. The estimated transmitted codeword at time t is that which maximises the *a posteriori* probability.

From observation, we notice that most of the *a posteriori* probabilities that $\mathbf{X}_t = \mathbf{X}^{(k)}$ for $k = 1, .., K$ are very small. In fact, most of them are almost zero. Thus, the

contribution of the channel estimates after being weighted by the small *a posteriori* probabilities to the mean channel estimate will be insignificant. Motivated by this observation, we will modify this receiver so that it will not compute the filtered channel estimate for the codeword with small *a posteriori* probability. Thus we can reduce the complexity of the receiver with little penalty on performance. The modified receiver will works as follow. After computing the *a posteriori* probability that $\mathbf{X}_t$ takes on a value in Θ as in (6.38), S codewords which have the greatest *a posteriori* probabilities will be selected, where $1 \leq S \leq K$. Lets s denotes a member of the set of S survivors. The *a posteriori* probabilities are then normalised so that $\sum_s \Pi_t(s) = 1$. For each of these survivor paths, a filtered channel estimate, $\hat{\mathbf{h}}_{t|t}^{(s)}$ and its associated covariance $\Sigma_{t|t}^{(s)}$ are calculated. These estimates are then combined to yield a single channel estimate and its covariance for time t. Thus,

$$\hat{\mathbf{h}}_{t|t} = \sum_s \Pi_t(s)\, \hat{\mathbf{h}}_{t|t}^{(s)} \; , \tag{6.41}$$

$$\Sigma_{t|t} = \sum_s \Pi_t(s) \left\{ \Sigma_{t|t}^{(s)} + \hat{\mathbf{h}}_{t|t}^{(s)}(\hat{\mathbf{h}}_{t|t}^{(s)})^T \right\} - \hat{\mathbf{h}}_{t|t}\, \hat{\mathbf{h}}_{t|t}^T \; .$$

This receiver has the flavour of the M-algorithm technique but it is not the same since at the end of each codeword period, this receiver retains only one channel estimate while the M-algorithm retains S channel estimates. The complexity of this receiver is $\mathrm{O}(S)$. When $S = K$ this receiver corresponds to the unmodified APP receiver as described previously.

Iterative EM Receiver

Another receiver which we proposed is the iterative receiver. In this approach, we use the Expectation-Maximisation algorithm to iteratively maximise the log likelihood function, $\log p(\mathbf{y}_t \mid \mathbf{y}_0, ..., \mathbf{y}_{t-1}, \mathbf{X}_t = \mathbf{X}^{(k)})$. The iteration is per-codeword basis. When the channel is known, the log likelihood is

$$L(t) = \log p(\mathbf{y}_t \mid \mathbf{h}_t, \mathbf{X}_t) = -\sigma^{-2}\, \|\mathbf{y}_t - \mathbf{X}_t\, \mathbf{h}_t\|^2 - M\, \log \sigma^2 \; , \tag{6.42}$$

Since the knowledge of the channel $\mathbf{h}_t$ is not available at the receiver, the conditional expectation of this log likelihood given the present and past measurements is used instead as the objective cost function,

$$J_t(\mathbf{X}_t, \hat{\mathbf{h}}_{t|t}, \Sigma_{t|t}) = E\{L(t) \,|\, \mathbf{y}_0, ..., \mathbf{y}_t\} = -\epsilon_t^T \,\Omega_t^{-1}\, \epsilon_t - \log \det \Omega_t \;, \qquad (6.43)$$

where $\epsilon_t = \mathbf{y}_t - \mathbf{X}_t \hat{\mathbf{h}}_{t|t}$ and $\Omega_t = \mathbf{X}_t \Sigma_{t|t} \mathbf{X}_t^T + \sigma^2 \mathbf{I}$. This motivates the iterative scheme whereby the conditional expectations is computed using $\mathbf{X}_t = \mathbf{X}^{(\hat{k}_\ell)}$ and then maximised J_t over all codewords to yield $\mathbf{X}^{(\hat{k}_{\ell+1})}$. The iterative process for determining the transmitted codeword at time t can be described in details as follow:

Initialisation: At the start of each iteration process, the channel estimate and its associated covariance are initialised to the Kalman filter one step prediction

$$\tilde{\mathbf{h}}_t^{(0)} = \hat{\mathbf{h}}_{t|t-1} = \mathbf{A}\,\hat{\mathbf{h}}_{t-1|t-1} + \mu$$

$$\tilde{\Sigma}_t^{(0)} = \Sigma_{t|t-1} = \mathbf{A}\,\Sigma_{t-1|t-1}\,\mathbf{A}^T + \mathbf{Q}\;. \qquad (6.44)$$

The initial codeword is chose to maximise the cost function

$$\hat{\mathbf{X}}_t^{(0)} = \arg\max_{\mathbf{X}_t \in \Theta}\; J(\mathbf{X}_t, \tilde{\mathbf{h}}_t^{(0)}, \tilde{\Sigma}_t^{(0)})\;. \qquad (6.45)$$

Iteration Update: We update from iteration $\ell - 1$ to ℓ for $\ell \geq 1$ via

$$\tilde{\mathbf{h}}_t^{(\ell)} = \hat{\mathbf{h}}_{t|t-1} + \mathbf{K}_t^{(\ell)} \left(\mathbf{y}_t - \hat{\mathbf{X}}_t^{(\ell-1)}\,\hat{\mathbf{h}}_{t|t-1} \right)$$

$$\mathbf{K}_t^{(\ell)} = \Sigma_{t|t-1}\,(\hat{\mathbf{X}}_t^{(\ell-1)})^T \left(\hat{\mathbf{X}}_t^{(\ell-1)}\,\Sigma_{t|t-1}\,(\hat{\mathbf{X}}_t^{(\ell-1)})^T + \sigma^2\,\mathbf{I} \right)^{-1}$$

$$\tilde{\Sigma}_t^{(\ell)} = \Sigma_{t|t-1} - \mathbf{K}_t^{(\ell)}\hat{\mathbf{X}}_t^{(\ell-1)}\,\Sigma_{t|t-1}$$

$$\hat{\mathbf{X}}_t^{(\ell)} = \arg\max_{\mathbf{X} \in \Theta}\; J_t(\mathbf{X}, \tilde{\mathbf{h}}_t^{(\ell)}, \tilde{\Sigma}_t^{(\ell)})\;. \qquad (6.46)$$

The algorithm terminates when the maximising codeword is unaltered. We then set the values of $\hat{\mathbf{h}}_{t|t}$ and $\Sigma_{t|t}$ to the final values of $\tilde{\mathbf{h}}_t^{(\ell)}$ and $\tilde{\Sigma}_t^{(\ell)}$ respectively and progress to codeword $t + 1$. In [60], an iterative receiver based on Kalman filtering similar to ours is also proposed. However, this receiver is designed specifically for decoding the Alamouti's space-time block code with two transmit and one receive antenna. Our proposed iterative EM receiver is more general.

6.4 Application to OFDM systems

One important example of a diversity system is (OFDM) which is used in various wireless local area networks. In an OFDM system data is transmitted over L orthogonal frequency channels using a discrete Fourier transform (DFT). Within our context, we can regard the DFT operation as a type of "time-frequency" coding. Generally, a number of different data symbols (from a single user, or a number of different users) are transmitted simultaneously on subsets of the available channels. Here we will consider the single-user case, where one data symbol is transmitted on all L channels simultaneously. This approach offers the maximal diversity gain. This is equivalent to using the precoder matrix $\mathbf{C}$ equal to the L-vector of all 1s. Thus corresponding to each data symbol (regarded as a member of a finite complex alphabet) s_t we transmit a segment of $M = L$ time samples containing all L carriers, and given by

$$x_{tL+m} = \sum_{\ell=0}^{L-1} s_t \ e^{2\pi i m \ell / L} \ , \tag{6.47}$$

for $m = 0, \ldots, L - 1$. In a standard OFDM system, a cyclic prefix is added to mitigate against inter-symbol interference, and this is stripped off in the receiver. The recovered received samples then have the form

$$r_{tL+m} = \sum_{\ell=0}^{L-1} s_t \ h_{t,\ell} \ e^{2\pi i m \ell / L} + v_{tL+m} \ , \tag{6.48}$$

where $h_{t,\ell}$ denote the channel path gains as a function of symbol time t and frequency channel ℓ. Using this model, the frequency channels are spaced by an amount equal to the symbol rate. Here v_τ is a zero mean white Gaussian noise process. Now blocking the received samples into a vector of length L corresponding to the symbol s_t we have

$$\tilde{\mathbf{y}}_t = \tilde{\mathbf{X}}_t \ \tilde{\mathbf{h}}_t + \tilde{\mathbf{n}}_t \ , \tag{6.49}$$

where $\tilde{\mathbf{X}}_t$ is a $L \times L$ matrix with elements $[\tilde{\mathbf{X}}_t]_{m,\ell} = s_t \ e^{2\pi i m \ell / L}$ and $\tilde{\mathbf{h}}_t$ is an $L \times 1$ vector with elements $[\tilde{\mathbf{h}}_t]_\ell = h_{t,\ell}$. We thus have placed this OFDM system model

in the general diversity form (6.1). Note that the assumption that only one user symbol is transmitted over the L channels is not restrictive as indicated in 6.2.1.

6.4.1 A Physical Model for Delay-Doppler Spread Channels

In this section we relate our statistical channel gain model to a commonly used physical model for multipath propagation. This provides a parameterisation of our state space model in terms of specified physical parameters such as delay and doppler spreads for the channel. In [49], a physical model is used to characterise the correlation between the time-frequency gains $h_{t,\ell}$ in (6.48). This model is using the statistics of the channel and uses a block representation of the channel based on its autocorrelation matrix. For a classical Doppler power spectrum and exponential multipath intensity profile, the correlation between two symbols spaced in time and frequency respectively with Δt and Δf is given by [51]:

$$\phi(\Delta t, \Delta f) = \phi_0 \ \phi_t(\Delta t) \ \phi_f(\Delta f) \tag{6.50}$$

with $\phi_0 > 0$,

$$\phi_t(\Delta t) = J_0(\pi B_d \Delta t) \ , \qquad \phi_f(\Delta f) = \frac{1}{1 + j2\pi T_m \Delta f} \ . \tag{6.51}$$

Here J_0 is the zero-order Bessel function of the first kind, with B_d and T_m being the Doppler and delay spreads respectively of the propagation channel. From (6.7), we can thus see that the $\tilde{\mathbf{P}}$ matrix represents by itself the frequency correlation and $\tilde{\mathbf{A}}^{t_1-t_2}$ is the time correlation between two symbols separated by a time shift equal to $(t_1 - t_2)$. Thus immediately,

$$[\tilde{\mathbf{P}}]_{k,l} = \frac{\phi_0}{1 + j2\pi T_m(k - l)F_s} \ , \tag{6.52}$$

F_s being the frequency spacing between two OFDM subcarriers.

Due to the separable nature of the physical channel correlation function (6.50), we take our state transition matrix $\tilde{\mathbf{A}}$ to be of the form $\tilde{\mathbf{A}} = a \ \mathbf{I}$, where a is a

complex constant. Comparing the time correlation between two OFDM symbols separated by k symbol periods, we need to have $a^k = J_0(\pi B_d k T_s)$ for every k, where T_s is the symbol period. This is theoretically impossible, so an approximation has to be made for the $\{a_k\}_{k=0}^{D-1}$ to fit the $\{J_0(\pi B_d k T_s)\}_{k=0}^{D-1}$, D being the number of OFDM symbols corresponding to the time-length after which the correlation is considered to be insignificant, see [50]. This approximation is given by solving the least squares problem:

$$\min_a \sum_{k=0}^{D-1} |J_0(\pi B_d k T_s) - a^k|^2 \ . \tag{6.53}$$

This problem can be solved for various values of normalised Doppler spreads $(B_d T_s)$ and the corresponding values of a used in the state space model. One could also choose to include additional "modes" in $\tilde{\mathbf{A}}$ to obtain a better approximation, but we do not provide details here.

6.5   Simulation

We considered a single user OFDM system with 16 carriers. The data sequence is arranged into a block of 16 OFDM symbols. Each symbol is transmitted over the 16 orthogonal frequency channels using the DFT. The first symbol of each OFDM block will be used as pilot symbol to generate an estimate to initialise the Kalman filters. We assumed that the fading channels are independent from block to block. However, within each block, the time-frequency covariance of the fading channels between any two symbols is given by the model in section 6.4. We assumed the channel is undergone Rayleigh fading, (i.e., $\mu = 0$). In this simulation, we used $B_d T_s = 0.25$. This Doppler spread value corresponds to a fast time-varying channel where the temporal correlation between any two consecutive symbols is only 0.852. Even with such low temporal correlation in the channels, a significant performance gain can be attained by exploiting this time correlation as will be shown later in the

simulation results. We examined two cases when $T_m F_s = 0.5$ and $T_m F_s = 0.025$. These values of $T_m F_s$ represent fading channels that are highly decorrelated and highly correlated in frequency dimension, respectively.

In this simulation we compared the performance of the proposed techniques with the NCML, MAP and GLRT receivers. In addition, we also compared them with the performance of the coherent Maximum Likelihood (ML) receiver which has ideal channel state information (CSI). The performance of the coherent ML receiver will be used as the benchmark and we will see how close the performance of the proposed receivers can approach this bound. The PSP and the M-algorithm receivers perform sequence estimation on the whole OFDM block of 16 symbols while the APP and the iterative EM receivers perform symbol by symbol detection. The M-algorithm receiver with two survivors (i.e. $S = 2$) and the unmodified APP receiver (i.e. the channel estimate is the weighted sum of the estimates from all the Kalman filters) are used in the simulation.

As from (6.49), the codeword $\tilde{\mathbf{X}}_t$ has elements $[\tilde{\mathbf{X}}_t]_{m,\ell} = s_t\ e^{2\pi i m \ell / L}$. In this simulation we considered the case where the data symbol $s_t \in \{\pm 1\}$. Thus, the codebook consists of two codes which are $\pm F$ where F is the Fourier matrix. Fig. 6.4 shows the performance of the receivers for the fading channel with $B_d T_s = 0.25$ and $T_m F_s = 0.5$. The symbol error rate (SER) is plotted for different values of Signal-to-Noise Ratio (SNR) where the SNR is defined as the average received SNR per symbol. The standard GLRT, MAP and NCML receivers all have an SER = 0.5 for all SNRs. This is because by using two codes which are scalar multiple of each other, the decision metrics of these receivers are indistinguishable (for the Rayleigh fading channels) between the two codewords as discussed in [44, 104]. Since the proposed receivers exploit the time correlation, then by using a pilot symbol to provide an estimate to initialise the Kalman filters, the proposed receivers do not suffer the same problem as those conventional noncoherent receivers. As shown in Fig. 6.4,
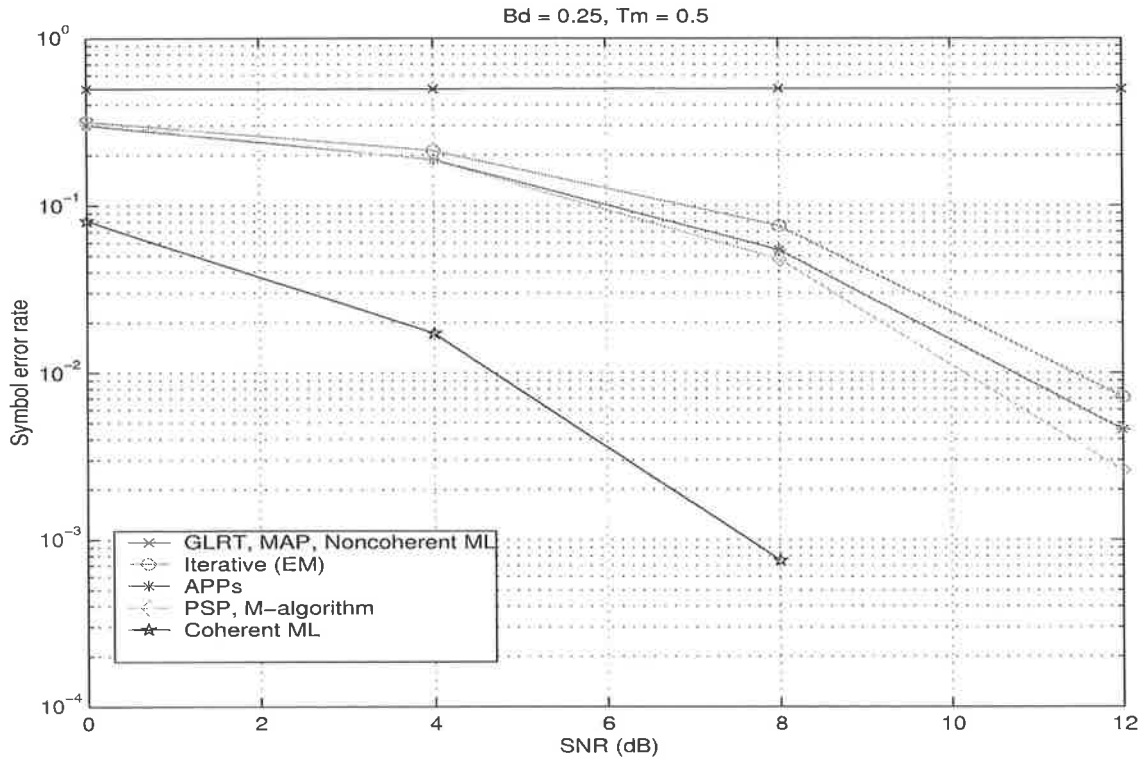
Figure 6.4: Codes $= \pm F$ with $B_d T_s = 0.25$ and $T_m F_s = 0.5$

at the $SER = 10^{-2}$ the performance of the proposed receivers is within $5 - 7dB$ of the coherent ML receiver. However, as the frequency correlation is increased, the performance of the proposed techniques is significantly reduced as shown in Fig. 6.5 for the highly frequency correlated channel with $T_m F_s = 0.025$. The reason for this huge degradation in performance is that as the channel become highly correlated in frequency, the diversity provided from the frequency domain is reduced. Thus, the probability that most of the channel coefficients faded at the same time is high, resulting in higher probability of error in detecting the transmitted code. Since the proposed receivers perform joint data detection and channel tracking, once an error is occurred it is likely to cause error propagation resulting in further error for the subsequence symbols in the block. The performance of the coherent ML receiver is also reduced in this case. However, this is mainly due to the loss in the diversity
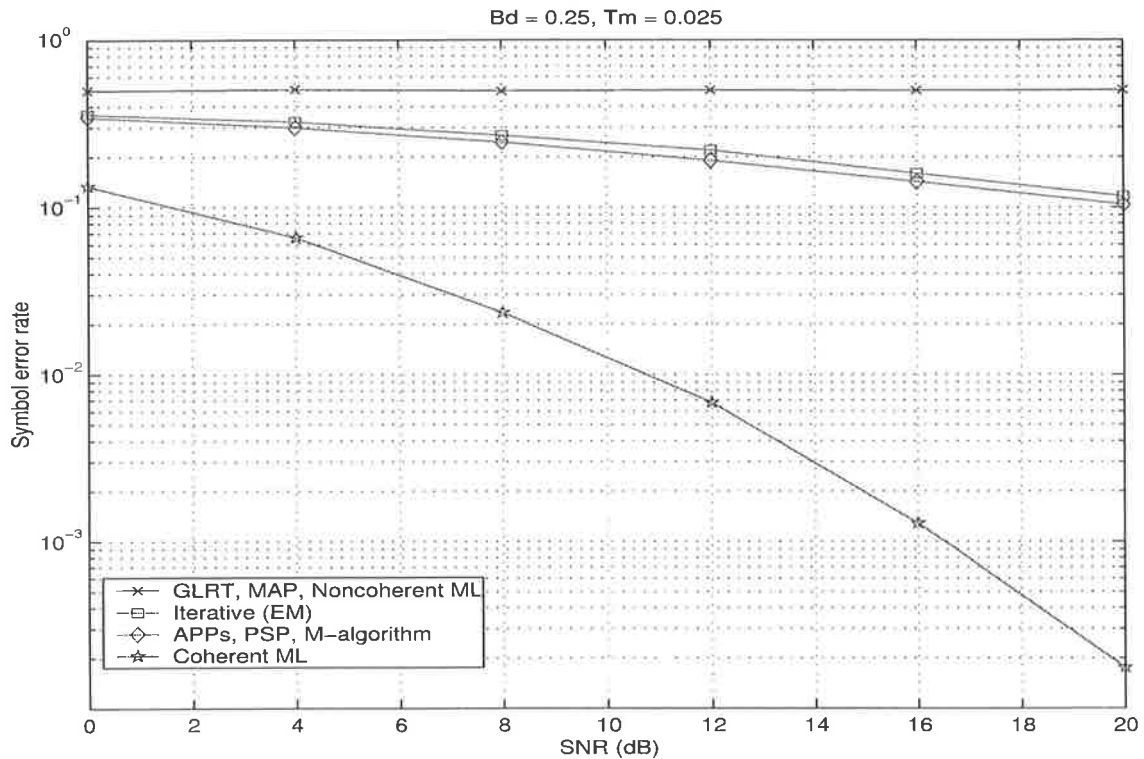
Figure 6.5: Codes $= \pm F$ with $B_d T_s = 0.25$ and $T_m F_s = 0.025$

rather than the error propagation problem since it has ideal CSI. This suggests that using codes which are scalar multiples of each other is not optimal for noncoherent detection.

To prevent the above problem, we chose two codes which are not scalar multiple of each other. The two codes are F and UF matrices where U is an arbitrarily unitary matrix. Fig. 6.6 shows the performance of the receivers using the new codebook for the fading channel with $B_d T_s = 0.25$ and $T_m F_s = 0.5$. The GLRT receiver still has the $SER = 0.5$ for this case since the orthogonal projector $\mathbf{P}_k^\perp$ to the codeword $\mathbf{X}^{(k)}$ is equal to zero for both codewords. The performance of the MAP and the NCML receivers are still very poor. However with this set of codes, the proposed receivers perform very well. At the $SER = 10^{-3}$ the proposed techniques obtained a performance which is within 4dB of the coherent ML receiver. Unlike
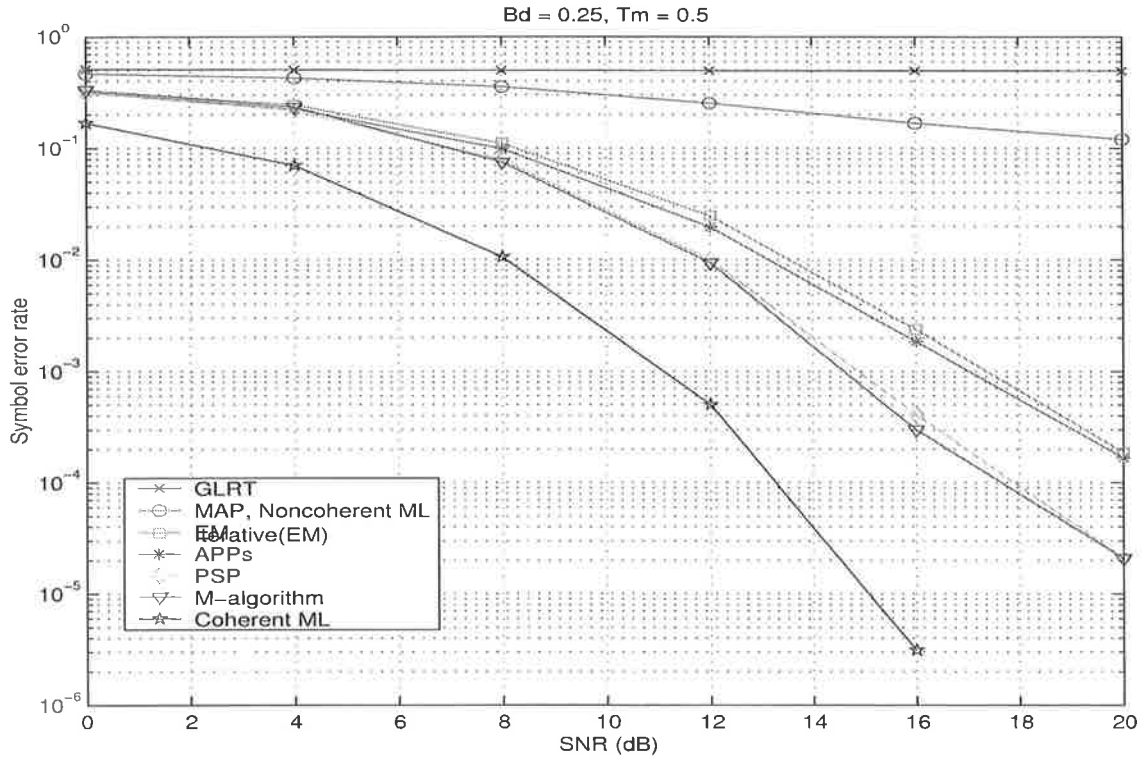
Figure 6.6: Codes $= F$ and UF with $B_dT_s = 0.25$ and $T_mF_s = 0.5$

the former set of codes, using these new codes reduces the performance gap between the proposed receivers and the coherent ML receiver when the fading channels are highly frequency correlated as shown in Fig. 6.7 for $T_mF_s = 0.025$. We also note an interesting observation that the performance of the MAP and NCML is improving as the fading channels are more correlated. This can be explained by examining the decision metrics of these receivers.

The simulation results for all the scenarios considered above confirm the superiority of the proposed receivers to other standard noncoherent receivers. They demonstrate the usefulness of exploiting the temporal and diversity correlations in the channels at the receivers. Simulations above show that even in fast time-varying channels where the temporal correlation is low, by exploiting such correlation could significant improve the receiver's performance. By utilising the temporal correlation,
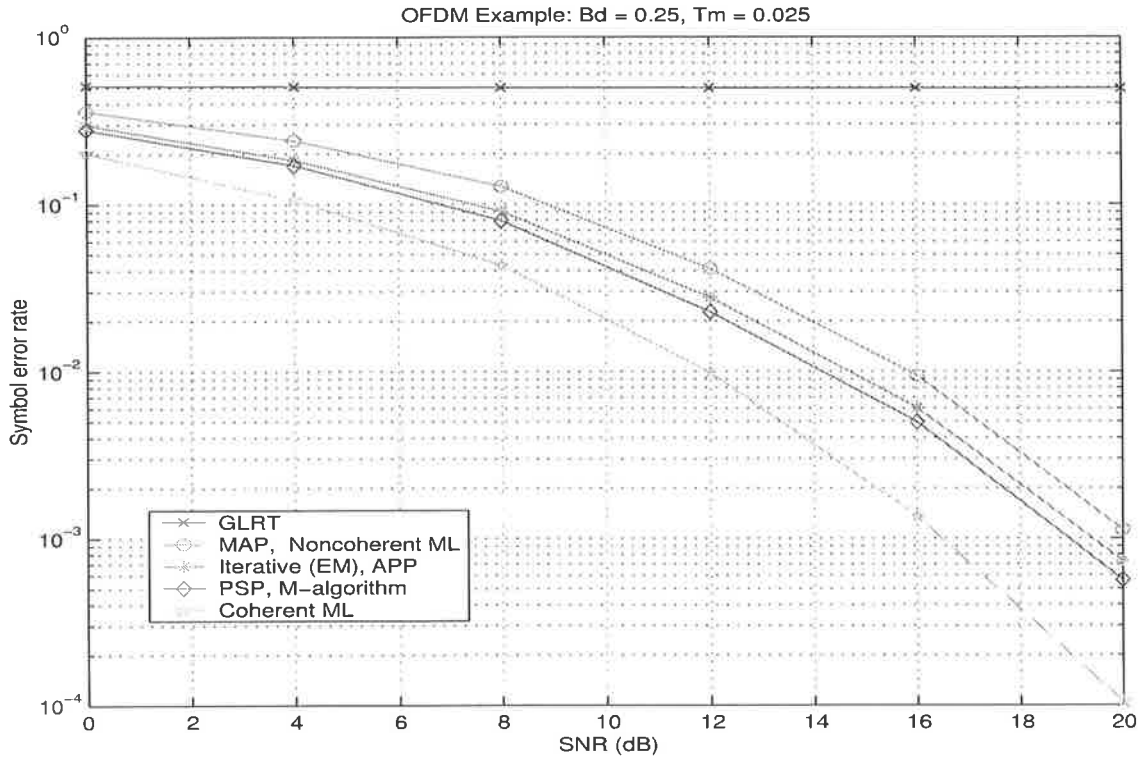
Figure 6.7: Codes $= F$ and UF with $B_d T_s = 0.25$ and $T_m F_s = 0.025$

the proposed receivers obtain a performance which is within few dB of the coherent receiver.

The above observations are pertinent from the point of code design for incoherent systems with temporal channel correlations. It appears that an approach consisting of modulation followed by a fixed diversity code will perform poorly. Such an observation is supported by previous work such as [14] as an example which does not seal with such correlation. We believe there is a deeper issue at hand here when temporal correlation is included in the signal model.

6.6   Conclusions

We have introduced a general state space model for a general diversity communications system with time correlated flat fading. Examples of diversity systems which fall within this framework include space-time coded systems, orthogonal frequency division multiplex (OFDM) systems, code division multiple access systems and hybrids of these systems. The model permits the design of a number of time recursive noncoherent receivers based either on sequence estimation or on symbol by symbol estimation. The receivers considered include Per-survivor processing, M-algorithm, two *a posteriori* probability techniques and a per-symbol iterative technique based on the EM algorithm. As an example of the utility of the approach, we examine an OFDM system model based on a physical delay doppler spread propagation medium, and we have shown how to approximate the resulting channel statistics by our model. The various receivers are compared to conventional designs which do not exploit the channel time correlations. These simulations have suggested that there can be significant gains in performance by incorporating time correlation into the signal model and the resulting receiver designs. Some implications for diversity code design in the noncoherent case are also briefly noted.

6.7   Appendix: Derivation of The Reduced State Model

In this appendix, we derived the reduced state model for diversity systems which do not have full diversity gain.

Using the eigenvalue decomposition, we can express the channel covariance $\tilde{\mathbf{P}} =$

$\tilde{\Phi} \, \Delta \, \tilde{\Phi}^H$ where $\tilde{\Phi}$ is unitary with $\tilde{\Phi} \, \tilde{\Phi}^H = \tilde{\Phi}^H \, \tilde{\Phi} = \mathbf{I}$, and Δ is diagonal of the form

$$\Lambda = \begin{bmatrix} \overline{\Lambda} & 0 \\ 0 & 0 \end{bmatrix} , \tag{6.54}$$

Let $\tilde{\mathbf{g}}_t = \tilde{\Phi}^H \, \tilde{\mathbf{h}}_t$, this implies:

- The covariance of $\tilde{\mathbf{g}}_t$, $Cov\{\tilde{\mathbf{g}}_t\} = \Delta$,

- If we partition $\tilde{\mathbf{g}}_t$ into the form of

$$\tilde{\mathbf{g}}_t = \begin{bmatrix} \overline{\mathbf{h}}_t \\ \hat{\mathbf{h}}_t \end{bmatrix} , \tag{6.55}$$

  where $\overline{\mathbf{h}}_t \in \mathbb{C}^{\overline{L} \times 1}$ and $\hat{\mathbf{h}}_t \in \mathbb{C}^{(L - \overline{L}) \times 1}$, then the corresponding covariances are: $Cov\{\overline{\mathbf{h}}_t\} = \overline{\Delta}$, $Cov\{\hat{\mathbf{h}}_t\} = \mathbf{0}$, and $Cov\{\overline{\mathbf{h}}_t, \hat{\mathbf{h}}_t\} = \mathbf{0}$.

- Since $Cov\{\hat{\mathbf{h}}_t\} = \mathbf{0}$, this implies $\hat{\mathbf{h}}_t$ is a constant. Thus, if $\tilde{\mathbf{h}}_t$ has zero mean, $\tilde{\mathbf{g}}_t$ will also has zero mean and hence $\hat{\mathbf{h}}_t = \mathbf{0}$. The resulting $\tilde{\mathbf{g}}_t$ can be then represented by

$$\tilde{\mathbf{g}}_t = \begin{bmatrix} \overline{\mathbf{h}}_t \\ \mathbf{0} \end{bmatrix} , \tag{6.56}$$

The measurement equation from (6.1) can be now expressed in term of $\overline{h}_t$ as:

$$\begin{aligned} \tilde{\mathbf{y}}_t &= \tilde{\mathbf{X}}_t \, \tilde{\mathbf{h}}_t + \tilde{\mathbf{n}}_t , \\ &= \tilde{\mathbf{X}}_t \, \tilde{\Phi} \, \tilde{\mathbf{g}}_t + \tilde{\mathbf{n}}_t , \\ &= \tilde{\mathbf{X}}_t \begin{bmatrix} \overline{\Phi} & \hat{\Phi} \end{bmatrix} \begin{bmatrix} \overline{\mathbf{h}}_t \\ \mathbf{0} \end{bmatrix} + \tilde{\mathbf{n}}_t , \\ &= \tilde{\mathbf{X}}_t \, \overline{\Phi} \, \overline{\mathbf{h}}_t + \tilde{\mathbf{n}}_t . \end{aligned} \tag{6.57}$$

The state equation of $\overline{\mathbf{h}}_t$ can be derive as follow:

$$\tilde{\mathbf{g}}_{t+1} = \Phi^H \, \tilde{\mathbf{h}}_{t+1} \, ,$$

$$= \Phi^H \left(\mathbf{A}\tilde{\mathbf{h}}_t + \tilde{\mathbf{n}}_t \right) \, ,$$

$$= \Phi^H \, \mathbf{A} \, \Phi\tilde{\mathbf{g}}_t + \Phi^H \, \tilde{\mathbf{v}}_t \, ,$$

$$\begin{bmatrix} \overline{\mathbf{h}}_{t+1} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \overline{\Phi}^H \mathbf{A}\overline{\Phi} & \overline{\Phi}^H \mathbf{A}\hat{\Phi} \\ \hat{\Phi}^H \mathbf{A}\overline{\Phi} & \hat{\Phi}^H \mathbf{A}\hat{\Phi} \end{bmatrix} \begin{bmatrix} \overline{\mathbf{h}}_{t+1} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \overline{\Phi}^H \tilde{\mathbf{v}}_t \\ \hat{\Phi}^H \tilde{\mathbf{v}}_t \end{bmatrix} \tag{6.58}$$

Thus,

$$\overline{\mathbf{h}}_{t+1} = \overline{\Phi}^H \, \mathbf{A} \, \overline{\Phi} \, \overline{\mathbf{h}}_t \, + \, \overline{\Phi}^H \tilde{\mathbf{v}}_t \, . \tag{6.59}$$

Equations (6.57) and (6.59) therefore represent the reduced state model of (6.1) and (6.5), respectively.

Chapter 7

Conclusions

7.1   Summary of Contributions

In this thesis, we have developed signal processing and space-time coding techniques to overcome the multiple access interference and channel fading impairments in wireless communication systems. We focused on the use of interference cancellation methods to mitigate the multiple access interference in CDMA systems, and the use of multiple transmit and multiple receive antennas for providing diversity to combat the effects of channel fading. The major contributions of the thesis are listed below:

Chapter 2: Multiuser Detectors for CDMA Systems

- Two new interference cancellation detectors, which are hybrid of the successive and parallel interference cancellation detectors, are proposed.

- We developed an adaptive multiuser detector which performs joint parameter estimation and symbol detection for CDMA systems.

154

Chapter 3: Convergence Behaviour Analysis of the PIC Technique

- We developed a general framework for analysing the convergence behaviour of the PIC detector for any type of tentative decision function.

- We derived general conditions from which the sufficient condition for convergence of the PIC detector for a wide range of tentative decision functions can be calculated. Several well-known conditions for convergence of the PIC detector with linear decisions and clip decisions can be obtained from this general framework.

- We proved that the PIC detector with any tentative decision function that is monotonically increasing at sublinear rate will either converge to a fixed point or enter a limit cycle of period-two.

Chapter 4: Space-Time Coding

- We derived two new upper bounds for the pairwise error probability of space-time coded systems with multiple transmit and multiple receive antennas in the presence of spatially correlated fading.

- We found new QPSK space-time trellis codes for two transmit antennas based on the design criterion of minimising the sum of the the pairwise error probability of all distinct pairs of codewords. Simulation results showed that these new space-time trellis codes are superior to other known codes.

Chapter 5: Joint Space-Time Trellis Decoding and Channel Estimation

- We showed how the physical model of the MIMO wireless fading channels with multiple transmit and multiple receive antennas can be approximated by a statistical channel model.

- We developed a space-time receiver for joint channel estimation and space-time trellis decoding in spatially correlated time-varying Rayleigh fading channels.

Chapter 6: Recursive Receivers for General Diversity Channels

- We introduced a general state space model for a general diversity communication system with time and diversity correlated flat fading.

- We developed a number of time-recursive receivers which exploit the temporal correlation in the fading channels.

7.2   Suggestions for Further Study

The works that we presented on the area of multiuser detection mainly assume that the CDMA systems are synchronous and there is no chip interference. However in a practical CDMA system, users transmit information independently and hence the signal of different users will arrive asynchronously at the receiver. In addition, since the chips are sent at a much faster rate than the symbol rate, it is likely to yield interchip interference and neighbouring symbol interference. Therefore it would be of practical interest to include those issues into our works. In addition, the convergence behaviour analysis in this thesis is only applied to the PIC detector that performs total interference cancellation. It would be of further research interest to investigate the convergence behaviour of the PIC detector that performs partial interference cancellation.

So far, all the works presented on the area of space-time coding have mainly focused on the narrowband flat fading case where only spatial diversity is available. However, future wireless communication systems will transmit information with symbol duration much smaller than the channel delay spread and consequently frequency-selectivity arises. In [90], theoretical analysis reveals that space-time codes designed for flat fading still provides at least the same diversity advantage even

in frequency-selective channels. Coding advantage however might decrease considerably due to the presence of intersymbol interference unless additional processing is employed [38]. Future work will look at developing space-time coding schemes for frequency selective channels that can exploit both spatial and multipath diversity. It will focus at coding schemes which allow low decoding complexity at the receiver while at the same time achieve full diversity advantage. Low decoding complexity is quite essential as mobile terminals are supposed to be small, light weight and low cost. The main technical challenge for designing space-time codes in frequency-selective fading channels is that signals are mixed both in space (due to the multiple transmit antennas) and time (due to the multipaths). Thus, optimal space-time coding for dispersive multipath channels is very complex and requires highly complex processing at the receiver. Since mobile terminals are required to small, light weight and low cost, suboptimal approaches which allow simple processing at the receivers would be needed. One approach is to combine space-time coding with orthogonal frequency division multiplexing (OFDM) which converts the frequency-selective fading channels into a set of flat fading channels. Another approach is to combine space-time coding with single carrier frequency domain equalization (SC-FDE) technique. Both OFDM and SC-FDE approaches have the advantage of lower processing complexity than the single carrier time domain equalization counterpart [34]. In [5], space-time codes are combined with OFDM to provide high data rate transmission over broadband channels. However, the codes being used are constructed for flat fading. Hence, it can only exploit the spatial diversity while fails to exploit the multipath diversity. In [8, 113], space-time block coding schemes based on single carrier frequency domain equalization approach are proposed for frequency selective fading channels. However, they follow the Alamouti's space-time block coding scheme and are restricted to only two transmit antennas case. It is of interest to seek for more general space-time codes that can exploit the full diversity advantage

using the OFDM and SC-FDE approaches.

Bibliography

[1] A. Abdi, J. A. Barger and M. Kaveh, "A parametric model for the distribution of the angle of arrival and the associated correlation function and power spectrum at the mobile station", *IEEE Trans. Veh. Technol.*, vol. 51, no. 3, pp. 425-434, May 2002.

[2] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels", *IEEE J. Select. Areas Commun.*, vol. 20, no. 3, pp.550-560, Apr. 2002.

[3] F. Adachi, M. Sawahashi and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems", *IEEE Commun. Magz.*, pp. 56-69, Sept. 1998.

[4] D. Agrawal, T. J. Richardson and R. Urbanke, "Multiple-antenna signal constellations for fading channels", *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2618-2626, Sept. 2001.

[5] D. Agrawal, V. Tarokh, A. Naguib and N. Seshadri, "Space-time OFDM for high data-rate wireless communication over wideband channels", *Proc. Veh. Technol. Conf.*, pp. 2232-2236, ON, Canada, May 1998.

[6] K. Aktas and M. P. Fitz, "Computing the distance spectrum of space-time trellis codes", *Proc. WCNC'00*, Chicago IL, Sept. 2000.

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

[8] N. Al-Dhahir, "Single-carrier frequency domain equalization for space-time block coded transmission over frequency-selective fading channels", *IEEE Commun. Letters*, vol. 5, no. 7, pp. 304-306, Jul. 2001.

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

[10] C. Anton-Haro, J. A. R. Fonollosa and J. R. Fonollosa, "Blind channel estimation and data detection using Hidden Markov Model. *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 241-247, Jan. 1997.

[11] K. E. Baddour and N. C. Beaulieu, "Autoregressive models for fading channel simulation", *Proc. GLOBECOM 2001*, Available at http://www.it.iitb.ac.in/ it612/resources/repository/GLOBECOM01/vol12/

[12] S. Baro, G. Bauch and A. Hansmann, "Improved codes for space-time trellis coded modulation", *IEEE Commun. Lett.*, Vol. 4, No. 1, pp. 20-22, Jan. 2000.

[13] Q. Bi, G. I. Zysman and H. Menkes, "Wireless mobile communications at the start of the 21st century", *IEEE Commun. Magz.*, pp. 110-116, Jan. 2001.

[14] M. Brehler and M. K. Varanasi, "Asymptotic error probability analysis of quadratic receivers in rayleigh fading channels with application to a unified analysis of coherent and noncoherent space-time receivers", *IEEE Trans. Inform. Theory*, vol. 47, No. 6, pp. 2383-2399, Sept. 2001.

[15] D. R. Brown III, M. Motani, V. V. Veeravalli, H. V. Poor and C. R. Johnson, Jr., "On the performance of linear parallel interference cancellation", *IEEE Trans. Inform. Theory*, vol. 47, no. 5, pp. 1957-1970, July 2001.

[16] J. Bruck and J. W. Goodman, "A generalized convergence theorem for neural networks", *IEEE Trans. Inform. Theory*, vol. 34, no. 5, pp. 1089-1092, Sept. 1988.

[17] R. M. Buehrer, S. P. Nicoloso and S. Gollamudi, "Linear versus non-linear interference cancellation", *IEICE J. on Commun. Networks*, vol. 1, no. 2, pp. 118-133, June 1999.

[18] R. M. Buehrer, "On the convergence of multistage interference cancelation" Conf. Rec. 33rd Asilomar Conf. Signals, Systems and Compuers, vol. 1, pp. 634-638, Pacific Grove, CA, Oct. 24-27, 1999.

[19] CellularOnline, "Latest global, handset, base station, & regional cellular statistic", Jan. 2003, available at http://www.cellular.co.za.

[20] T-A. Chen, M. P. Fitz, W-Y. Kuo, M. D. Zoltowski and J. H. Grim, "A space-time model for frequency nonselective rayleigh fading channels with application to space-time modems", *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1175-1190, July 2000.

[21] Z. Chen, J. Yuan and B. Vucetic, "Improved space-time trellis coded modulation scheme on slow Rayleigh fading channels", *Electronics Letters*, Vol. 37, No. 7, pp. 440-441, March 2001.

[22] G. Colman, S. D. Blostein and N. C. Beaulieu, "An ARMA multipath fading simulator", *Proc. 7th Annual Virginia Tech. Symposium on Wireless Personal Commun.*, Blacksburg, VA, Jun. 11-13, 1997.

[23] C. Cozzo and B. L. Hughes, "Joint channel estimation and data symbol detection in space-time communications", in *Proc. Int. Conf. on Communications*, New Orleans, LA, June 18-22, 2000, pp. 287-291.

[24] C. Cozzo and B. L. Hughes, "An adaptive receiver for space-time trellis codes based on per-survivor processing", *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1213-1216, Aug. 2002.

[25] Q. Dai and E. Shwedyk, "Detection of bandlimited signals over frequency selective rayleigh fading channels", *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 941-950, Feb./Mar./Apr. 1994.

[26] E. Dahlman, B. Gudmundson, M. Nilsson and J. Skold, "UMTS/IMT-2000 Based on wideband CDMA", *IEEE Commun. Magz.*, pp. 70-80, Sept. 1998.

[27] R. T. Derryberry, S. D. Gray, D. M. Ionescu, G. Mandyam and B. Raghothaman, "Transmit diversity in 3G CDMA systems", *IEEE Commun. Magz.*, pp. 68-75, Apr. 2002.

[28] C. A. Desoer and M. Vidyasager *Feedback Systems : Input-Output Properties*, Academic Press, 1975.

[29] D. Divsalar and M. Simon, "Improved CDMA performance using parallel interference cancellation", *IEEE MILCOM.*, New York, USA, 1994, vol. 3, pp. 911-917.

[30] D. Divsalar, M. Simon and D. Raphaeli, " A new approach to parallel interference cancellation for CDMA. *IEEE GLOBECOM*, New York, NY, USA; 1996 vol. 3, pp. 1452-1457.

[31] D. Divsalar, M. Simon and D. Raphaeli, "Improved parallel interference can-
cellation for CDMA", *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 258-268, Feb.
1998.

[32] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for syn-
chronous code-division multiple access channel", *IEEE Trans. Commun.*, vol.
41, no. 2, pp. 285-290, Feb. 1993.

[33] A. Duel-Hallen, J. Holtzman and Z. Zvonar, " Multiuser detection for CDMA
systems", *IEEE Personal Communications*, vol. 2, no. 2, pp. 46-58, April 1995.

[34] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson, "Fre-
quency domain equalization for single-carrier broadband wireless channels",
IEEE Commun. Magz., pp. 58-66, Apr. 2002.

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

[36] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a
fading environment when using multiple antennas", *Wireless Personal Commu-
nications*, vol. 6, no. 3, pp. 311-335, Mar. 1998.

[37] J. Fuhl, J-P. Rossi and E. Bonek, "High-resolution 3-D direction of arrival
determination for urban mobile radio", *IEEE Trans. Antennas Propagat.*, vol.
45, no. 4, pp. 672-682, Apr. 1997.

[38] Y. Gong and K. B. Lataief, "Performance evaluation and analysis of space-time
coding in unequalised multipath fading links", *IEEE Trans. Commun.*, vol. 48,
no. 11, pp. 1778-1782, Nov. 2000.

[39] A. Grant and C. Schlegel, "Convergence of linear interference cancellation multiuser receiver", *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1824-1834, Oct. 2001.

[40] S. D. Gray, M. Kocic and D. Brady, "Multiuser detection in mismatched multiple-access channels", *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 3080-3089, Dec. 1995.

[41] J. C. Guey, M. P. Fitz, M. R. Bell and W. Y. Kuo, "Signal design for transmitter diversity wireless systems over Rayleigh fading channels", *Proc. IEEE Vehicular Technology Conference*, pp.135-140, Atlanta, US, 1996.

[42] T. Hashimoto, "A list-type reduced-constraint generalization of the Viterbi algorithm", *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 866-876, Nov. 1987.

[43] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communications in rayleigh flat fading", *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 543-564, Mar. 2000.

[44] B. M. Hochwald, T. L. Marzetta, T. J. Richardson, W. Sweldens and R. Urbanke, "Systematic design of unitary space-time constellations", *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 1962-1973, Sept. 2000.

[45] B. M. Hochwald and W. Sweldens, "Differential unitary space-time modulation", *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2041-2052, Dec. 2000.

[46] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons", *Proc. Natl. Acad. Sci. USA*, vol. 81, pp. 3088-3092, May 1984.

[47] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[48] B. L. Hughes, "Differential space-time modulation", *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2567-2578, Nov. 2000.

[49] E. Jaffrot, M. Siala and I. Fijalkow, "Maximum a posteriori semi-blind channel estimation for OFDM systems operating on highly frequency and time selective channels", submitted to *IEEE Trans. Communications*, February 2002.

[50] E. Jaffrot, V. K. Nguyen, M. Soamiadana, L. B. White, and I. Fijalkow, "Symbol by symbol reduced complexity highly selective OFDM channel estimation", Proc. EUSIPCO 2002, Toulouse, France, September 2002.

[51] W. C. Jakes Jr. (Ed.), *Microwave Mobile Communications*, New York : Wiley, 1974.

[52] C. R. Johnson, Jr., *et al*, "Blind equalisation using the constant modulus criterion : A review", *Proc. IEEE*, vol. 86, no. 10, Oct. 1998, pp. 1927-1949.

[53] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown dispersive channels", *IEEE Trans. Commun.*, vol. 42, No. 7, pp. 2406-2413, July 1994.

[54] G. Kang, P. Zhang, H. Haas and E. Schulz, "Good space-time codes in terms of distance spectrum", *Proc. VTC Fall-2002*, Vancouver, Cannada, pp. 252-255, Sept. 2002.

[55] S. M. Kay, *Fundamentals of statistical signal processing: Estimation Theory*, Prentice-Hall signal processing series, 1993.

[56] G. I. Kechriotis and E. S. Manolakos, "Hopfield neural network implementation of the optimal CDMA multiuser detector", *IEEE Trans. Neural Networks*, vol. 7, no. 1, pp. 131-141, Jan. 1996.

[57] C. Komninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Adaptive multi-input multi-output fading channel equalization using kalman estimation", Proc. ICC 2000, pp. 1655-1659, New Orleans, Louisiana, Jun. 18-22, 2000.

[58] D. N. Knisely, Q. Li and N. S. Ramesh, "cdma2000: A third-generation radio transmission technology", *Bell Labs Tech. Journal*, pp. 63-78, Jul.-Sept. 1998.

[59] V. Krishnamurthy and J. B. Moore, "On-line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure", *IEEE Trans. Signal Process.*, vol. 41, no. 8, pp. 1557-1572, Aug. 1993.

[60] Z. Liu, X. Ma, G. B. Giannakis, "Space-time coding and kalman filtering for time-selective fading channels", *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183-186, Feb. 2002.

[61] Z. Liu, G. B. Giannakis and B. L. Hughes, "Double differential space-time block coding for time-selective fading channels", *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1529-1539, Sept. 2001.

[62] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels", *IEEE Trans. Inform. Theory*, vol. 35, no. 1, pp. 123-136, Jan. 1989.

[63] U. Madhow and M. Honig, "MMSE interference suppression for direct sequence spread spectrum CDMA", *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3178-3188, Dec. 1994.

[64] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with applications in statistics and econometrics*, New York: McGraw-Hill, 1999.

[65] L. Mailaender and R. A. Iltis, "Multiuser detectors with single user parameter estimation on Quasi-Synchronous CDMA channels", *IEEE Trans. Commun.*, vol. 48, no. 2, pp. 200-203, Feb. 2000.

[66] C. M. Marcus and R. M. Westervelt, "Dynamics of iterated-map neural networks", *Physical Review A*, vol. 40, no. 1, pp. 501-504, July 1989.

[67] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple antenna communication link in rayleigh flat fading", *IEEE Trans. Inform. Theory*, vol. 45, no. 3, pp. 139-157, Jan. 1999.

[68] M. Motani and D. R. Brown, "On the convergence of linear parallel interference cancellation", *ISIT 2001, Washington DC*, June 24-29, 2001.

[69] R. D. Murch and K. B. Letaief, "Antenna system for broadband wireless access", *IEEE Commun. Magz.*, pp. 76-83, Apr. 2002.

[70] V. K. Nguyen and L. B. White, "Iterative multiuser detection with parameter estimation", *Digital Signal Processing*, vol. 12, no. 2,3, pp. 145-158, Apr./July 2002.

[71] V. K. Nguyen and L. B. White, "Interference cancellation schemes for CDMA systems", *Information, Decision and Control 2002 Conf.*, Adelaide, Feb. 2002.

[72] V. K. Nguyen and L. B. White, E. Jaffrot, M. Soamiadana, I. Fijalkow, "Recursive receiver structures for general diversity channels with time correlated flat fadings", *IEEE J. Select. Areas Commun.*, vol. 21, no. 5, pp. 754-764, June 2003.

[73] V. K. Nguyen and L. B. White, "Recursive receiver structures for general diversity channels with time correlated flat fadings", *Proc. Third Australian Communications Theory Workshop*, Canberra, Australia, pp. 48-52, Feb. 2002.

[74] M. J. Omidi, S. Pasupathy, P. G. Gulak, "Joint data and Kalman estimation of fading channel using a generalized viterbi algorithm", *International Conf. on Commun.* pp. 1198-1203, June 1996.

[75] P. Patel and J. Holtzman, "Analysis of a simple successive interference cancellation scheme in a DS/CDMA system", *IEEE J. Select. Areas Commun.*, vol. 12, No. 5 pp. 796-807, June 1994.

[76] S. Perreau and L. B. White, "Nonlinear iterative multiuser detection and equalization for CDMA receivers in the presence of interchip interference", *Digital Signal Processing*, vol. 11, no. 2, pp. 94-109, Apr. 2001.

[77] E. Pittampalli, "Third-generation CDMA wireless standards and harmonization", *Bell Labs Tech. Journal*, pp. 6-18, Jul.-Sept. 1999.

[78] M. B. Priestley, *Spectral Analysis and Time-Series*, Academic Press, London, 1981.

[79] R. Raheli, A. Polydoros, C. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environments", *IEEE Trans. Commun.*, vol. 43, pp. 354-364, Feb./Mar./Apr. 1995.

[80] T. S. Rappaport, *Wireless Communications: Princles and Practice*, Prentice Hall, New Jersey, 1996.

[81] L. K. Rasmussen, T. J. Lim and A. Johansson, "A matrix-algrebraic approach to successive interference cancellation in CDMA", *IEEE Trans. Commun.*, vol. 48, no. 1, pp. 145-151, Jan. 2000.

[82] E. B. Saff and A. D. Snider, *Fundamentals of complex analysis for mathematics, science, and engineering.* New Jersey: Prentice-Hall, 1976.

[83] M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*. New York: McGraw-Hill, 1966.

[84] N. Seshadri and J. H. Winters, "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity", *Int. J. Wireless Inform. Networks*, vol. 1, no. 1, Jan. 1994.

[85] N. Sellami, I. Fijalkow and M. Siala, "Low-complexity iterative receiver for space-time coded signals over frequency selective channels", *EURASIP Journal on Applied Signal Processing*, special issue on space-time coding and its applications, No. 5, pp.517-524, May 2002.

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

[87] S. Siwamogsatham, F. P. Fitz and J. H. Grimm, "A new view of performance analysis of transmit diversity schemes in correlated Rayleigh fading", *IEEE Trans. Inform. Theory*, Vol. 48, No. 4, pp. 950-956, April 2002.

[88] F. Swarts and H. C. Ferreira, "Markov characterization of channels with soft decision outputs", *IEEE Trans. Commun.*, vol. 41, no. 5, pp. 678-682, May 1993.

[89] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction", *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.

[90] V. Tarokh, A. Naguib, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility and multipaths", *IEEE Trans. Commun.*, vol. 47, no. 2, pp. 199-207, Feb. 1999.

[91] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456-1467, Jul. 1999.

[92] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity", *IEEE J. Select. Areas Commun.*, vol. 18, no. 7, pp. 1169-1174, Jul. 2000.

[93] I. E. Telatar, "Capacity of multi-antenna Gaussian channels", Technical Report, AT&T Bell Laboratories, Lucent Technologies, 1995.

[94] D. M. Titterington, "Recursive Parameter Estimation using Incomplete Data", *J. R. Statist. Soc. B*, vol. 46, no. 2, pp. 257-267, 1984.

[95] M. K. Tsatsanis, G. B. Giannakis and G. Zhou, "Estimation and equalization of fading channels with random coefficients", *Signal Processing* 53, pp. 211-229, 1996.

[96] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: effective interference, effective bandwidth and user capacity", *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 641-657, Mar. 1999.

[97] M. K. Varanasi and B. Aazhang, " Multistage detection in asynchronous code-division multiple-access communications", *IEEE Trans. Commun.*, vol. 38, no. 4, pp. 509-519, Apr. 1990.

[98] M. K. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems", *IEEE Trans. Commun.*, vol. 39, no. 5, pp. 725-736, May 1991.

[99] S. Verdu, "Minimum probability of error for asynchronous gaussian multiple-access channels", *IEEE Trans. Inform. Theory*, vol. 32, no. 1, pp. 85-96, Jan. 1986.

[100] S. Verdu *Multiuser Detection*, Cambridge University Press, 1998.

[101] E. Visotsky and U. Madhow, "Noncoherent multiuser detection for CDMA systems with nonlinear modulation: a non-Bayesian approach", *IEEE Trans. Information Theory*, vol. 47, no. 4, pp. 1352-1367, May 2001.

[102] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels", *IEEE J. Select. Areas Commun.*, vol. 8, No. 4, pp. 641-649, May 1990.

[103] H. S. Wang and N. Moayeri, "Finite-state Markov channel - a useful model for radio communication channels", *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163-171, Feb. 1995.

[104] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication". *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 651-668, Mar. 2002.

[105] F. R. Waugh and R. M. Westervelt, "Analog neural networks with local competition. I. Dynamics and stability", *Physical Review E*, vol. 47, no. 6, pp. 4524-4536, June 1993.

[106] Y. Xue and X. Zhu, "PSP decoder for space-time trellis code based on accelerated self-tuning LMS algorithm", *Electronic Letters*, vol. 36, no. 17, pp.1472-1474, Aug. 2000.

[107] Q. Yan and R. S. Blum, "Optimum space-time convolutional codes", *Proc. IEEE WCNC'00*, Chicago, IL, pp. 1351-1355, Sep. 2000.

[108] A. Yener, R. D. Yates and S. Ulukus, "CDMA multiuser detection: a nonlinear programming approach", *IEEE Trans. Commun.*, vol. 50, no. 6, pp. 1016-1024, June 2002.

[109] Y. C. Yoon, R. Kohno and H. Imai, "A spread-spectrum multiaccess system with cochannel interference cancellation for multipath fading channels", *IEEE J. Select. Areas Commun.*, vol. 11, no. 7, pp. 1067-1075, Sep. 1993.

[110] J. Yuan, Z. Chen, B. Vucetic and W. Firmanto, "Performance analysis and design of space-time coding on fading channels", Submitted to *IEEE Trans. Commun.*, Sept. 2000.

[111] M. Zeng, A. Annamalai and V. K. Bhargava, "Recent advances in cellular wireless communications", *IEEE Commun. Magz.*, pp. 128-138, Sept. 1999.

[112] L. Zheng, D. N. C. Tse, "Packing spheres into the Grassmann manifold: A geometric approach to noncoherent multi-antenna channels", *IEEE Trans. Inform. Theory*, Submitted for publication.

[113] S. Zhou and G. B. Giannakis, "Space-time coding with maximum diversity gains over frequency-selecitve fading channels", *IEEE Signal Process. Letters*, vol. 8, no. 10, pp. 269-272, Oct. 2001.