Reduced Complexity Detection Techniques for Multi-Antenna Communication Systems

Khawaja Tauseef Tasneem

A thesis submitted for the degree of
Doctor of Philosophy
in
Electrical and Computer Engineering
at the
University of Canterbury,
Christchurch, New Zealand.

June 2013

ABSTRACT

In a multiuser system, several signals are transmitted simultaneously within the same frequency band. This can result in significant improvements both in spectral efficiency and system capacity. However, a detrimental effect of the shared transmissions (both in time and bandwidth), is that the signal received at the base station (BS) or access point (AP) suffers from cochannel interference (CCI) and inter-symbol interference (ISI). This situation presents challenges to receiver design. To combat the destructive nature of multipath fading, a receiver often employs multiple antennas to collect the faded superimposed versions of the transmitted signals. The multiple signals are combined and processed in such a way that the effects of CCI and ISI are minimized and the desired information is reliably recovered. The situation is even more challenging when the system is operating under overload, i.e. when there are fewer receive antennas than there are transmitted signals. Multiuser detection (MUD) is used to simultaneously estimate the information sent by the transmitters. To do this, the receiver exploits differences among the cochannel signals (through unique spatial signatures in this case).

We consider a cochannel communication system where multiple transmitted signals arrive at a receiver (equipped with multiple receive antennas) after propagating through a Rayleigh fading channel. It is assumed that the receiver is operating in an overloaded scenario. For such systems, an optimum maximum a posterior probability (MAP) detector estimates the transmitted signal by maximizing the probability of correct decision. The MAP detector reduces to the maximum likelihood (ML) detector when all the transmitted signals are equiprobable. The computational complexity of both MAP and ML detectors increases exponentially with the number of transmitted signals and the channel memory. For large systems suffering severe CCI and ISI, this is clearly not a

good choice for real-time implementation due to the associated computational expenses. The main factors that influence the complexity of MAP / ML detection are: (i) the number of transmitted signals (or equivalently the number of users sharing the system resources), (ii) modulation alphabet size, and (iii) length of the channel memory. On the other hand, linear detection approaches fail to offer acceptable performance while other nonlinear sub-optimum approaches incur high computational costs for reasonably improved system performance and exhibit an irreducible error-floor at medium to high signal to noise ratio (SNR) values [1–11].

We develop receiver signal processing techniques for the frequency-flat fading channel (where all the multipaths of the transmitted signal arrive at the receiver within a symbol period). We develop an ant colony optimization (ACO) assisted soft iterative detection approach for binary phase-shift keying (BPSK) modulated signals which employs a simplified MAP criteria to extract the most probable signals from the search space. The structure of the receiver is such that it can continue operating under overloaded conditions. The technique achieves near maximum likelihood (ML) performance in critically loaded cases using much lower complexity. For the challenging case of overload it still offers performance close to ML at low to moderate SNR values. Second, an integrated framework comprising of ACO metaheuristic and a recursively defined ML search criteria is developed to handle multilevel modulations. The proposed receiver is capable of achieving near-ML performance for the considered system with significant savings in computational complexity. The receiver framework is independent of the system loading condition, and therefore it remains suitable for overloaded scenarios. Due to the branch and bound nature of the algorithm, an exact expression for the complexity cannot be determined. Instead, an upper bound on computational complexity is developed.

ACKNOWLEDGMENT

First of all, I would like to thank my supervisors Assoc. Prof. Philippa A. Martin and Prof. Desmond P. Taylor for their valuable guidance and consistent support during my studies. Their suggestions and opinions have been proved invaluable and contributed immensely towards my research. I highly acknowledge our discussions ranging from theoretical aspects of communication to algorithm specifications and computer simulations. All this will surely remain with me as a valuable asset for the rest of my professional career.

I would like to express my gratitude to the occupants of Communications Lab during the past years. In particular, I would like to thank Gayathri, Michael, Rui, Zahid and Dushyantha for their help and cooperation during my stay. The interaction with all of them has been very positive and helped me in keeping my spirits high. Special thanks to the friends from Power Electronics Research Group, Bhaba, Thahirah, Michael and Kalyan. I am grateful to all the other friends and colleagues. In particular, James Johnson, Ejaz Khokhar, Kerry Khokhar, Mirza Nasir-ul-Hassan, Abdul Sattar, Ansr Ali, Gulraiz Ahmed and Mohsin Ahmed Shaikh.

I wish to express my appreciation to the departmental staff for their cooperation and the administration of the University of Canterbury for providing tremendous support during and after the earthquakes. My special thanks to everyone at the Campus Health Center for their help and support during the injury times.

Last but not least, I would like to thank my family (specially my parents) for their never ending love, support and patience for all these years.

CONTENTS

ABSTRACT **iii**

ACKNOWLEDGMENT **v**

LIST OF ACRONYMS **xvii**

CHAPTER 1 INTRODUCTION **1**
1.1 Introduction 1
1.2 Background and Motivation 4
1.3 Overview and Contributions 7
1.3.1 Thesis Overview 7
1.3.2 Thesis Contributions 8

CHAPTER 2 COMMUNICATION OVER WIRELESS FADING CHANNELS **11**
2.1 Introduction 11
2.2 The Wireless Channel 12
2.2.1 Multipath Fading 14
2.2.2 Interference 17
2.3 Statistical Modeling of Multipath Channels 17
2.3.1 Rayleigh Fading 18
2.3.2 Rician Fading 18
2.4 Introducing Diversity in Multipath Fading Channels 19
2.4.1 Basic Techniques 21
Frequency Diversity 21
Time Diversity 22
Antenna Diversity 22
2.4.2 Diversity Combining 23
Selection Combining 24
Switched Combining 24
Maximal Ratio Combining 25
Equal Gain Combining 26
2.5 Multiple Access Techniques 26
2.5.1 Frequency Division Multiple Access 27
2.5.2 Time Division Multiple Access 27

2.5.3 Code Division Multiple Access 28
2.6 Spatial Multiplexing System 30
2.6.1 System Model 32
2.6.2 System Loading Conditions 34
2.7 Summary 35

CHAPTER 3 EXISTING DETECTION TECHNIQUES 37
3.1 Introduction 37
3.2 Optimum Detection 39
3.2.1 Maximum A Posteriori Probability Detector 39
3.2.2 Maximum Likelihood Detector 40
3.3 Linear Detection 41
3.3.1 Zero-Forcing Detector 43
3.3.2 Minimum Mean Square Error Detector 45
3.4 Sphere Detector 46
3.5 Successive Interference Cancellation Detector 50
3.6 Group Detector 53
3.7 Genetic Algorithm Based Soft Biased Detector 56
3.8 Ant-Colony Optimization Based Detector 60
3.9 Summary 64

CHAPTER 4 DETECTION OF COCHANNEL SIGNALS USING ANT COLONY OPTIMIZATION 67
4.1 Introduction 67
4.2 System Model 69
4.3 ACO Assisted Detector for Binary Signals 69
4.3.1 The Heuristic Framework 70
4.3.2 Preprocessing 72
4.3.3 Iterative Solution Construction and Refinement 74
Initialization 74
ACO Search Process 75
Fitness Evaluation 77
Soft Signal Estimation 78
Pheromone Update 78
Output Solution 79
4.3.4 Computational Complexity 80
4.3.5 Simulation Results 81
4.4 ACO Assisted Detector for Non-Binary Signals 85
4.4.1 The Heuristic Framework 86
4.4.2 Solution Construction using ACO 87
4.4.3 Pheromone Update 88
4.4.4 Computational Complexity 90
4.4.5 Simulation Results 91
4.5 Summary 96

CHAPTER 5 NEAR-ML DETECTION OF COCHANNEL SIGNALS USING ANT COLONY OPTIMIZATION 99
5.1 Introduction 99
5.2 System Model and Receiver Structure 100
5.3 The Heuristic Framework 101
5.4 Solution Construction using ACO 104
5.4.1 Graph-Based Search Model 104
5.4.2 Tree-Based Search Model 107
5.5 Transition Probability and Pheromone Update 110
5.5.1 Transition Probability 110
5.5.2 Pheromone Update 110
5.6 Higher Order Modulations: The $|\chi|$-ary QAM 114
5.7 Computational Complexity of Proposed Algorithm 115
5.7.1 Complex-to-Real Conversions 116
5.7.2 Upper Bound on Complexity 117
Complexity for QPSK / 4-QAM Modulation 117
Complexity for Higher Order QAM 118
5.8 Simulation Results 119
5.9 Summary 125

CHAPTER 6 CONCLUSION AND FUTURE RESEARCH 129
6.1 Introduction 129
6.2 Conclusion 129
6.3 Future Research Directions 131
6.3.1 Frequency-Selective Channels 131
6.3.2 Coded Systems 131
6.3.3 Joint Channel Estimation and Data Detection 132
6.3.4 Antenna Selection 132
6.3.5 Cooperative Diversity Systems 133

APPENDIX A THE GENETIC ALGORITHM FOR SIGNAL DETECTION IN OVERLOADED COCHANNEL SYSTEMS 135
A.1 Introduction 135
A.2 The Genetic Engine 135
A.2.1 Initialization 136
A.2.2 Fitness Evaluation 137
A.2.3 Selection 137
A.2.4 Crossover 137
A.2.5 Mutation 138
A.2.6 Elitism 138
A.2.7 Termination 138

APPENDIX B A TUTORIAL ON ANT-COLONY OPTIMIZATION 139
B.1 Introduction 139
B.2 Real Ants and the Shortest Path Problem 140
B.3 Combinatorial Optimization 141
B.4 The ACO Metaheuristic 142
B.5 Ant Based Algorithms 143
B.6 Summary 144

REFERENCES 156

LIST OF FIGURES

1.1 System model for a generic time-varying, frequency-selective cochannel communication system. 3

2.1 Block diagram of a baseband wireless communication system. 14

2.2 Multipath phenomena and impulse response of a time-varying wireless channel [12]. 15

2.3 Combiner structure in a diversity aided receiver. 23

2.4 General model of a multiple access communication system. 27

2.5 Channelization scheme for FDMA systems. 28

2.6 Channelization scheme used in TDMA systems. 29

2.7 Channelization for CDMA systems. 29

2.8 Block diagram of a generic time-varying fading cochannel communication system. 31

2.9 Block diagram of a slowly varying, flat-fading cochannel communication system. 34

3.1 Classification of signal detection techniques. 38

3.2 Receiver for linear detection techniques. 43

3.3 Geometric illustration of the sphere decoding algorithm. 48

3.4 Tree generated by the branch and bound SD algorithm. 49

3.5 SIC receiver employing N_T linear receivers and interference cancellers for successive signal estimation. 51

3.6 Group detection based receiver. 54

3.7 Initialization phase of the biased GA detector. 58

4.1 Structure of the soft ACO based receiver 72

4.2 Flowchart of the proposed ACO based receiver for binary signalling. 73

4.3 Problem graph and an illustration of solution construction by the ants. 77

4.4 BER performance of a critically loaded system with $N_T = N_R = 16$ and $A = 500$. 81

4.5 BER versus I, at 13dB for a critically loaded system with $N_T = N_R = 16$ and $A = 100 - 400$. 82

4.6 BER performance of an overloaded system with $N_T = 16$, $N_R = 12$ and $A = 500$. 83

4.7 BER versus I at 13dB for an overloaded system with $N_T = 16, N_R = 12$ and $A = 100 - 500$. 84

4.8 BER performance of an overloaded system with $N_T = 16$, $N_R = 10$ and $A = 500$. 85

4.9 BER performance for a critically loaded system with $N_T = N_R = 8$. 92

4.10 BER performance of $N_T = N_R = 8$ cochannel system at $\gamma = 10$ dB. 93

4.11 BER performance for an overloaded system with $N_T = 8, N_R = 6$. 94

4.12 BER performance of $N_T = 8, N_R = 6$ cochannel system at $\gamma = 12$ dB. 95

4.13 BER performance for an overloaded system with $N_T = 8, N_R = 5$. 96

4.14 BER performance of $N_T = 8, N_R = 5$ cochannel system at $\gamma = 14$ dB. 97

5.1 Receiver structure for the proposed detector. 101

5.2 ACO detection process for 4-QAM modulation. 105

5.3 Heuristic value calculation using PMM. 106

5.4 Tree-based model for the detection problem with QPSK / 4QAM in use. 106

5.5 PMM computation for the $|\chi| = 4$ nodes at the i-th level for QPSK / 4QAM. 107

5.6 Transition probability calculation using the pheromone and the heuristic value for QPSK / 4QAM. 108

5.7 PMM generation on a graph during the detection process for a $N_T = N_R = 4$ system. 111

5.8 Construction of a 16-QAM symbol through addition of two QPSK symbols. 114

5.9 Performance of $N_T = 8$, $N_R = 5$ cochannel system employing QPSK modulation. 121

5.10 Performance of $N_T = 8$, $N_R = 5$ cochannel system at $\gamma = 16dB$. 121

5.11 Performance of $N_T = 8$, $N_R = 4$ cochannel system employing QPSK modulation. 123

5.12 Performance of $N_T = 8$, $N_R = 4$ cochannel system at $\gamma = 18dB$. 123

5.13 Performance of $N_T = 9$, $N_R = 6$ cochannel system employing QPSK modulation. 124

5.14 Performance of $N_T = 9$, $N_R = 6$ cochannel system at $\gamma = 14dB$. 124

5.15 Performance of $N_T = 4$, $N_R = 3$ cochannel system employing 16-QAM modulation. 126

5.16 Performance of $N_T = 4$, $N_R = 2$ cochannel system employing 16-QAM modulation. 126

A.1 Genetic Processing Engine. 136

B.1 Algorithmic framework of the ACO metaheuristic [13]. 143

LIST OF TABLES

4.1 ACO Detection Parameters (Modulation is BPSK for all the cases) 82

4.2 Number of Real Multiplications ($N_T = 16$ in all cases) 85

4.3 Detector parameters used in simulations 91

4.4 Complexity Comparison ($N_T = 8$, $N_R = 8$) 92

4.5 Complexity Comparison ($N_T = 8$, $N_R = 6$) 93

4.6 Complexity Comparison ($N_T = 8$, $N_R = 5$) 95

5.1 Conversion from complex to real operations 116

5.2 Number of Real Multiplications for various Detection Techniques 119

5.3 Number of Real Additions for various Detection Techniques 119

5.4 Detector parameters used in simulations 120

5.5 Complexity Comparison ($N_T = 8, N_R = 5, |\chi| = 4$) 120

5.6 Number of Real Multiplications ($N_T = 4$ in all cases) 125

LIST OF ACRONYMS

AWGN	Additive White Gaussian Noise
ACI	Adjacent Channel Interference
ACO	Ant-Colony Optimization
AMPS	Advanced Mobile Phone System
APP	A Posteriori Probability
ACS	Ant-Colony System
BPSK	Binary Phase-Shift Keying
BER	Bit-Error-Rate
CMM	Complete ML Metric
CCI	Cochannel Interference
CO	Combinatorial Optimization
CDMA	Code-Division Multiple Access
CSI	Channel State Information
DS-CDMA	Direct-Sequence CMDA
EGC	Equal Gain Combining
FDMA	Frequency Division Multiple Access
GA	Genetic Algorithm

GSM	Global System for Mobiles
ISI	Intersymbol Interference
LOS	Line-of-Sight
LLR	Loglikelihood Ratio
MACO	Modified ACO
MMAS	MAX-MIN Ant System
MIMO	Multiple-Input Multiple Output
MRC	Maximum Ratio Combining
MMSE	Minimum Mean-Square Error
MAP	Maximum A Posterior Probability
ML	Maximum Likelihood
MUD	Multiuser Detection
NP-Hard	Non-Deterministic Polynomial-Time Hard
PDF	Probability Density Function
PAM	Pulse-Amplitude Modulation
PSO	Particle Swarm Optimization
PMM	Partial ML Metric
PDA	Probabilistic Data Association
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
SD	Sphere Detection
SRSJD	Spatially Reduced-Search Joint Detection

SER Symbol-Error Rate

SNR Signal-to-Noise Ratio

SISO Single-Input Single-Output

SM Spatial Multiplexing

TDMA Time Division Multiple Access

VLSI Very-Large-Scale Integration

WCDMA Wideband CDMA

ZF Zero-Forcing

Chapter 1

INTRODUCTION

1.1 INTRODUCTION

A typical electronic communication system is composed of an information source, a transmitter, channel, a receiver and the intended destination [14]. The source emits electrical signals which are processed by the transmitter for efficient transmission over the channel. The channel is a medium (e.g. a wire, an optical fiber or a radio link) through which the transmitted signals travel toward the receiver. It can be modeled as a filter which distorts and attenuates the transmitted signal [12]. The receiver is responsible to extract the correct message upon receiving the distorted source signal. To do so, it attempts to cancel the effects of the channel and the signal processing techniques applied at the transmitter (in reverse order).

The maximum error-free information transmission rate supported by a channel is known as channel capacity[1]. The capacity of traditional single-input single-output (SISO) communication systems was first presented in [15] assuming transmission over an additive white Gaussian noise (AWGN) channel. In contrast to AWGN channels, it was later discovered in [16, 17] that multiple antenna channels are capable of increasing the available capacity in proportion to the number of transmit antennas.

In a wireless communication system, the channel or the medium for information transfer is a radio link [18]. Both, the source and destination are equipped with a device called an antenna which radiates the information signal into the air (in the form of electromagnetic waves). The transmitted signal propagates through the radio channel

[1] Channel capacity represents a fundamental constraint on information transmission rate over any communication channel.

and impinges on the receiver antenna. The received signal is also contaminated by AWGN (mainly thermal noise from the receiver front-end). The resulting signal is then passed to the receiver for processing.

The electromagnetic spectrum available for wireless communication application is already limited by nature (and thus very expensive), and is regulated by government agencies. Consequently, the spectrum usage needs to be highly efficient and be allocated in a well-planned manner. In modern wireless communication systems where multiple signals are transmitted simultaneously within the same frequency band, both spectral efficiency and system capacity are improved [16, 19–21]. However, the channel induces the undesirable effects of multipath fading and intersymbol interference (ISI). Additionally, the overlapping transmissions lead to cochannel interference (CCI) as the received signal is a weighted superposition of the transmitted signals perturbed by noise. This presents a challenge to receiver design.

To combat multipath fading, a receiver often employs multiple receive diversity branches to collect the faded superimposed versions of the transmitted signals [1, 22, 23]. The multiple received signals are combined and processed in such a way that the effects of CCI and ISI are minimized.

In this research, we consider a multiple signal cochannel communication system as shown in Fig. 1.1 where a total of N_T signals are transmitted which, after propagating through the wireless fading channel, arrive at the receiver. It is assumed that the receiver employs a total of N_R antennas for collecting the faded superimposed copies of all the transmitted signals. Assuming $x_j(t)$ to be the signal emitted by the j-th transmitter, the received signal at the i-th receive antenna can be expressed as

$$\begin{aligned} y_i(t) &= \sum_{j=1}^{N_T} \int_{-\infty}^{\infty} h_{ij}(\tau, t)\, x_j(t-\tau)\, d\tau + n_i(t), \qquad i = 1, 2 \ldots N_R \\ &= \sum_{j=1}^{N_T} h_{ij}(\tau, t) \star x_j(\tau) + n_i(t), \end{aligned} \tag{1.1}$$

where $\star$ denote convolution and $\{n_i(t)\}$ is the AWGN noise process at i-th receive antenna. The symbol $h_{ij}(\tau, t)$ represents the time-varying channel impulse response

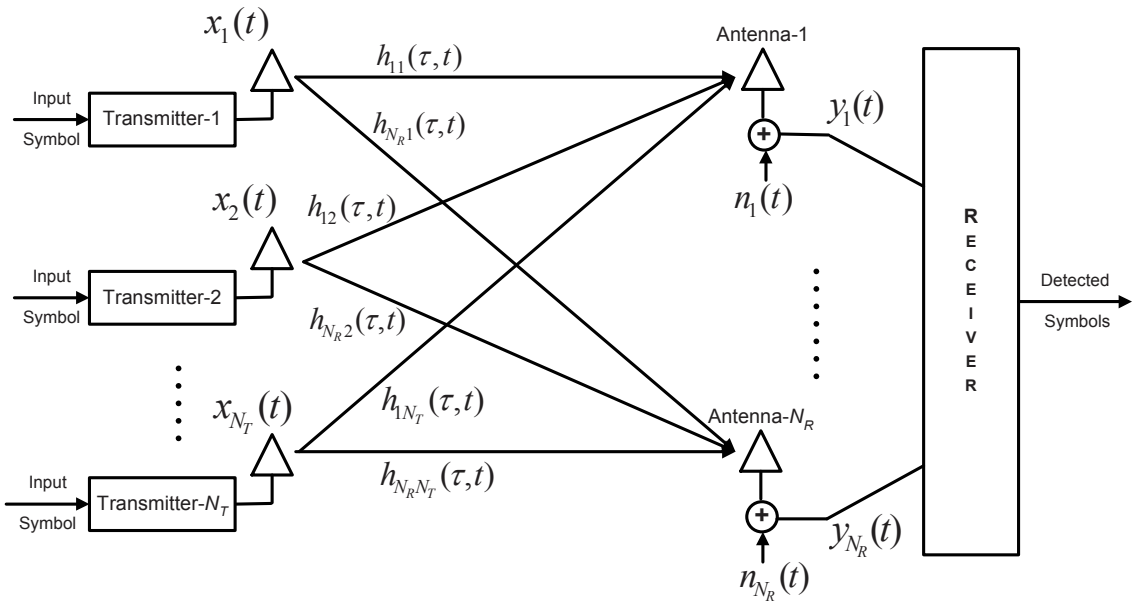

Figure 1.1 System model for a generic time-varying, frequency-selective cochannel communication system.

between the j-th transmitter and the i-th receive antenna (where τ is the multipath delay and t represents time).

A closer look at (1.1) reveals that the i-th received signal is a faded superposition of N_T transmitted signals perturbed by noise. Also there are a total of N_R signals observed at the receiver (one at each receiving antenna). By collecting these signals, a linear system of equations can be formed with N_T unknowns (i.e. the N_T transmitted signals) and N_R equations (i.e. the N_R received signals). The detection problem then aims to estimate the N_T unknowns based on N_R observations.

In critically loaded or underloaded systems, the number of receive antennas are either equal to, or exceed the number of transmitted signals (i.e. $N_R = N_T$ or $N_R > N_T$), respectively. For such systems, the optimum maximum *a posterior* probability (MAP) detector estimates the transmitted signal by minimizing the probability of making a decision error or equivalently it maximizes the probability of a correct decision. The MAP detector becomes the maximum likelihood (ML) detector when all possible transmitted signals are equally likely to be chosen for transmission. Both the MAP and the ML detectors have computational complexity that increases exponentially with the

number of transmitted signals [23, 24]. Due to this reason, a number of low complexity detection techniques are developed in the past offering various performance complexity tradeoffs [4, 25–32].

The situation, however, becomes more challenging when the number of transmitted signals exceeds the number of receive antennas i.e $N_R < N_T$. This condition is referred to as *overloaded*. It is clearly evident from (1.1) that there are more unknowns than equations and thus one have to deal with an underdetermined system. Due to this reason, most of the detection techniques developed for the systems with $N_R \geq N_T$ are no longer applicable.

From the optimum detection viewpoint, it has been shown in [23] that it is still possible to recover the transmitted data, but the receiver performance tends to degrade slowly with each additional signal. For large systems suffering with severe CCI and ISI, the computational complexity makes real-time implementation difficult. The main factors that influence the complexity of MAP/ML detection are: (i) the number of transmitted signals, (ii) the modulation alphabet size, and (iii) the length of the channel memory[2].

The rest of the chapter is organized as follows: Section 1.2 starts with a brief discussion on existing signal detection approaches and presents the motivation behind our work. Section 1.3 presents a thesis overview and a list of publications resulting from the work.

1.2 BACKGROUND AND MOTIVATION

In a cochannel system supporting multiple simultaneous transmissions within the same frequency band, the received signals suffer from CCI. The received signal is a weighted superposition of the transmitted signals perturbed by noise. Assuming a rich-scattering environment, a receiver equipped with multiple diversity branches (in the form of appropriately spaced receiving antennas) can separate the coupled signals by exploiting the differences in the spatial channels between each of the transmitters and the receiver antenna array. Multiuser detection (MUD) is the process used to simultaneously detect the multiple transmitted signals [12, 24].

[2]In the case of a frequency-selective fading channel.

An interesting and rather challenging situation occurs when the system is operating under *overload* where there are fewer receive diversity antennas than there are transmitted signals. The optimum detection techniques for this system are joint MAP and joint ML detection [23, 24]. Unfortunately, their implementation in real time becomes infeasible as the number of transmitted signals is increased. The complexity of optimum detection increases exponentially with the number of transmitted signals, the number of bits per symbols and the length of the channel memory.

Several attempts have been made in the past to develop low complexity near optimal detection techniques [2–6, 8–11, 29–31, 33–37]. Sphere detection (SD) [3–5, 8, 30, 34, 36] yields ML or near ML performance at reduced complexity, but its complexity is exponential in the number of excess signals when the system is overloaded[3]. Other suboptimum MUD techniques perform badly in overloaded conditions due to singularity of the channel matrix which contains the channel responses (gains) interconnecting the transmitters and receivers. For example, zero forcing (ZF) detection [1] and the probabilistic data association (PDA) algorithm [29] cannot be applied due to the singularity of the channel matrix. A stable ZF solution can be obtained by carefully defining the linear transformation matrix [10, 38]. The minimum mean-square error (MMSE) detector performance is badly affected in overloaded conditions due to the loss of degrees of freedom [1]. It is important to note that the performance degradation of the linear detection techniques is essentially equivalent to singularity of the channel matrix [39].

In [40], a spatially reduced-search joint detection (SRSJD) technique is proposed for overloaded receivers under AWGN channel assumption. A line-of-sight (LOS) link is also assumed to be present between the transmitters and the receiving station. It offers a reduction in complexity by partitioning the receiver structure into two stages. During the first stage, spatial separation between interfering signals is exploited with a linear preprocessor. In the second stage, iterative joint detection is performed with a reduced-state trellis over space. The problem with this method is that it requires the geometry of the antenna array to be circular with low elevation angle. The complexity of SRSJD depends on several detector parameters; specifically it varies exponentially

[3]Excess signals is defined as the number of transmitted signals minus the number of receiving antennas and is always greater than zero for an overloaded system.

with the size of the reduced search trellis.

To improve performance over linear detection, a number of algorithms based on a group detection principle have been developed [2, 9, 11]. In [2], all the interferers out of the group currently being detected are nulled using a ZF approach. The number of excess signals cannot be greater than the group size minus one. Similarly, another suboptimal group detection scheme with improved performance is proposed in [9], where soft symbol estimates are calculated and iteratively canceled from the received signal. Complexity in this case is proportional to the number of groups, but increases exponentially with group size. Performance improves as group size is increased or alternatively when the number of groups is decreased. Hence, the main reduction in complexity comes from dividing the number of transmitted signals into detection groups. In [11], a list-based iterative MUD scheme for an overloaded antenna array is proposed for frequency-flat Rayleigh fading channels. The receiver structure is again divided into two stages. During the first stage a linear preprocessor is applied for cochannel interference reduction. In the second stage, a multiuser detector operates, based on iterative processing and groupwise detection to extract a list of the most likely user symbols. It is shown that an overloaded system operating with nine transmitted signals and six receive diversity branches offer symbol-error rate (SER) performance close to that of an ML detector at an SNR of 7 dB with a complexity approximately half that of the ML detector. However, the detector exhibits an error floor with increasing signal-to-noise ratio (SNR).

Recently, nature inspired *metaheuristic* optimization techniques such as genetic algorithm (GA), particle swarm optimization (PSO) and ant-colony optimization (ACO) algorithms have been developed for solving non-deterministic polynomial-time hard (NP-hard) optimization problems. Since the search space of the optimal MAP/ML detector is discrete with a solution having integer components, it is therefore an integer least-squares optimization problem [41]. Furthermore, it belongs to the class of NP-hard, combinatorial optimization (CO) problems [24, 42]. The heuristic approaches, the metaheuristic techniques (i.e. GA, ACO, PSO) can also be used for solving the MUD problem. These algorithms have already been successfully applied to MUD in

code-division multiple access (CDMA) systems [43–52].

In the context of overloaded array processing, GA assisted techniques are developed in [6] and [10], for low complexity suboptimal signal detection. In [6], a technique based on a combination of adaptive filtering and GA is presented. It is shown that GA can achieve good performance in an overloaded system by exploiting the *a priori* information generated by a specially designed spatial filter. In [10], a soft-biased technique based on GA is proposed. This scheme performs satisfactorily provided channel estimation is sufficiently accurate. Our work in this area is based on the fact that, for a direct-sequence CMDA (DS-CDMA) system, the performance of ACO based MUD converges to that of optimal MUD much faster (with significant complexity savings) in comparison to its GA based counterpart [51]. In this regard, ACO has already been used for near ML detection for multiple-input multiple-output (MIMO) systems [31]. This last approach relies on the QR factorization of the channel matrix and cannot easily handle the singularity problem seen in the overloaded scenario.

1.3 OVERVIEW AND CONTRIBUTIONS

The aim of this thesis is to design reduced complexity near-optimum performance detection techniques for multiple signal cochannel communication systems. We develop receiver architectures to work in underloaded, critically loaded and overloaded scenarios. Ant colony optimization as an outer framework is integrated with simplified ML/MAP criteria. The resulting algorithms achieve ML or near-ML performance while offering significant savings in terms of required computational efforts.

1.3.1 Thesis Overview

The thesis is divided into seven chapters addressing different aspects of this work. In chapter 2, the impact of multipath fading is examined followed by a summary of diversity approaches to combat its effects. Multiple access methods are introduced for the purpose of better utilization of the radio spectrum. The chapter ends with a discussion on spatial multiplexing systems and presents a system model for multiple

signal cochannel communication systems.

Chapter 3 summarizes existing approaches for signal detection. The different detection methods fall in one of two main categories: linear and nonlinear detection techniques. Key developments are discussed for each category along with their advantages and shortcomings.

In Chapter 4, we present ACO assisted reduced-complexity detection techniques for binary and non-binary signals. The developed algorithms are discussed along with the associated heuristic designs. An overview of the associated computational costs along with a performance-complexity tradeoff is provided at the end.

A near-ML approach for the detection of non-binary signals based on ACO is presented in Chapter 5. The heuristic framework and its integration with the ACO algorithm is discussed. A complexity bound is developed and performance comparisons are provided.

Finally, Chapter 6 presents thesis conclusion and discusses the proposed future research.

1.3.2 Thesis Contributions

This thesis develops reduced complexity near-optimum performance detection algorithms based on an ACO metaheuristic for cochannel systems. It is assumed that the receiver has multiple receiving antennas to collect multiple copies of the Rayleigh faded superimposed transmitted signals. We consider an overloaded receiver antenna array where the linear suboptimum approaches fail to offer acceptable performance while other approaches either incur high computational costs or exhibit an irreducible error-floor at medium to high SNR values.

We have developed an ACO based receiver which treats the detection problem as finding a shortest path over a weighted multigraph representing all possible solutions. The artificial ants *walk* on the graph and probabilistically build candidate solution vectors. The search process is guided by pheromone trails and the heuristic values. An overview of our contributions is presented below:

- We first developed an ACO based soft iterative detection approach which relies

on a simplified MAP criteria to extract the most probable signals from the entire search space. The scheme can handle only binary phase-shift keying (BPSK) modulated signals. Due to iterative soft signal processing, channel coding schemes can be easily incorporated into the developed receiver. Performance results show that our approach can achieve near ML performance at significantly reduced complexity for critically loaded systems and continues to provide good performance under overloaded scenarios.

- Second, an integrated framework is developed which comprises the ACO metaheuristic and a recursively defined ML search criteria to tackle multilevel modulations. The receiver is capable of handling overloaded environments and also achieves near-ML performance with significantly reduced complexity. The proposed receiver consists of a preprocessing stage and a solution construction phase. The preprocessing stage obtains an MMSE estimate and performs a Cholesky factorization of the channel matrix. The solution construction phase then employs the ACO algorithm and partial ML metric (PMM) to extract a list of most likely candidate solution vectors. Due to the branch and bound nature of the algorithm, an exact expression for the complexity cannot be determined. Instead, an upper bound on computational complexity is developed. The performance results suggest that the proposed receiver achieves ML performance with considerable computational savings.

The following publications have been produced during the course of this research

1. *K. T. Tasneem*, *P. A. Martin*, *D. P. Taylor*,"Iterative soft detection of cochannel signals using ant colony optimization," *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp.1617-1621, 9-12 *Sept.* 2012.

2. *K. T. Tasneem*, *P. A. Martin*, *D. P. Taylor*,"Detection of Multiple Cochannel Signals using Ant Colony Optimization in Overloaded Environments", submitted to *IEEE Transactions on Vehicular Technology.*

Chapter 2

COMMUNICATION OVER WIRELESS FADING CHANNELS

2.1 INTRODUCTION

Over the past few decades, wireless communication has emerged as a rapidly growing field of communication offering many design and optimization challenges to communication engineers. Wireless communication offers connectivity between different devices free of wires thus avoiding the work and cost of laying cables. This also increases the reachability to places where wireline telephony is either not cost-effective or in some cases impossible. Secondly, the significant growth and improvements in very-large-scale integration (VLSI) technology have resulted in miniature and low-powered electronic devices capable of performing sophisticated signal processing tasks in real-time.

On a commercial note, the popularity and acceptance of cellular telephony is evident as there are approximately six billion users worldwide [53]. In addition, wired local-loops from telephone exchange to end-users' premises and the wired local area networks installed in many homes and businesses are being replaced by their wireless counterparts, wireless local-loop and wireless local area networks, respectively. Other applications of wireless communication include wireless sensor networks, remote controlled homes and wireless personal area networks. More recently, the development of new wireless devices, e.g. smart phones and tablets, and improvements in laptop technology have resulted in mounting pressure on existing wireless communication systems and standards to support higher data rates, greater system capacity, efficient energy consumption and better quality of service.

From the point of view of researchers and engineers in the field, all the associated phys-

ical processes and requirements translate into mathematical models and constraints that need to be optimized for the intended solutions. Some of those have already been resolved while many of them are under consideration by researchers. Furthermore, many of those resolved are still under constant optimization and tuning for further improvements in terms of various figures of merits mentioned above. Practically, wireless transmission suffers from many different types of impairments that are inherent in the wireless environment.

This chapter starts with a brief overview of propagation phenomena in Section 2.2 followed by a summary of different statistical models used for multipath fading channels in Section 2.3. In Section 2.4, the role of diversity and the different types of diversity combining techniques are discussed. A preview of some of the major multiple access techniques is provided in Section 2.5. A discussion of spatial multiplexing systems is presented in Section 2.6 and the system model of a multiple-antenna cochannel communication system is introduced. At the end, a short summary of the chapter is provided in Section 2.7.

2.2 THE WIRELESS CHANNEL

In wireless communication, free-space is the medium for carrying information from source to destination. Before information is transmitted over the air, several transformations and signal processing operations are performed resulting in its digital representation [54]. The digital data in the form of binary pulses is filtered or shaped and then modulated to produce a signal which is finally transmitted as an electromagnetic wave. The terminal equipment is fabricated with antennas for the purpose of radiating the signal into the atmosphere. Once the signal is transmitted, it propagates through the medium in the form of a radio wave towards the intended receiver. Here it is important to note two fundamental impairments affecting the signal: *fading* and *interference*.

Fading happens at both *small-scale levels* and *large-scale levels*. Small-scale fading, simply referred to as *fading*, is the result of constructive and destructive interference of multiple copies of the same signal arriving at the receiver along different paths. It

is important to note that the fading process is sensitive to receiver movements, and therefore, a slight change in the receiver position[1] may change the received signal amplitude [55].

On the other hand, variation in the average received signal strength due to increasing distance and shadowing by large objects (e.g. high-rise buildings, mountains, etc.) placed between the transmitter and the receiver is called large-scale fading [56]. It is the result of movement over distances large enough to cause gross variations in the signal strength and is usually on the order of the cell size [55]. Large-scale fading phenomena typically do not depend on frequency.

Interference can happen between the signals transmitted by a single transmitter due to the delay dispersion inherent in the wireless channel or between the signals transmitted by different sources to a common destination [57]. In the former case, it is termed inter-symbol interference (ISI), whereas in the latter case the interference is called cochannel interference (CCI).

Fig. 2.1 depicts a baseband model of a typical SISO wireless communication system. The source emits information which is first converted to digital format suitable for processing by digital signal processors. The modulator then produces the information signal, $x(t)$, which is radiated into free-space by the source antenna. The transmitted signal propagates through the wireless channel represented by its time-varying impulse response, $h(\tau, t)$, and finally impinges on the receiver antenna where noise, $n(t)$, is added to form the received signal, $y(t)$. The channel acts as a linear filter transforming the input signal $x(t)$ into the output signal $y(t)$ according to the relation [12]

$$y(t) = h(\tau, t) \star x(t) + n(t) \tag{2.1}$$

where $(\star)$ denotes convolution. It is the task of communication engineers to design efficient and robust signal processing/coding techniques which are able to offer the desired quality of service in the presence of multipath fading and interference.

[1] Usually of the order of carrier wavelength.

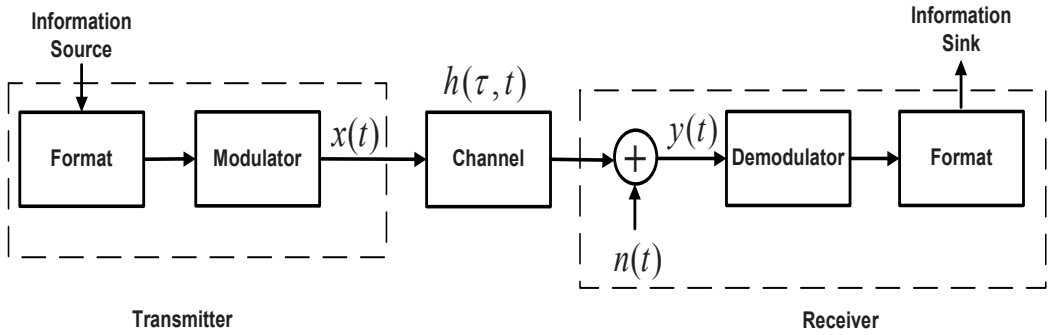

Figure 2.1 Block diagram of a baseband wireless communication system.

2.2.1 Multipath Fading

Radio waves representing the information signal originate from the transmitting antenna and travel through free space to reach the destination via a number of different spatial paths. During propagation, they are affected by the terrain and different physical objects in their path including buildings, hills and trees. As a result, the waves undergo reflection, refraction, diffraction and scattering. These phenomena combined with the nature of the atmosphere and the weather conditions define the channel behavior and the characteristics of the received signal [55].

The channel may be characterized by its impulse response, $h(\tau, t)$, which is the channel output at time t to an extremely short-duration pulse or an impulse at time $t-\tau$. It can be observed that the received signal is composed of a train of impulses when the input signal, $x(t)$, is an impulse as depicted in Fig 2.2. The time elapsed between the impulses is specified by the variable τ. An individual impulse in the impulse-train actually represent a *multipath component*. Associated with i-th such multipath component is its attenuation factor, $a_i(t)$, propagation delay $\tau_i(t)$ and phase shift $\phi_i(t)$. The total number of such multipath components is often very large and collectively termed as *multipath*. Furthermore, the attenuation factor and propagation delay of each impulse (or equivalently the multipath component) inside an impulse-train also changes if the same channel probing experiment is performed at different times thus depicting the time-varying nature of the channel. Therefore, the channel impulse response for discrete

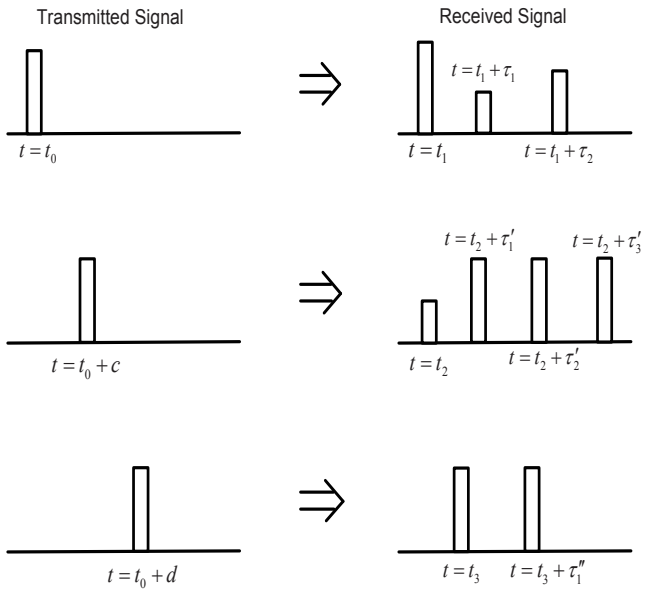

Figure 2.2 Multipath phenomena and impulse response of a time-varying wireless channel [12].

multipath components can be written as [12]

$$h(\tau, t) = \sum_i a_i(t)\, e^{j\phi_i(t)} \delta(t - \tau_i(t)), \tag{2.2}$$

where $\delta(\cdot)$ represents the unit impulse function. Now, suppose the input to the channel is an unmodulated carrier at frequency f_c, i.e. $x(t) = 1$ at baseband. In the absence of noise, the baseband received signal can be expressed using (2.1) and (2.2) as [12]

$$y(t) = \sum_i a_i(t)\, e^{j\phi_i(t)}, \tag{2.3}$$

where $j = \sqrt{-1}$ and $\phi_i(t) = -2\pi f_c \tau_n(t)$. Clearly, a conventional receiver is unable to distinguish between the different multipath components and therefore it superimposes them. Sometimes the receiver adds these multipath components constructively (or in-phase) resulting in a better overall received signal strength, whereas at other times the multipath components are added destructively (or out-of-phase) resulting in a rapid fall in the received signal strength which is termed a *fade*. This fluctuation in the received signal strength is called *multipath fading*. Note that the effect of fading strongly depends on the spatial position of the receiver i.e. a slight change in receiver position in space

can result in a change in $\phi_i(t)$ and $a_i(t)$, thus changing the overall received signal quality. Usually, the number of multipath components is large and the central limit theorem can be invoked, [58], to model the channel impulse response as a complex-valued Gaussian random process in the variable t.

Small-scale fading can be further classified as *frequency-flat/flat* or *frequency-selective* (due to channel time-dispersion) and *slow* or *fast* (depending on the frequency-dispersive nature of the channel) [12, 56]. A received signal is said to have experienced flat fading if the channel gain, $a_i(t)$, remains constant whereas the phase response, $\phi_i(t)$, varies linearly over a bandwidth greater than the bandwidth of the transmitted signal. In such conditions, the received signal amplitude fluctuates due to the variations in the channel gain resulting from the superimposed multipath components. However, the spectral characteristics of the transmitted signal remain intact at the receiver and all the multipath components are confined within one signaling interval, T_s. On the other hand, if the radio channel has a constant gain, $a_i(t)$, and linear phase response, $\phi_i(t)$, over a bandwidth smaller than that of the transmitted signal, the transmitted signal is said to have undergone frequency-selective fading. As a result, the various spectral components of the signal are distorted differently. Viewed in the time-domain, the received signal in a given signaling interval is composed of multipath components of the signal transmitted in the current interval (which is considered to be the desired signal) as well as signals transmitted in previous intervals. This phenomenon is called intersymbol interference (ISI) and can cause severe degradation in the received signal quality. It can be said that the channel possesses memory and acts like a finite-state machine if the memory is finite.

Time-selective fading occurs when there is a relative motion between the transmitter and the receiver or due to the presence of moving scatterers in the surroundings. Then *Doppler* spread appears in the received signal spectrum, causing frequency dispersion. If the Doppler spread is significant relative to the bandwidth of the transmitted signal, the received signal is said to undergo fast fading. On the other hand, if the Doppler spread of the channel is insignificant compared to the bandwidth of the baseband signal, the signal is said to undergo slow fading.

2.2.2 Interference

In wireless communication, the presence of signal interference limits the system coverage, capacity and performance. Most modern wireless communication systems exist in challenging environments. These include wireless networks ranging from commercial cellular systems to specialized mobile radio and paging/broadcast systems. Also many wireless applications coexist in the unlicensed ISM band. These systems may result in mutual interference.

ISI and CCI are the two fundamental type of interference in wireless communication systems. The notion of ISI is already described in detail in the last section. In wireless cochannel systems, multiple signals (belonging to single or many different users) transmitted simultaneously within the same frequency band also interfere with each other causing CCI [57]. CCI can be among the signals transmitted by different sources communicating with a common destination (e.g. uplink of a cellular system where multiple users communicate with a common base-station) or among the signals transmitted by a single source to multiple destinations (e.g. downlink of a cellular system where base-station communicates with multiple users). Another type of interference called adjacent channel interference (ACI) is also present in wireless channels and occurs between signals which are on neighboring frequencies [56]. ACI occurs due to poorly designed receiver filters which allow the adjacent frequencies to penetrate into the passband of the desired signal or due to relative Doppler shifts.

2.3 STATISTICAL MODELING OF MULTIPATH CHANNELS

The wireless channel acts as a filter to the input signals emitted by the source and transform them into faded, noisy received signals. The receiver has to extract correct information from the received signal. Clearly, this is not possible without an accurate understanding of the nature of the channel impulse or frequency response. Theoretically, a multipath channel is modeled as a random process and several probability distributions are considered for modeling its stochastic nature depending on the assumptions possible in the real world. This helps in developing appropriate robust and

efficient signal processing techniques. In the presence of a large number of scatterers in the environment, the number of multipath components arriving at the receiver is large. In such a situation, the central limit theorem is applicable and the resulting model for the channel impulse response is a complex Gaussian random process.

2.3.1 Rayleigh Fading

In the presence of a large number of buildings, trees, houses, mountains etc. the existence of a direct signal component between the transmitter and receiver is not possible. The signal observed at the receiver is the superposition of various reflected, refracted and diffracted multipath components resulting in a zero mean complex Gaussian fading process. At any particular instant, the amplitude of this fading process is Rayleigh distributed while the phase follows a uniform distribution in the interval $[0, 2\pi]$ [12].

The probability density function (PDF) of the Rayleigh distribution can be expressed as follows [12, 54]

$$p(r) = \begin{cases} \frac{2r}{\sigma^2} \exp\left(-\frac{r^2}{\sigma^2}\right) & \text{if } r \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

The symbol σ^2 represents average power of the fading process.

2.3.2 Rician Fading

In some cases, particularly in fixed wireless systems, there is a direct dominant signal component, termed the LOS component, between the transmitter and receiver in addition to many reflected multipath components. The presence of this LOS component results in the fading process being complex Gaussian with a non-zero mean. The appropriate model used to describe the statistical fluctuations of signals received from all the paths in this case is known as the Rician distribution [59].

The Rician distribution can be expressed by its PDF as [12, 56]

$$p(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left(-\frac{r^2+A^2}{2\sigma^2}\right) I_0\left(\frac{rA}{\sigma^2}\right) & \text{if } r \geq 0, A \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

Here, A is the peak amplitude of the LOS component and σ^2 is again the noise power. $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind defined as

$$I_0(r) = \int_0^{2\pi} \exp(-r\cos\phi)\, d\phi. \tag{2.6}$$

The Rician distribution is usually defined in terms of a parameter called the K-factor, which is the ratio of the power in the deterministic component of the signal to the variance of the multipath component. The Rician factor K is given as **??**

$$K = \frac{A^2}{2\sigma^2}. \tag{2.7}$$

It is important to note that as the LOS component vanishes ($A \to 0$ or equivalently $(K)_{dB} \to -\infty$) the fading process only possess reflected multipath components and the Rician PDF converges to the Rayleigh PDF.

2.4 INTRODUCING DIVERSITY IN MULTIPATH FADING CHANNELS

It is evident, that the multipath fading in traditional SISO wireless links adversely affects the propagating signals. In order to mitigate its effects, diversity techniques can be used to improve the reliability of the transmission without increasing the transmitted power or sacrificing bandwidth. The concept of diversity revolves around the idea that if the receiver can be supplied with more than one independently faded version of the same signal then the probability that all those signals will be in a deep fade at the same time can be significantly reduced. This implies that some copies of the same signals may be severely faded (experiencing deep fade) while others can be less attenuated (from other independent paths offering less severe fading). Apart from diversity, multipath fading phenomena also plays a positive role in increasing the system capacity as in

MIMO wireless communication systems by employing multiple transmit and receive antennas at both the transmitting and receiving sides.

For the purpose of better understanding, it is convenient to express the continuous-time signal model of (2.1) in its discrete-time equivalent counterpart. The signal in (2.1) can be symbol-spaced sampled to obtain the equivalent discrete-time representation as [18]

$$y_k = \sum_{i=1}^{N} \sqrt{E_s} x_i h_{k-i} + n_k, \qquad \text{for } k = 1, 2 \text{ and } i = 0, 1, 2 \ldots N-1 \tag{2.8}$$

where N represents the channel response length and E_s is the average transmitted energy.

The transmitted information signals $\{x_i\}$ are chosen from a linear signal modulation (e.g. pulse-amplitude modulation (PAM), phase-shift keying (PSK) or quadrature amplitude modulation (QAM)). For the purpose of normalization, we assume $T_s = 1$ second which also makes E_s the transmitted power during the signaling period. The channel gain h_k is assumed to be complex Gaussian with zero mean and unit variance i.e. the channel energy is normalized to $\sigma_h^2 = 1$. This assumption is used for both flat-fading channels where $N = 1$, and frequency-selective channels where $N > 1$ and the channel impulse response is composed of multiple taps (i.e. the total energy in all the taps is normalized, $\sum_{k=1}^{N} E\left(|h_k|^2\right) = 1$). Finally, the noise is assumed to be zero mean complex Gaussian with variance σ_n^2.

To explain the concept of diversity, a frequency-flat fading scenario is considered (i.e. $h_k = 0$ for $k > 1$). In such a case, the input-output relation in (2.8) can be further simplified to

$$y_k = \sqrt{E_s}\, h x_k + n_k. \tag{2.9}$$

In this discrete-time baseband representation, the channel gain can be interpreted as a phasor, $h = |h|\, e^{j\phi_h}$, which is multiplied with the transmitted signal, $x_k = |x_k|\, e^{j\phi_{x_k}}$, resulting in its scaling and rotation thus transforming it into the received signal, $y_k = |y_k|\, e^{j\phi_{y_k}}$. Here, ϕ_{x_k}, ϕ_{y_k} and ϕ_h denote the phases associated with x_k, y_k and h, respectively. When this scaling and rotation is severe in nature (in other words when the channel is in a deep fade, $|h| \to 0$), it is an extremely difficult task for the receiver

to recover the transmitted information. To overcome this, a receiver is provided with L independently faded replicas of the signal, $y_1, y_2 \ldots y_L$. The receiver attempts to process these copies so as to recover the transmitted signal reliably.

Now, if p is the probability that one of the copies will fade below some threshold value then the probability that all L copies will fade simultaneously is proportional to p^L. Diversity techniques are commonly classified as spatial diversity, polarization diversity, time diversity or frequency diversity. These are the core diversity techniques. Some practical systems use a combination of two or more of these methods to provide multi-dimensional diversity in order to meet the specified system performance requirements. For example, multiple receive antennas at the base stations are used in conjunction with interleaving and error control coding to simultaneously exploit space and time diversity in GSM [59].

2.4.1 Basic Techniques

Frequency Diversity

One way to obtain L independently faded samples of the same signal is to send the signal on L sufficiently separated carrier frequencies [12]. Before specifying adequate frequency spacing we define the coherence bandwidth of a channel as follows [56].

Definition 2.1: Coherence bandwidth is a statistical measure of the range of frequencies over which the channel passes all the spectral components with almost equal gain and linear phase.

Notice that the spectral components of a transmitted signal will remain intact if it occupies a bandwidth less than the coherence bandwidth of the channel. Viewing it from another angle, if two signals are transmitted with a frequency-domain separation greater than the coherence bandwidth, they will fade independently. Thus, if the data is sent on carriers that are spaced by more than the coherence bandwidth of the channel, frequency diversity can be achieved.

Time Diversity

Another way to achieve the multiple independently faded versions of the same signals is time diversity. If the transmitted signal is repeated in time such that the time duration between successive repetitions is large enough for the channel characteristics to change from one state to another, the received signals can undergo independent fading [12]. To specify a time interval sufficient for this purpose, we define coherence time of the channel [56]

Definition 2.2: Coherence time is the statistical measure of time duration over which the channel impulse response is invariant.

In this case, if two signals are transmitted with a time-domain separation of more than the coherence time, they will experience independent fading. Thus it can be said that, if data is repeated in time with time duration greater than the coherence time of the channel, then time diversity can be guaranteed [56]. If the data rate is fixed, higher Doppler will provide greater time diversity though not necessarily better reception quality as higher Doppler means greater ISI.

Antenna Diversity

System performance in multipath fading environments can be improved using spatial diversity which can be achieved by placing L sufficiently spaced antennas in space (typically half a wavelength or more between the antennas is sufficient [55]). This also requires appropriate signal processing that combines the received signals from the various antennas in an effective manner. Spatial or space diversity is also called antenna diversity. It can be achieved by using multiple antennas at either the transmitter or the receiver. If multiple antennas are used at the receiver to collect independently faded copies of the transmitted signal then it is called receive diversity. On the other hand, if multiple antennas are employed at the transmitter then it is called transmit diversity [17]. Spatial diversity does not introduce any loss in bandwidth efficiency which is a very attractive feature for designing systems with high data rate.

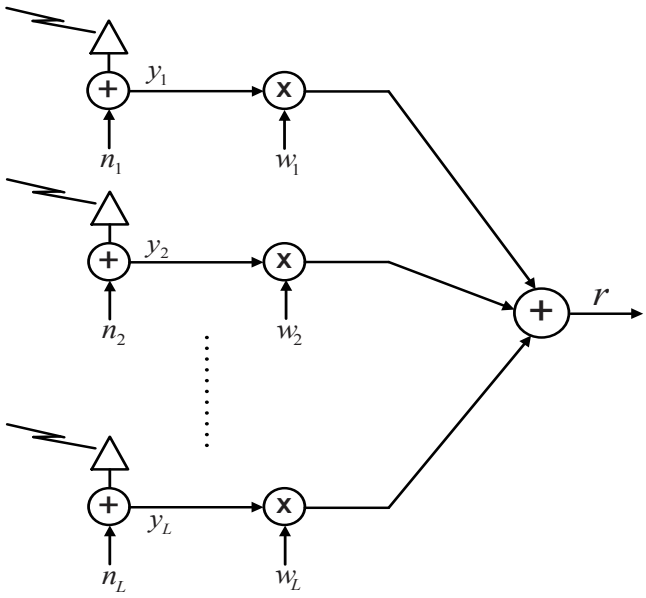

Figure 2.3 Combiner structure in a diversity aided receiver.

2.4.2 Diversity Combining

The main objective of all diversity techniques is to increase SNR at the receiver thereby improving the error-rate performance of the system. This enhancement actually depends upon how the different replicas of the signal are combined at the receiver. The aim of diversity combining is to extract maximum gain out of those faded copies by maximizing the received SNR. In this section, the diversity combining techniques are described with reference to spatial diversity. These combining techniques are equally applicable to the other diversity methods discussed above.

For the interference-free environment, the most common space diversity combining techniques employed at the receiver for combining the received signal replicas are selection combining, switched combining, maximal ratio combining, and equal gain combining [59–61]. Before discussing each of these techniques individually, we assume that the receiver is equipped with L sufficiently spaced antennas to collect L independently faded copies of the same signal. The received signal at the l-th antenna can be expressed as

$$y_l = \sqrt{E_s}\, h_l x + n_l, \tag{2.10}$$

where $l = 1, 2 \ldots L$. Each of the channel gains are an independent and identically distributed (iid) Gaussian random variable with mean zero, $E\left(|h_l|^2\right) = 1$ and $E\left(h_l h_k^*\right) = 0$. Also, the instantaneous received SNR at the l-th antenna is defined as

$$\gamma_l^{inst} = |h_l|^2 \frac{E_s}{N_0}, \tag{2.11}$$

where N_0 is the noise power spectral density which is the same as the noise variance as we are assuming the transmission bandwidth to be 1 Hz for normalization purposes. Fig. 2.3 depicts the general structure of the diversity receiver where the weights are determined according to the chosen diversity combining method.

Selection Combining

In selection combining, a receiver equipped with L antennas selects the signal that offers the best decision statistic, y_m at the m-th receive antenna. The decision statistic used in selection combining can be based on either the maximum SNR or the lowest bit-error-rate (BER). In the maximum SNR case, diversity branch m corresponding to the maximum instantaneous SNR, γ_m^{inst}, is selected for further processing. Therefore, the weights are chosen according to [61, 62]

$$w_m = \begin{cases} 1 & m \in \gamma_m^{inst} \text{ where, } \gamma_m^{inst} = \max_l \left\{\gamma_l^{inst}\right\} \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

When the lowest BER criterion is used, the diversity branch offering the lowest BER is employed for signal detection. The source transmits a training sequence to the receiver. The receiver demodulates each of the received signals, calculates the associated BER and selects the antenna that offers the smallest BER.

Switched Combining

One of the drawbacks of selection combining is that the receiver must continuously monitor all the antennas in order to select a particular antenna. As an alternative solution, switched combining can be employed. In switched combining, the receiver

scans all L antennas and selects the first signal it finds offering the SNR γ_m^{inst} greater than a threshold value γ_T^{inst} [62]. The combiner chooses the weights w_m according to [61, 62]

$$w_m = \begin{cases} 1 & m \in \gamma_m^{inst} \text{ where, } \gamma_m^{inst} = \left(\gamma_l^{inst} > \gamma_T^{inst}\right) \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

When γ_m^{inst} falls below the threshold γ_T^{inst}, the receiver starts the scanning process again and switches to a branch offering a selection value greater than the threshold value. A distinctive feature of both selection and switched combining is that they do not require the knowledge of the channel state information (CSI) and can therefore be used in combination with both coherent and non-coherent modulation schemes.

Maximal Ratio Combining

In both selection and switched combining, the receiver exploits only one of the L available independent received copies of the same signal. In other words, the energy of the other $L-1$ available signal replicas is simply wasted. This drawback is overcome by the maximum ratio combining (MRC) method. Assuming the availability of channel-state information (CSI), the MRC combiner first calculates L complex weights $w_m = |w_m|\, e^{(j\theta_m)}$ from the estimated channel gains as follows [61, 62]

$$w_m = h_m^*. \tag{2.14}$$

An attempt is then made to maximize the output SNR by multiplying each of the signal replicas with the corresponding weight and adding them all to produce the final decision statistic. The output signal in MRC is thus a weighted linear combination of the signal replicas $\{y_i\}_{l=1}^{L}$ expressed as

$$y = \sum_{l=1}^{L} w_l y_l. \tag{2.15}$$

Since MRC can only operate based on the availability of CSI, it can work only with coherent signaling schemes.

Equal Gain Combining

Equal gain combining (EGC) also employs the superposition principle for signal combining with the only difference being that the magnitude of the complex weights are set equal to 1 i.e. $|w_m| = 1$ which implies w_m can be expressed as

$$w_m = e^{j\phi_m}. \tag{2.16}$$

In other words, the combining is carried out to align the phases without worrying about the effects of the fading amplitudes, which clearly is a suboptimum approach. The benefit of EGC over MRC is complexity savings as it does not require the estimation of the fading amplitudes.

2.5 MULTIPLE ACCESS TECHNIQUES

A wireless communication system operates within a certain frequency band. Spectrum is a limited and expensive resource that needs to be used wisely. Most practical systems allow multiple users to share a common channel to exploit the available bandwidth more efficiently. Some examples of wireless communication systems using this idea, where multiple transmitters share a common communication channel, include: cellular systems where multiple users simultaneously transmit their signals to a common base-station, satellite communication systems where several ground stations send information to a satellite, wireless local area networks where multiple devices are connected to a common access point etc. Each of the message sources in such a communication system is called a user and the whole system is termed a multiuser communication system. A generic model of a multiuser communication system is shown in Fig.2.4.

For such systems, access to the communication channel by system users must be controlled in a manner that avoids the detrimental effects of multiple access interference. Several different multiple access techniques are possible in order to accomplish this task.

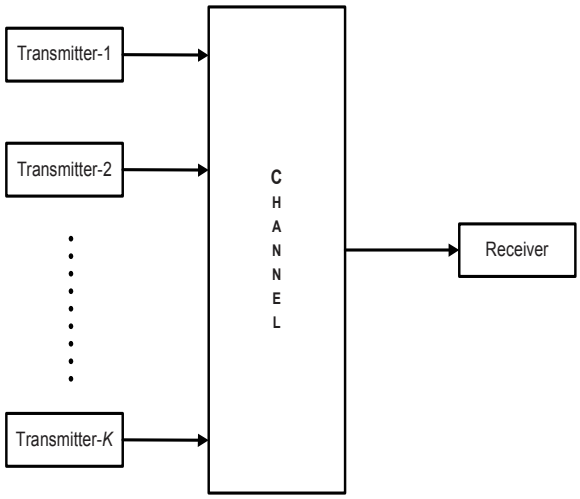

Figure 2.4 General model of a multiple access communication system.

2.5.1 Frequency Division Multiple Access

In frequency division multiple access (FDMA) [56], the available transmission bandwidth B is partitioned into K non-overlapping frequency subbands or subchannels each having a bandwidth B/K. Upon request, one of the available subchannels can be allocated to a particular user. Fig. 2.5 depicts the channelization for a typical FDMA system.

The first generation Advanced Mobile Phone System (AMPS) is an example of an analog cellular communication system. In AMPS, the uplink operates in the frequency band $824-849$ MHz while the downlink is supported in the $869-894$ MHz band. Each requesting user is assigned one radio channel consisting of a pair of simplex channels with 45 MHz separation [56, 62].

2.5.2 Time Division Multiple Access

Another simple strategy to control the channel access is to divide the frame interval of T_F seconds into K non-overlapping sub-intervals each of duration T_F/K seconds [56]. As an access request arrives from a user, a subinterval is assigned for transmission. This strategy is called Time Division Multiple Access (TDMA). In TDMA, each of the subintervals requires synchronization and channel estimation, which increases the

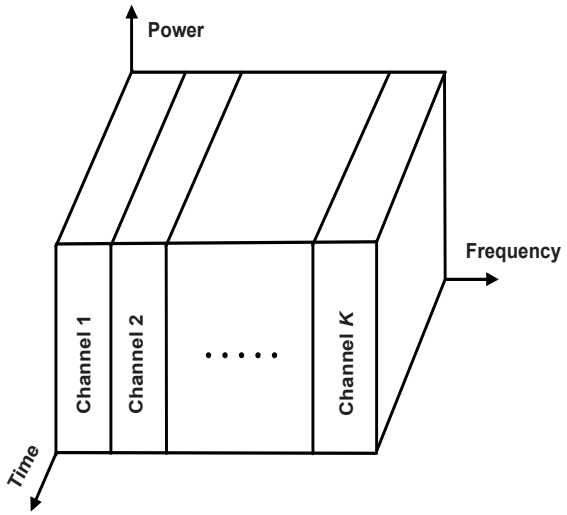

Figure 2.5 Channelization scheme for FDMA systems.

signal processing requirements. The duration of each subinterval needs to be calculated appropriately as an insufficient duration may require synchronization and channel estimation tasks more frequently in which case the overall information interchange will be less than the overhead. Similarly, if the subinterval length is large then the channel may change within the subinterval and the equalizer will have to track the channel variations [56, 62]. Fig. 2.6 depicts the channelization scheme used in a typical TDMA system.

The second generation Global System for Mobiles (GSM) standard is an example of commercial digital communication system based on TDMA principles. The uplink frequency range for GSM is 890-915 MHz whereas the downlink frequency range is 935-960 MHz. Each of the transmission channels consists of an uplink channel and a downlink channel separated by 45 MHz and time-shared by eight users [63].

2.5.3 Code Division Multiple Access

As an alternative to both FDMA and TDMA, CDMA [24] allows multiple transmissions to be possible within the same time and frequency. Each of the requesting users is assigned a unique user-specific code or signature sequence. These codes are designed in such a way that they possess zero or very low cross correlations. The user spreads

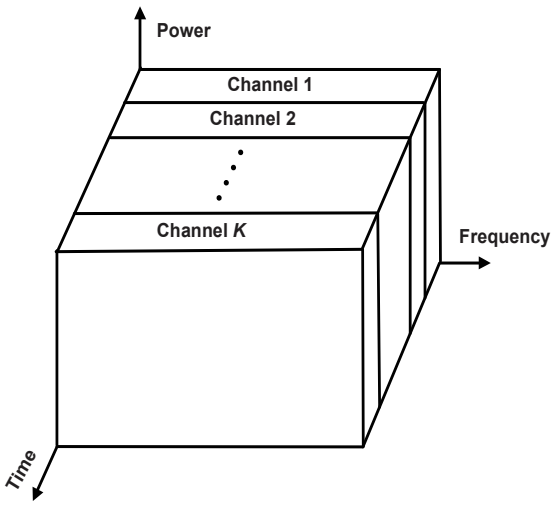

Figure 2.6 Channelization scheme used in TDMA systems.

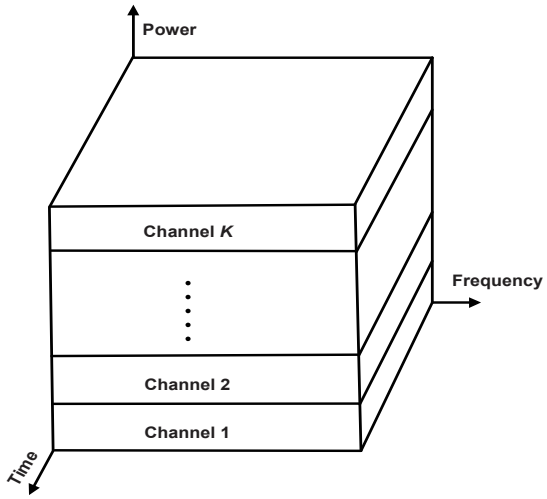

Figure 2.7 Channelization for CDMA systems.

his data using the signature sequence to a much larger bandwidth. At the receiver, the signals arriving from different users are separated by correlating the received signal with each of the user-specific signature sequence [12]. If there are some users whose signature sequences are not fully orthogonal then they may cause interference into the desired user's decision statistics. This interference can become severe if the power level of those users is larger than the desired user. This situation is called the near-far effect and it requires a robust power control mechanism [12, 24, 62]. Fig. 2.7 depicts the structure of a typical CDMA system.

An example of a practical CDMA system is the Interim Standard 95 standard developed by Telecommunications Industry Association (TIA). The standard recommended the uplink operation in the frequency band $824 - 849$ MHz and the downlink within the frequency band $869 - 894$ MHz. A CDMA channel is composed of a pair of simplex channels separated by 45 MHz with many users sharing the channel simultaneously [64]. Others examples include wideband CDMA (WCDMA) and similar systems.

2.6 SPATIAL MULTIPLEXING SYSTEM

In Section 2.4.2, the use of multiple antennas at the receiver is described for the purpose of achieving spatial diversity to effectively combat multipath fading. The attractive feature of this technique is that it does not compromise the transmitted data rate and bandwidth. Another interesting situation occurs when multiple transmit antennas simultaneously send independent data streams to a receiver equipped with multiple antennas at the same carrier frequency. Such a configuration of multiple transmit and receive antennas is often called a multiple-input multiple output (MIMO), spatial multiplexing (SM) or a cochannel communication system. Originally proposed in [1, 22], MIMO/SM systems captured great attraction because of promising theoretical results later discovered in [16, 19–21], it was shown that the information-theoretic capacity of a SM system increases linearly with the number of transmitted signals. From the detection point of view, it is shown in [23] that the independent spatial channels among different users can be exploited to reliably separate the transmitted signals.

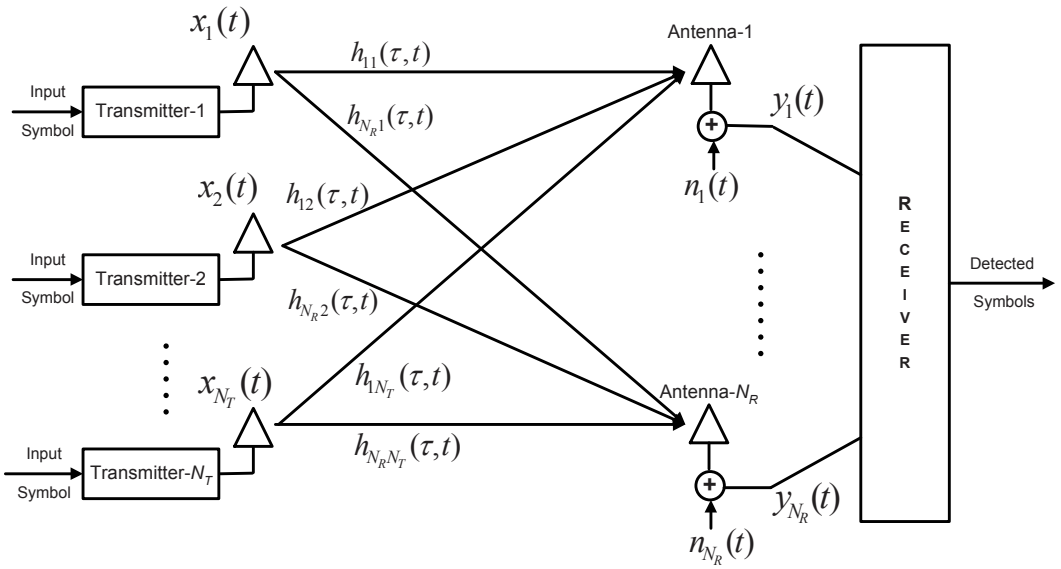

Figure 2.8 Block diagram of a generic time-varying fading cochannel communication system.

Multiple data streams can either originate from a single source or they may represent data transmitted by several users[2] to a common base-station. In the first case, the system is called point-to-point (PP) MIMO whereas in the later case it is referred to as a multiuser (MU) MIMO system. Unlike PP MIMO systems, joint encoding of information is not possible in a MU MIMO system since the users are geographically dispersed throughout the coverage area. Another characteristic of the two systems is the difference in their power constraints and control mechanisms. In the first case, since the transmitting antennas are co-located, the total transmitted power can be kept constant. Adding an antenna at the source causes the power transmitted by each individual antenna to be equally reduced. In the case of a MU MIMO system, a power control mechanism for multiuser systems is implemented to compensate for differences in the path loss and to maintain a peak power constraint. Fig. 2.8 shows a typical block diagram of a cochannel communication system. The multiple-antenna configuration becomes PP MIMO if all the transmitters are co-located or it may represent a MU MIMO system if the transmitters are assumed to be individual users.

[2]Each of the user may be assumed to be equipped with one or more antennas.

2.6.1 System Model

We consider a multiple-antenna system with N_T transmitters and a base-station equipped with N_R receive antennas as shown in Fig. 2.8. The time-varying channel impulse response between the j-th transmitter and the i-th receive antenna is denoted $h_{ij}(\tau, t)$, where τ is the multipath delay and t represents the channel variation with time. The overall channel between all the transmitters and receiver antenna array is characterized by a $N_R \times N_T$ channel matrix $\mathbf{H}(\tau, t)$ as

$$\mathbf{H}(\tau, t) = \begin{bmatrix} h_{11}(\tau, t) & h_{12}(\tau, t) & \dots & h_{1N_T}(\tau, t) \\ h_{21}(\tau, t) & h_{22}(\tau, t) & \dots & h_{2N_T}(\tau, t) \\ \vdots & \vdots & & \vdots \\ h_{N_R1}(\tau, t) & h_{N_R2}(\tau, t) & \dots & h_{N_RN_T}(\tau, t) \end{bmatrix}. \tag{2.17}$$

Assuming $x_j(t)$ to be the signal transmitted by the j-th transmitter, the received signal at the i-th receive antenna can be expressed as

$$\begin{aligned} y_i(t) &= \sum_{j=1}^{N_T} \int_{-\infty}^{\infty} h_{ij}(\tau, t)\, x_j(t-\tau)\, d\tau + n_i(t), \qquad i = 1, 2 \dots N_R \\ &= \sum_{j=1}^{N_T} h_{ij}(\tau, t) \star x_j(\tau) + n_i(t), \end{aligned} \tag{2.18}$$

where $\star$ denotes convolution and $\{n_i(t)\}$ are samples of i.i.d. complex Gaussian random process with mean zero and power spectral density N_0. Collecting the received signals from all antennas, the above received signal can be written in vector-matrix notation as

$$\mathbf{y}(t) = \mathbf{H}(\tau, t) \star \mathbf{x}(\tau) + \mathbf{n}(t). \tag{2.19}$$

When the channel is slowly varying and flat-fading, the variables τ and t can be removed because the channel matrix $\mathbf{H}(\tau, t)$ is constant over these variables. The received signal can thus be written as

$$\mathbf{y}(t) = \mathbf{H}\,\mathbf{x}(t) + \mathbf{n}(t), \tag{2.20}$$

where

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1N_T} \\ h_{21} & h_{22} & \dots & h_{2N_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_R1} & h_{N_R2} & \dots & h_{N_RN_T} \end{bmatrix}. \tag{2.21}$$

In scalar form, the received signal in (2.20) can be expressed as

$$y_i(t) = \sum_{j=1}^{N_T} h_{ij}\, x_j(t) + n_i(t), \qquad i = 1, 2 \dots N_R. \tag{2.22}$$

Assuming symbol-synchronous transmission, the symbol rate sampled, received signal at the output of the demodulator associated with the i-th antenna is a linear combination of the N_T transmitted signals perturbed by noise. It is expressed as

$$y_i = \sum_{j=1}^{N_T} \sqrt{\frac{E_{x,j}}{N_T}}\, h_{ij}\, x_j + \nu_i, \qquad i = 1, 2 \dots N_R, \tag{2.23}$$

where x_j is the $|\chi|$-ary symbol transmitted by the j-th transmitter and is chosen from a signal modulation such as PSK or QAM. For the sake of simplicity, all signals are assumed to be chosen from the same modulation. $E_{x,j}\,(j = 1, 2 \dots N_T)$ is the total average transmitted energy of the j-th transmitted signal during a symbol duration. The composite signal arriving at each receive antenna is perturbed by a sample of i.i.d. complex additive white Gaussian noise (AWGN). The samples of the AWGN process, denoted ν_i, have mean zero and variance σ_n^2 equal to the noise power spectral density, N_0. The channel gain from the j-th transmitter to the i-th receive antenna is represented by h_{ij} and follows a complex Gaussian distribution with zero mean and unit variance. This sampled, flat-faded equivalent system is depicted in Fig. 2.9.

The received signal in (2.23) can also be expressed in an equivalent matrix form as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\,\mathbf{x} + \boldsymbol{\nu}, \\ &= \sum_{j=1}^{N_T} \mathbf{h}_j\, x_j + \boldsymbol{\nu}, \end{aligned} \tag{2.24}$$

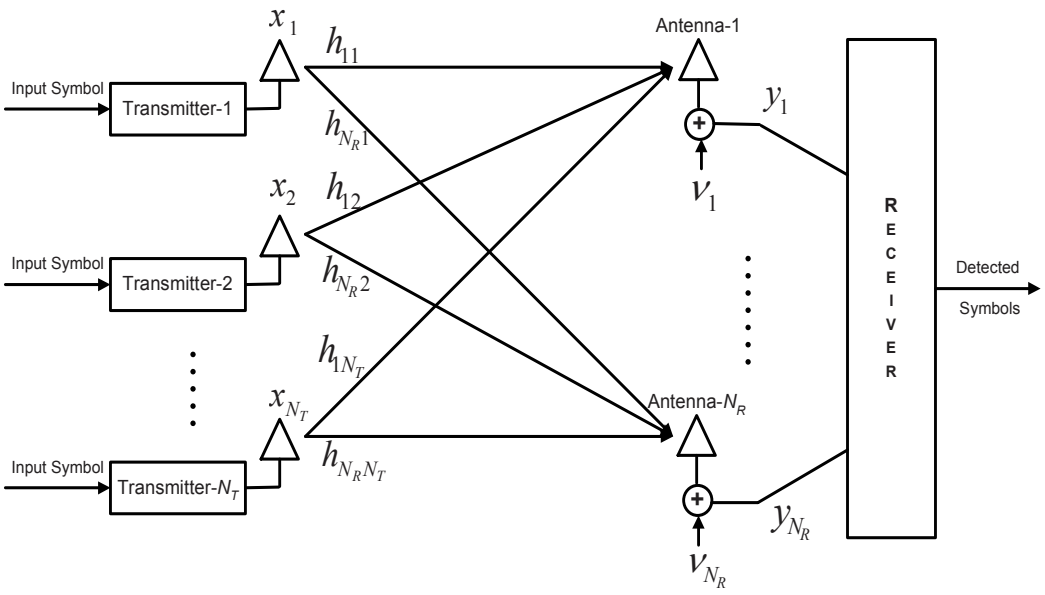

Figure 2.9 Block diagram of a slowly varying, flat-fading cochannel communication system.

where $\mathbf{x} = [x_1, x_2 \dots x_{N_T}]^T$ is the vector of transmitted symbols, $\mathbf{y} = [y_1, y_2 \dots y_{N_R}]^T$ is the received vector and $\boldsymbol{\nu} = [\nu_1, \nu_2 \dots \nu_{N_r}]^T$ is the AWGN noise vector. The channel matrix $\mathbf{H}$ was obtained in (2.21) and $\mathbf{h}_j$ is the j-th column of the $\mathbf{H}$ representing the array of channel gains between the j-th transmitter and the receiver.

2.6.2 System Loading Conditions

The performance of existing detection algorithms of the multi-antenna system in Fig. 2.9 depends on the number of antennas employed at the receiver. Using joint detection of the cochannel signals, it is shown in [23] that such a system offers good performance when the number of receiving antennas is greater than or equal to the number transmitted signals. Furthermore, performance slowly deteriorates when the number of transmitted signals exceed the number of receiver antennas. To describe this, the system load factor is defined as

$$L_F = \frac{N_T}{N_R}. \tag{2.25}$$

Based on this definition, the system can operate under one of the following three loading conditions:

1. Underloaded if $L_F < 1$.

2. Critically loaded if $L_F = 1$.

3. Overloaded if $L_F > 1$.

The first two cases can be combined to form a single condition $L_F \leq 1$ representing both underloaded and critically loaded systems. Most of the near-ML detection approaches for these systems depend upon QR factorization of the channel matrix [4, 29–31]. For such systems, the channel matrix is full-rank and QR decomposition produces an upper triangular matrix having as many non-zero rows as there are transmitted signals. In contrast, QR decomposition for a system operating under overload, where $L_f > 1$, results in only the first N_R rows of the upper triangular matrix being nonzero due to the ill-conditioned channel matrix[3]. Thus, the detection schemes developed for $L_f \leq 1$ are no longer applicable. Efforts are made to address this problem in [2–6, 8–11], but there is still a need to develop techniques that can provide near-ML performance with affordable complexity.

2.7 SUMMARY

In this chapter, the effect of multipath fading on wireless communication systems was examined. It was seen that the fading process introduces multiplicative distortion that severely degrades the system performance. Statistical Rayleigh and Ricean models which closely mimic the multipath fading phenomena in most real-world scenarios were discussed. The use of diversity to combat the destructive effects of fading were also examined along with different diversity approaches. Diversity combining methods were explained in detail. They act to gain benefit from the multiple independently faded copies of the signal available at the receiver. Multiple access methods are introduced for the purpose of better utilization of radio spectrum. The chapter ends with a discussion of spatial multiplexing systems where a system model for multiple signal cochannel communication systems is presented and possible operating environments of the system are outlined.

[3]For details, see Section 3.4.

Chapter 3

EXISTING DETECTION TECHNIQUES

3.1 INTRODUCTION

This chapter focuses on detection techniques for cochannel communication systems in Rayleigh flat-fading environments. As mentioned in Chapter 1, due to the simultaneous transmissions from multiple transmitters at the same carrier frequency, the received signal at each of the receiving antennas is a superposition of the transmitted signals. Multiuser detection (MUD) is the process of simultaneously estimating the data sent in all those transmissions. To do this, the receiver exploits an array of independent channel impulse responses between the transmitters and the receiver.

Signal detection techniques can primarily be classified into two main categories namely *linear* detection and *non-linear* detection. This classification is based on achievable performance and the associated computational efforts. At one extreme, the non-linear and optimum MAP and ML detection methods offer the best performance at the cost of complexity that varies exponentially with the number of transmitted signals. In contrast, the linear ZF and MMSE detectors save considerable computational effort, but the performance is significantly degraded compared to ML and MAP. Various attempts are made to develop algorithms which can bridge the performance gap between the optimum MAP/ML and the linear ZF/MMSE detectors with as much computational saving as possible.

Fig. 3.1 depicts the classification of the existing detection techniques including those developed during the course of this thesis. It is important to note that the performance of these detectors also depends on the loading condition under which the receiver is

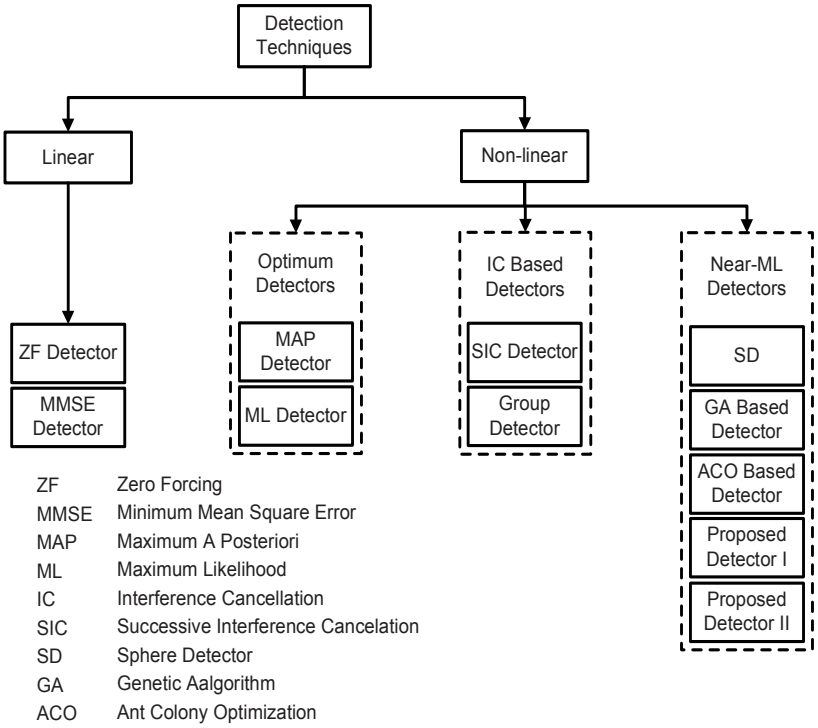

Figure 3.1 Classification of signal detection techniques.

operating. As an example, the performance of MMSE and ZF detectors degrades significantly in overloaded scenarios which limits their usage to underloaded and critically loaded system conditions.

Before starting the discussion of existing detection techniques, for the purpose of derivations and referencing in this chapter, the signal received at the receiver antenna array (Section 2.6) can be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}\,\mathbf{x} + \boldsymbol{\nu}, \\ &= \sum_{j=1}^{N_T} \mathbf{h}_j\, x_j + \boldsymbol{\nu}, \end{aligned} \tag{3.1}$$

where $\mathbf{x} \in \mathbb{C}^{N_T}$ is the transmitted symbol vector, $\mathbf{y} \in \mathbb{C}^{N_R}$ is the received signal vector and $\boldsymbol{\nu} \in \mathbb{C}^{N_R}$ is the AWGN noise vector. The matrix of channel gains interconnecting the transmitters and the receiver antenna array is denoted $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ and $\mathbf{h}_j \in \mathbb{C}^{N_T}$ is the j-th column of $\mathbf{H}$ which represents the vector of channel impulse responses between the j-th transmitter and the receiver. The objective of all the de-

tection techniques in this chapter is to obtain an estimate $\hat{\mathbf{x}}$ of the transmitted signal $\mathbf{x}$ based on the received signal $\mathbf{y}$ and assumed knowledge of the channel matrix $\mathbf{H}$.

In this chapter, fundamental signal detection techniques are described for the cochannel system described by (3.1). A systematic approach is followed where the optimum detection methods are discussed first in Section 3.2. This is followed by the suboptimum linear detection techniques in Section 3.3. The focus then shifts to other nonlinear approaches for signal detection. We consider sphere detection in Section 3.4, ordered successive interference cancellation detection in Section 3.5, group detection in Section 3.6, genetic algorithm based detection in Section 3.7 and ant colony optimization based detection in Section 3.8.

3.2 OPTIMUM DETECTION

3.2.1 Maximum A Posteriori Probability Detector

Assuming the availability of channel-state information (CSI) at the receiver, the optimum, maximum *a posteriori* probability (MAP) detector [23, 24] attempts to recover the most likely transmitted symbol vector $\mathbf{x}$ by maximizing the probability of correct decision

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x} \in \chi^{N_T}} P\left(\mathbf{x}|\mathbf{y}, \mathbf{H}\right), \tag{3.2}$$

where χ is the set of $|\chi|$ distinct complex symbols belonging to the linear modulation constellation employed for signaling. Also, χ^{N_T} represents the list of all possible symbol combination vectors or equivalently the entire search space

$$\chi^{N_T} = \left\{ \mathbf{x} = \left(x_1\, x_2 \ldots\, x_{N_T}\right)^T \mid x_j \in \chi \right\}. \tag{3.3}$$

Using Baye's theorem, the a posteriori probability in (3.2) can be written as [58, 65]

$$P\left(\mathbf{x}|\mathbf{y}, \mathbf{H}\right) = \frac{p\left(\mathbf{y}|\mathbf{x}, \mathbf{H}\right) P\left(\mathbf{x}\right)}{p\left(\mathbf{y}\right)}, \tag{3.4}$$

where $P(\mathbf{x})$ represents the *a priori* probability of $\mathbf{x}$ being the transmitted symbol vector

chosen from the entire search space. The a posteriori PDF of the received vector $\mathbf{y}$ conditioned on $\mathbf{x}$ and $\mathbf{H}$ is given by [23]

$$p(\mathbf{y}|\mathbf{x},\mathbf{H}) = \frac{1}{\left(\pi\sigma_\nu^2\right)^{N_T}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma_\nu^2}\right). \tag{3.5}$$

Since $p(\mathbf{y})$ does not depend on the detector argument, maximizing the a posteriori probability, $P(\mathbf{x}|\mathbf{y},\mathbf{H})$, is the same as maximizing the product $p(\mathbf{y}|\mathbf{x},\mathbf{H})P(\mathbf{x})$. Hence, (3.2) can be rewritten as

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}\in\chi^{N_T}} P(\mathbf{x}|\mathbf{y},\mathbf{H}) = \arg\max_{\mathbf{x}\in\chi^{N_T}} \left(p(\mathbf{y}|\mathbf{x},\mathbf{H})P(\mathbf{x})\right). \tag{3.6}$$

Using (3.5), the product, $p(\mathbf{y}|\mathbf{x},\mathbf{H})P(\mathbf{x})$, can be expressed as

$$p(\mathbf{y}|\mathbf{x},\mathbf{H})P(\mathbf{x}) = \frac{1}{\left(\pi\sigma_\nu^2\right)^{N_R}} \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma_\nu^2}\right) P(\mathbf{x}). \tag{3.7}$$

Taking the natural logarithm on both sides of (3.7), neglecting the terms independent of $\mathbf{x}$ and solving with appropriate sign changes, the simplified expression for the MAP detector can be obtained as [12]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\chi^{N_T}} \left[\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2 - \sigma_\nu^2 \ln\{P(\mathbf{x})\}\right]. \tag{3.8}$$

3.2.2 Maximum Likelihood Detector

When all possible symbol vectors in χ^{N_T} have equal probability of being chosen then the second term in (3.8) has no effect on the minimization and therefore it can be neglected. Then, the MAP detector reduces to the ML detector as [12, 65]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\chi^{N_T}} \Lambda(\mathbf{x}), \tag{3.9}$$

where

$$\Lambda(\mathbf{x}) = \|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2. \tag{3.10}$$

Note that the metric in (3.10) represents the Euclidean distance between the received signal vector, $\mathbf{y}$, and channel transformed hypothesized signal vector $\mathbf{x}$. This simplifies the geometrical interpretation of the ML receiver. From (3.9) and (3.10), it follows that the ML receiver looks for the signal vector $\mathbf{x}$ in χ^{N_T} which is closest to the received signal vector $\mathbf{y}$ using the standard Euclidean distance. The metric in (3.10) can also be expressed in equivalent form as

$$\Lambda\left(\mathbf{x}\right) = \sum_{i=1}^{N_R} \left| y_i - \sum_{j=1}^{N_T} h_{ij} x_j \right|^2 . \tag{3.11}$$

Clearly, the minimization in both the MAP and ML detectors is performed over all $|\chi|^{N_T}$ possible transmitted vectors, and therefore the complexity of the detector grows exponentially with the number of transmitted signals, N_T. In terms of the number of real operations, it can be easily verified that the evaluation of (3.10) for one candidate vector $\mathbf{x}$ takes $(4N_RN_T + 4N_R)$ real multiplications and $(4N_RN_T + 4N_R - 2)$ real additions. Thus, the overall complexity of the ML detector in terms of the number of real multiplications and additions, respectively, is given as

$$\mathcal{R}_{ML}^{\times} = \left(4N_RN_T + 4N_R\right)|\chi|^{N_T} \tag{3.12}$$

and

$$\mathcal{R}_{ML}^{+} = \left(4N_RN_T + 4N_R - 2\right)|\chi|^{N_T} . \tag{3.13}$$

This can be performed as a brute-force search which is computationally very expensive and impractical in most practical real-time scenarios.

3.3 LINEAR DETECTION

Due to enormous computational expenses, the optimal MAP and ML detectors can not be deployed in practical systems. As a result, various low-complexity suboptimum techniques have been developed offering acceptable performance. In this section, linear detection techniques are explored [18, 66], which significantly reduce the computational

burden on the receiver. The price paid is degradation in error-rate performance.

Linear detection comprises the class of receivers in which the symbol estimate $\tilde{\mathbf{x}}$ is found by passing the received signal $\mathbf{y}$ through a linear filter specified by a matrix $\mathbf{W}$ [41] to obtain

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{W}^H\mathbf{y} = \mathbf{W}^H\left(\mathbf{H}\mathbf{x}+\boldsymbol{\nu}\right)\\ &= \mathbf{W}^H\mathbf{H}\mathbf{x}+\mathbf{W}^H\boldsymbol{\nu},\end{aligned} \tag{3.14}$$

where the transformation matrix $\mathbf{W}$ is dependent on the channel $\mathbf{H}$. Each component of the filter output is then mapped to the nearsest constellation point to obtain the detected symbol vector $\hat{\mathbf{x}}$ as follows

$$\hat{x}_j = Q_\chi\left(\tilde{x}_j\right), \qquad j = 1, 2 \ldots N_T \tag{3.15}$$

where $Q_\chi(.)$ is defined as

$$Q_\chi\left(\tilde{x}_j\right) = \arg\min_{x\in\chi}\left|\tilde{x}_j - x\right|^2. \tag{3.16}$$

The j-th component of $\tilde{\mathbf{x}}$ in (3.14) which is the estimate of the j-th transmitted signal can also be written as

$$\begin{aligned}\tilde{x}_j &= \mathbf{w}_j^H\mathbf{y}\\ &= \underbrace{\mathbf{w}_j^H\,\mathbf{h}_j x_j}_{\text{signal}} + \underbrace{\sum_{\substack{l=1\\ l\neq j}}^{N_T}\mathbf{w}_j^H\,\mathbf{h}_l x_l}_{\text{interference}} + \underbrace{\mathbf{w}_j^H\boldsymbol{\nu}}_{\text{noise}},\end{aligned} \tag{3.17}$$

where $\mathbf{w}_j$ and $\mathbf{h}_j$ are the j-th columns of the matrices $\mathbf{W}$ and $\mathbf{H}$, respectively. The first term in (3.17) contains the desired signal while the middle term represents CCI. The last term is the transformed noise at the filter output. Keeping in view (3.17), each of the linear detectors computes $\mathbf{W}$ according to a particular detection criterion. The structure of a linear receiver is shown in Fig. 3.2.

Finally, the complexity of a linear detector depends on the linear transformation oper-

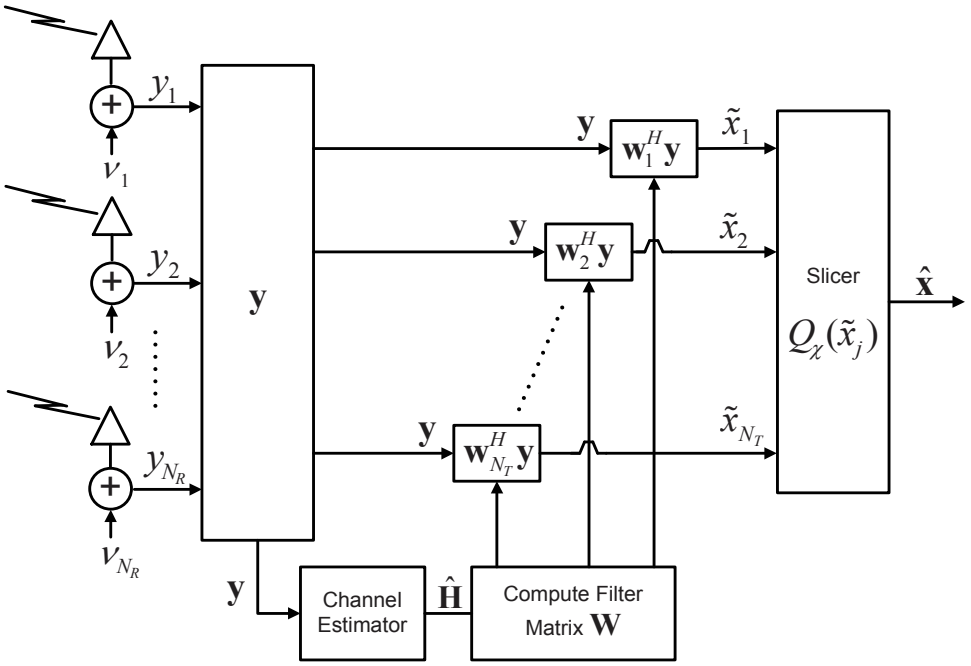

Figure 3.2 Receiver for linear detection techniques.

ation in (3.14) and the filter matrix $\mathbf{W}$ which is computed differently for each of the linear detectors. The transformation operation in (3.14) takes a total of $(4N_RN_T)$ real multiplications and $(4N_RN_T - 2N_T)$ real additions. Therefore, the overall complexity of a linear detector can be expressed in terms of the number of real multiplications and additions, respectively, as

$$\mathcal{R}_{LD}^{\times} = \mathcal{R}_{\mathbf{W}}^{\times} + 4N_RN_T \tag{3.18}$$

and

$$\mathcal{R}_{LD}^{+} = \mathcal{R}_{\mathbf{W}}^{+} + 4N_RN_T - 2N_T, \tag{3.19}$$

where $\mathcal{R}_{\mathbf{W}}^{\times}$ and $\mathcal{R}_{\mathbf{W}}^{+}$ denote the complexity in terms of number of real multiplications and additions, respectively, associated with calculation of matrix $\mathbf{W}$.

3.3.1 Zero-Forcing Detector

A zero-forcing (ZF) detector is a linear detector that aims to perfectly eliminate the CCI between the signals to be estimated. It solves the integer least squares problem by removing the discreteness constraint on the components of $\mathbf{x}$. The estimate of $\mathbf{x}$ is

performed using the filter matrix, $\mathbf{W}$ as [32, 66]

$$\begin{aligned}\mathbf{W} &= \left(\mathbf{H}^{+}\right)^{H} \\ &= \mathbf{H}\left(\mathbf{H}^{H}\mathbf{H}\right)^{-1},\end{aligned} \tag{3.20}$$

where $\mathbf{H}^{+} = \left(\mathbf{H}^{H}\mathbf{H}\right)^{-1}\mathbf{H}^{H}$ is the pseudo-inverse of the channel matrix $\mathbf{H}$. For overloaded scenarios, a more consistent method can be used to obtain the pseudo-inverse of the channel matrix [10]. Using the singular value decomposition $\mathbf{H} = \mathbf{V}_1\boldsymbol{\Omega}\mathbf{V}_2^{H}$ the pseudo-inverse is defined as [38]

$$\mathbf{H}^{+} = \mathbf{V}_2\bar{\boldsymbol{\Omega}}\mathbf{V}_1^{H}, \tag{3.21}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are left and right singular matrices of $\mathbf{H}$, and $\boldsymbol{\Omega}$ is a diagonal matrix containing ordered singular values $\{\omega_1, \omega_2 \ldots \omega_{N_R}\}$ on its main diagonal. The matrix $\bar{\boldsymbol{\Omega}}$ is defined using the singular values as

$$\bar{\boldsymbol{\Omega}}_{ii} = \begin{cases} \frac{1}{\omega_i}, & \omega_i > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{3.22}$$

The drawback of ZF detection is degraded BER performance due to noise enhancement. The AWGN noise $\boldsymbol{\nu}$ loses its whiteness property because it tends to be enhanced and correlated across the data streams corresponding to each of the transmitters. In addition, ZF detection achieves a diversity order of $N_R - N_T + 1$ out of a possible N_R for $N_T \times N_R$ system.

The complexity associated with the computation of the filter matrix $\mathbf{W}$ for a ZF detector is calculated in [67] to be

$$\mathcal{R}_{\mathbf{W}}^{\times} = 4N_T^3 + 8N_T^2N_R \tag{3.23}$$

and

$$\mathcal{R}_{\mathbf{W}}^{+} = 4N_T^3 + N_T^2\left(8N_R - 2\right) - 2N_RN_T. \tag{3.24}$$

Substituting $\mathcal{R}_{\mathbf{W}}^{\times}$ and $\mathcal{R}_{\mathbf{W}}^{+}$ into (3.18) and (3.19) yields the overall complexity of the ZF detector in terms of the number of real operations, which is clearly of the order of N_T^3.

3.3.2 Minimum Mean Square Error Detector

In minimum mean-square error (MMSE) detection, the transformation matrix $\mathbf{W}$ is formed in such a way that after applying the transformation, the combined effects of CCI, $\sum_{l=1,\, l\neq j}^{N_T} \mathbf{w}_j^H \mathbf{h}_l x_l$, and the filtered noise, $\mathbf{w}_j^H \boldsymbol{\nu}$, are minimized [66]. This leads to finding the filter matrix $\mathbf{W}$ according to the criterion [18]

$$\min_{\mathbf{W}} \mathrm{E}\left(J\left(\mathbf{W}\right)\right), \tag{3.25}$$

where the cost function $J\left(\mathbf{W}\right)$ represents the mean-square error, meaning

$$\begin{aligned} J\left(\mathbf{W}\right) &= \left\|\mathbf{x}-\tilde{\mathbf{x}}\right\|^2 \\ &= \left(\mathbf{x}-\tilde{\mathbf{x}}\right)^H \left(\mathbf{x}-\tilde{\mathbf{x}}\right) \\ &= \left(\mathbf{x}-\mathbf{W}^H\mathbf{y}\right)^H \left(\mathbf{x}-\mathbf{W}^H\mathbf{y}\right). \end{aligned} \tag{3.26}$$

Minimization of the cost function in (3.26) produces the optimum filter matrix $\mathbf{W}$ as [26, 28]

$$\mathbf{W} = \mathbf{H}\left(\mathbf{H}^H\mathbf{H} + \sigma_\nu^2 \mathbf{I}_{N_T}\right)^{-1}, \tag{3.27}$$

where $\mathbf{I}_{N_T}$ is an identity matrix of dimension N_T.

The MMSE detector offers superior BER performance compared to ZF detection because it minimizes the overall error energy instead of finding the channel inverse (which causes noise enhancement). The MMSE solution obtained from (3.14), (3.15) and (3.27) has much lower computational complexity than the MAP/ML solutions in (3.8) and (3.9), but this comes at the cost of significant performance degradation. This degradation is even more pronounced in the overloaded scenario considered in this thesis (when

$N_R < N_T$).

The expression for the MMSE filter matrix $\mathbf{W}$ in (3.27) is similar to that of the ZF filter matrix in (3.20) except for the term $\sigma_\nu^2 \mathbf{I}_{N_T}$ which has a small impact on $\mathcal{R}_{\mathbf{W}}^{+}$ only. The complexity for the filter matrix therefore is given as [67]

$$\mathcal{R}_{\mathbf{W}}^{\times} = 4N_T^3 + 8N_T^2 N_R \tag{3.28}$$

and

$$\mathcal{R}_{\mathbf{W}}^{+} = 4N_T^3 + N_T^2\,(8N_R - 2) - 2N_R N_T + N_T. \tag{3.29}$$

Substituting $\mathcal{R}_{\mathbf{W}}^{\times}$ and $\mathcal{R}_{\mathbf{W}}^{+}$ in (3.18) and (3.19) yields the overall complexity of the MMSE detector in terms of the number of real operations.

3.4 SPHERE DETECTOR

The ML detection problem can be rearranged and transformed into an integer least squares problem for which classical sphere decoding (SD) approaches exist in the literature [42, 68]. The main idea behind SD is to restrict the ML search space to within a hypersphere of radius r_0 around the received signal instead of performing a brute-force search of the entire lattice. This can significantly reduce complexity compared to an exhaustive search.

The sphere decoding technique begins by transforming the complex-valued system model described in (3.1) into an equivalent real-valued system model. To do this, the vector $\tilde{\mathbf{x}}$ composed of the real and imaginary parts of the vector $\mathbf{x}$ is defined as

$$\check{\mathbf{x}} = \left[\Re(\mathbf{x})^T \;\; \Im(\mathbf{x})^T\right]^T_{2N_T \times 1}. \tag{3.30}$$

Similarly, the complex received signal vector $\mathbf{y}$, the AWGN noise vector $\boldsymbol{\nu}$ and the complex channel matrix, $\mathbf{H}$, can be stated respectively as

$$\check{\mathbf{y}} = \left[\Re(\mathbf{y})^T \;\; \Im(\mathbf{y})^T\right]^T_{2N_R \times 1} \tag{3.31}$$

$$\check{\boldsymbol{\nu}} = \left[\Re(\boldsymbol{\nu})^T \;\; \Im(\boldsymbol{\nu})^T\right]^T_{2N_R \times 1} \tag{3.32}$$

and

$$\check{\mathbf{H}} = \begin{bmatrix} \Re(\mathbf{H}) & \Im(\mathbf{H}) \\ -\Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix}_{2N_R \times 2N_T}. \tag{3.33}$$

Hence, the real-valued equivalent version of the model of (2.24) can be expressed as

$$\check{\mathbf{y}} = \check{\mathbf{H}}\check{\mathbf{x}} + \check{\boldsymbol{\nu}}. \tag{3.34}$$

Finally, it can be shown that the ML detection rule for the equivalent real-valued channel model obtained in (3.34) is given by

$$\check{\mathbf{x}}_* = \arg\min_{\check{\mathbf{x}}} \|\check{\mathbf{y}} - \check{\mathbf{H}}\check{\mathbf{x}}\|^2, \tag{3.35}$$

where the components of $\check{\mathbf{x}}$ are integers and depend upon the constellation in use (for example, $\check{x}_j$ takes the values from the integer set $\{-3, -1, 1, 3\}$ if the constellation in use is 16-QAM).

Fincke and Pohst in [42] introduced the sphere decoding (SD) algorithm as an exact solution to the integer least-squares optimization problem which can therefore be applied to the detection problem in (3.35).

The transmitted symbol vector $\check{\mathbf{x}}$ is represented as a point on a rectangular $2N_T$-dimensional lattice. The MIMO channel $\check{\mathbf{H}}$ is assumed to be a lattice generating matrix. $\check{\mathbf{H}}\check{\mathbf{x}}$, which is $2N_R$-dimensional, spans the transformed skewed lattice. Hence considering a received vector $\check{\mathbf{y}}$ at N_R receivers and $\check{\mathbf{H}}\check{\mathbf{x}}$ the transformed lattice, the ML detection problem reduces to determining the nearest lattice point to $\check{\mathbf{y}}$ in terms of minimum Euclidean distance. The basic concept of SD is to search for the lattice points within a radius r_0 around the received vector $\check{\mathbf{y}}$, instead of traversing the complete search space. The SD searches for the point that gives the minimum Euclidean distance within the hypersphere with radius r_0, as shown in Fig. 3.3.

The resultant point in hypersphere with least distance to $\check{\mathbf{y}}$ is the solution returned by the SD algorithm. The selection of the radius r_0 and finding the nodes that fall inside

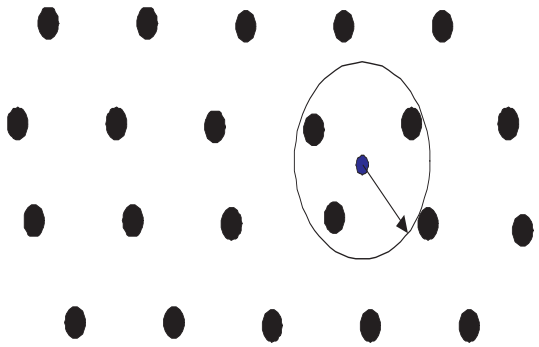

Figure 3.3 Geometric illustration of the sphere decoding algorithm.

this radius determine the complexity of the SD algorithm.

SD can also be viewed as branch and bound tree search algorithm [69]. To reorganize the ML detection into a tree like structure, first the channel matrix $\check{\mathbf{H}}$ in (3.35) is replaced with the product of a unitary matrix $\mathbf{Q}$ and an upper triangular matrix $\mathbf{R}$ using the well-known QR factorization [38]. The resulting expression is then pre-multiplied by $\mathbf{Q}^H$ to form the equivalent problem representation:

$$\check{\mathbf{y}} = \mathbf{QR}\check{\mathbf{x}} + \check{\boldsymbol{\nu}} \tag{3.36}$$

$$\bar{\mathbf{y}} \equiv \mathbf{Q}^H\check{\mathbf{y}} = \mathbf{Q}^H\left[\mathbf{QR}\check{\mathbf{x}} + \check{\boldsymbol{\nu}}\right] \tag{3.37}$$

$$\bar{\mathbf{y}} = \mathbf{R}\check{\mathbf{x}} + \bar{\boldsymbol{\nu}}, \tag{3.38}$$

where $\bar{\boldsymbol{\nu}} = \mathbf{Q}^H\check{\boldsymbol{\nu}}$. Now the detection rule of (3.35) becomes

$$\check{\mathbf{x}}_* = \arg\min_{\check{\mathbf{x}}} \|\bar{\mathbf{y}} - \mathbf{R}\check{\mathbf{x}}\|^2. \tag{3.39}$$

Note that due to the upper triangular nature of $\mathbf{R}$, the right hand side of (3.39) can be written as

$$\|\bar{\mathbf{y}} - \mathbf{R}\check{\mathbf{x}}\|^2 = \sum_{i=1}^{2N_T}\left|\bar{y}_i - \sum_{j=i}^{2N_T} R_{ji}\check{x}_j\right|^2. \tag{3.40}$$

As mentioned above, SD constrains the search process to within a sphere of radius r_0.

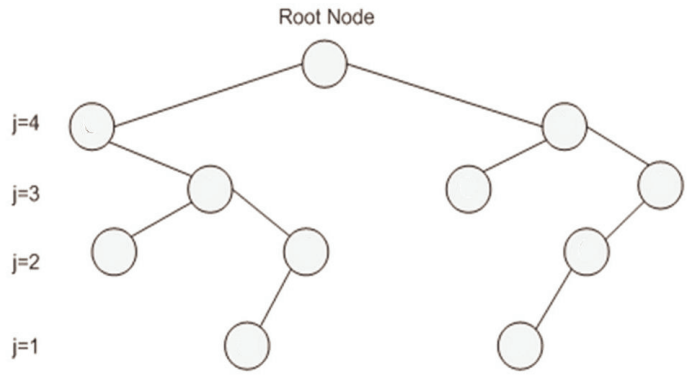

Figure 3.4 Tree generated by the branch and bound SD algorithm.

A candidate solution lies inside this sphere if and only if

$$r_0^2 \geq \sum_{i=1}^{2N_T} \left| \bar{y}_i - \sum_{j=1}^{2N_T} R_{ji} \check{x}_j \right|^2 \tag{3.41}$$

is satisfied for all the components $\check{x}_j$ of $\check{\mathbf{x}}$. A close look at (3.41) reveals the tree structure of ML detection problem where all the possible solutions in the search space can be mapped onto a $\log_2 |\chi|$-ary tree of depth $2N_T + 1$. Topmost level of the tree only contains the root node. Following either a depth-first or breadth-first strategy, the tree can be traversed from top to bottom, to form probable solutions. The sphere radius constraint provides a bound for the SD algorithm which helps in pruning unnecessary branches of the tree. The algorithm begin searching at the root node and then traverses the successive nodes ensuring that, at each node, the radius constraint in (3.41) is satisfied. In this way, the SD algorithm outputs the lattice points having least Euclidean distance from $\bar{\mathbf{y}}$ and lying within a sphere of radius r_0. As an example, Fig. 3.4 depicts a sample pruned tree constructed by the SD algorithm and selected lattice points in a four-dimensional sphere. The j-th level of the tree denotes dimension-j of the lattice point.

Significant research work has been done on SD to develop methods to efficiently search a lattice tree structure, such as lower bound pruning [27], lattice reduction [69] and

incrementing the search radius [70]. The worst case computational complexity of SD remains exponential in the number of transmitted signals[1] [70]. An exponential lower bound on SD average complexity has been derived in [36]. It is also important to note here that in the case of an overloaded system (i.e. when $L_F > 1$), the upper-triangular matrix $\mathbf{R}$ obtained by performing the QR factorization of the channel matrix $\check{\mathbf{H}}$ has zero elements in its lower half. Due to this fact the standard SD algorithm can not be directly applied. Efforts have been made to address this problem in [3–5, 8], but the complexity remain exponential in the number of excess signals[2]. Hence, there is still a need to develop techniques that can provide near-ML performance with affordable complexity.

3.5 SUCCESSIVE INTERFERENCE CANCELLATION DETECTOR

As suggested by the name, the successive interference cancelation (SIC) detector is based on the principle of interference cancellation. With a linear detector at its core, the SIC detector is designed to improve performance without a significant increase in computational complexity. The idea is to detect the transmitted signals sequentially while cancelling the effects of each of the detected signals from the one to be detected next. The structure of the SIC receiver consists of a bank of N_T linear receivers and N_T interference cancellation branches as shown in Fig. 3.5.

The detection process is completed in N_T stages where one signal is detected in each stage using a linear detection method as outlined in Section 3.3. For the purpose of explanation, MMSE detection in each stage is adopted in the rest of this section. The process starts by estimating the first signal, $\tilde{x}_{j_1}$, using the linear detector

$$\tilde{x}_{j_1} = \mathbf{w}_1^H \mathbf{y}_1, \tag{3.42}$$

where j_1 is the subscript associated with the signal to be detected, $\mathbf{w}_1 = (\mathbf{W})_{j_1}$ is the j_1-th column of the linear filter matrix $\mathbf{W}$ in (3.27), and $\mathbf{y}_1 = \mathbf{y}$. The estimated signal,

[1] The expected or average complexity of SD is regarded as cubic over certain ranges of SNR and problem dimensions [70]

[2] The number of signals exceeding the number of receive antennas is referred to as number of excess signals.

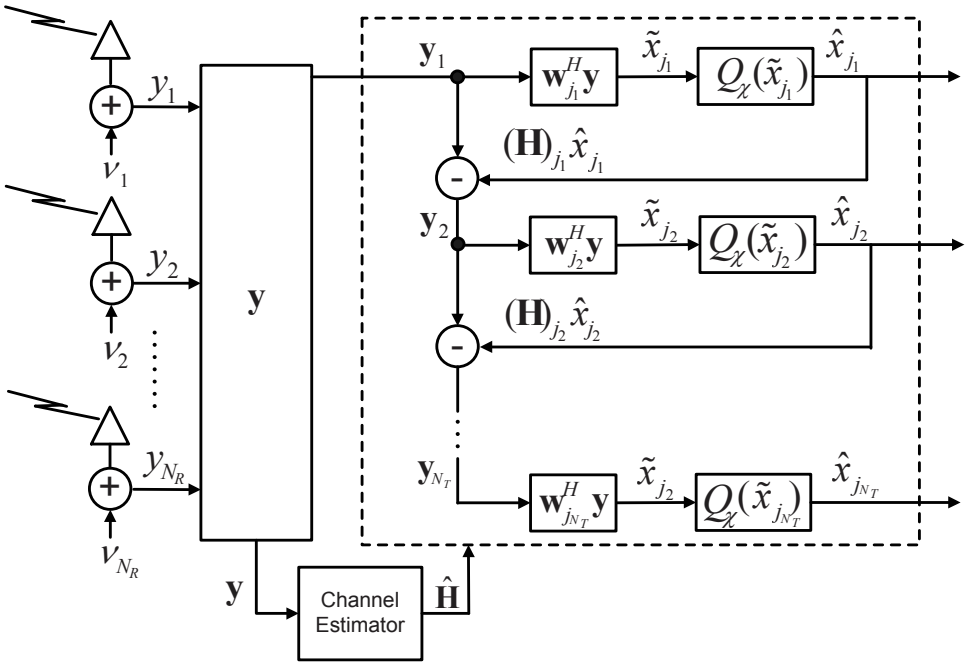

Figure 3.5 SIC receiver employing N_T linear receivers and interference cancellers for successive signal estimation.

$\tilde{x}_{j_1}$, is then mapped to the closest point in the signal constellation according to (3.15) to generate the hard output, $\hat{x}_{j_1}$.

Assuming $\hat{x}_{j_1}$ is correctly detected, the signal $\mathbf{y}_1$ is modified by removing the interference due to $\hat{x}_{j_1}$ in the next step to produce

$$\mathbf{y}_2 = \mathbf{y}_1 - (\mathbf{H})_{j_1} \hat{x}_{j_1}, \tag{3.43}$$

where $\mathbf{H}_{j_1}$ is the j_1-th column of the channel matrix $\mathbf{H}$.

To estimate the next signal, x_{j_2}, the updated filter matrix $\mathbf{W}_2$ is first calculated according to (3.27) using

$$\mathbf{W}_2 = \mathbf{H}_{\bar{j_1}} \left(\mathbf{H}_{\bar{j_1}}^H \mathbf{H}_{\bar{j_1}} + \sigma_\nu^2 \mathbf{I}_{N_T} \right)^{-1}, \tag{3.44}$$

where $\mathbf{H}_{\bar{j_1}}$ is the matrix obtained by deleting the j_1-th column of $\mathbf{H}$. The estimate of the signal x_{j_2} is then formed as

$$\tilde{x}_{j_2} = \mathbf{w}_2^H \mathbf{y}_2, \tag{3.45}$$

where j_2 is the subscript of the second signal, $\mathbf{w}_2 = (\mathbf{W}_2)_{j_2}$ is the j_2-th column of the linear filter matrix $\mathbf{W_2}$ in (3.44). The signal $\tilde{x}_{j_2}$ is then sliced to the nearest constellation point according to (3.15) to produce the detected signal, $\hat{x}_{j_2}$. This process continues until the last signal, $\hat{x}_{j_{N_T}}$, is detected. The filter matrix for the n-th stage is computed as

$$\mathbf{W}_n = \mathbf{H}_{\bar{j_n}} \left(\mathbf{H}_{\bar{j_n}}^H \mathbf{H}_{\bar{j_n}} + \sigma_\nu^2 \mathbf{I}_{N_T} \right)^{-1}, \tag{3.46}$$

where $\mathbf{H}_{\bar{j_n}}$ is a matrix obtained by deleting the columns $\{j_1, j_2 \dots j_n\}$ of the channel matrix $\mathbf{H}$.

It is important to note that, if $\hat{x}_{j_1} = x_{j_1}$, then the interference due to the first signal is perfectly canceled. However, if $\hat{x}_{j_1} \neq x_{j_1}$, error will be propagated to the next stage which may affect the decision on $\hat{x}_{j_2}$. In general, the error in deciding a particular signal during the detection process will propagate through the rest of the stages and will be amplified in case of further wrong decisions. For this reason, the order of signal detection has a vital impact on the overall performance achieved by the SIC receiver. A simple SNR based ordering criteria was presented in [71] while a more effective log-likelihood ratio (LLR) based detection ordering scheme was proposed in [72, 73]. In [74], it is shown that the optimum ordering is based on minimizing the symbol error rate (SER) and therefore achieves the best performance.

The complexity of the MMSE based SIC receiver in terms of the number of real multiplications and additions can be expressed, respectively, as [67]

$$\mathcal{R}_{SIC}^{\times} = N_T^2 \left(N_T + 1\right)^2 + \frac{4}{3} N_R N_T \left(N_T + 1\right) \left(2N_T + 1\right) + 8 N_R N_T \tag{3.47}$$

and

$$\mathcal{R}_{SIC}^{+} = \frac{N_T}{3} \left(3N_T^3 + 4N_T^2 \left(1 + 2N_R\right) + 3N_T \left(1 + 3N_R\right) + N_R - 1\right) + 2N_T \left(4N_R - 1\right). \tag{3.48}$$

3.6 GROUP DETECTOR

Another reduced complexity suboptimum approach to detection is based on the idea of group detection [75–77]. Here, a linear detector is first used to separate the transmitted signals and then a ML or MAP detector operates on subsets of the signals called *groups*. The subset G contains the indices of $|G|$ signals in the group and is sufficient to characterize the detector. The total number of groups is denoted $N_G = \frac{N_T}{|G|}$.

To specify the detection algorithm, notations and terminologies are first introduced. The channel matrix $\mathbf{H}$ is partitioned as $\mathbf{H} = [\mathbf{H}_G \ \mathbf{H}_{\bar{G}}]$, where $\mathbf{H}_G$ and $\mathbf{H}_{\bar{G}}$ are the submatrices of $\mathbf{H}$ corresponding to the signals inside and outside the group G, respectively. The matrix $\mathbf{C}$ is defined as $\mathbf{C} = \left(\mathbf{H}^H\mathbf{H}\right)^{-1}$ and $\mathbf{C}_G$ represents the submatrix of $\mathbf{C}$ corresponding to the signals in G.

Following the group detection idea, the ZF based group detector first applies the linear transformation $\mathbf{W} = \mathbf{H}^+ = \mathbf{H}\left(\mathbf{H}^H\mathbf{H}\right)^{-1}$ to the received signal vector as

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{W}\mathbf{y} = \mathbf{H}^+\mathbf{y} \\ &= \mathbf{x} + \check{\boldsymbol{\nu}},\end{aligned} \tag{3.49}$$

where the transformed noise $\check{\boldsymbol{\nu}}$ has a covariance of $\mathbf{K} = \sigma_\nu^2\mathbf{C}$. The resulting vector $\tilde{\mathbf{x}}$ can be partitioned as $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_G \ \tilde{\mathbf{x}}_{\bar{G}}]$, where $\tilde{\mathbf{x}}_G$ and $\tilde{\mathbf{x}}_{\bar{G}}$ are the subvectors of $\tilde{\mathbf{x}}$ corresponding to the signals in and outside group G, respectively. The linear filter output in (3.49) can now be expressed for group G as

$$\tilde{\mathbf{x}}_G = \mathbf{x}_G + \check{\boldsymbol{\nu}}_G, \tag{3.50}$$

where $\check{\boldsymbol{\nu}}_G$ is the transformed noise subvector having covariance $\mathbf{K}_G = \sigma_\nu^2\mathbf{C}_G$ [2]. The joint ML detector can be applied to $\tilde{\mathbf{x}}_G$ as [2]

$$\hat{\mathbf{x}}_G = \arg \min_{\mathbf{x}_G \in \chi^{|G|}} \left(\mathbf{x}_G - \tilde{\mathbf{x}}_G\right)^H \mathbf{K}_G^{-1} \left(\mathbf{x}_G - \tilde{\mathbf{x}}_G\right). \tag{3.51}$$

Since the search in (3.51) is performed on a subset of size $|G|$, the complexity therefore

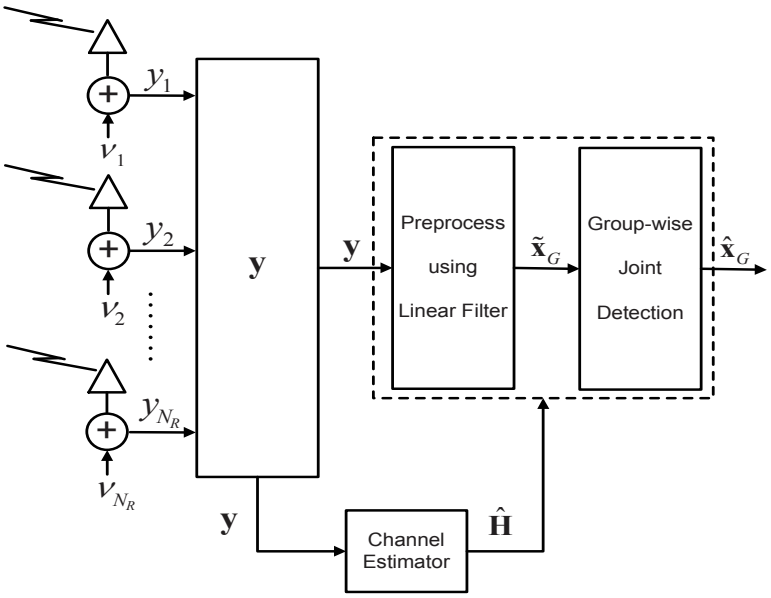

Figure 3.6 Group detection based receiver.

grows exponentially with group size. It is interesting to note that when $|G| = 1$, the group detector becomes the linear detector offering the performance of the ZF detector along with its complexity benefits. On the other hand, when $|G| = N_T$ it represents the ML detector with a complexity that varies exponentially with N_T. Hence, a tradeoff can be achieved by choosing a value of $|G|$ that can offer savings in complexity while performing within the acceptable range. The structure of the group detector consists of a linear preprocessing stage and a joint group detection phase as depicted in Fig. 3.6 and governed by (3.49) and (3.51), respectively.

The overall complexity of the group detector in terms of the number of real multiplications and additions can be expressed as

$$\mathcal{R}^{\times}_{GD} = \mathcal{R}^{\times}_{LD} + \mathcal{R}^{\times}_{JD} \tag{3.52}$$

and

$$\mathcal{R}^{+}_{GD} = \mathcal{R}^{+}_{LD} + \mathcal{R}^{+}_{JD}, \tag{3.53}$$

where $\mathcal{R}^{\times}_{LD}$ and $\mathcal{R}^{+}_{LD}$ are the complexities of the linear receiver in terms of the number of real multiplication and addition obtained in (3.18) and (3.19), respectively. The

terms $\mathcal{R}_{JD}^{\times}$ and $\mathcal{R}_{JD}^{+}$ represent the complexity (in terms of real multiplications and additions) of the joint group detection stage specified by (3.51).

To derive the expressions for $\mathcal{R}_{JD}^{\times}$ and $\mathcal{R}_{JD}^{+}$, note that the evaluation of the inverse of the group noise covariance matrix $\mathbf{K}_G$ and the entire vector-matrix product $(\mathbf{x}_G - \tilde{\mathbf{x}}_G)^H \mathbf{K}_G^{-1} (\mathbf{x}_G - \tilde{\mathbf{x}}_G)$, for each candidate vector $\mathbf{x}_G$ in each group is required. It can be easily shown that the complexity of (3.51) for one candidate vector is equal to $4|G|(|G|+1)$ real multiplications and $2|G|(2|G|+3)-2$ real additions. Also, the calculation of $\mathbf{K}_G^{-1}$ requires a total of $4|G|^3$ real multiplications and $4|G|^3$ real additions. Adding all these figures produces the complexity of the detection stage as

$$\mathcal{R}_{JD}^{\times} = N_G \left(4|G|^3\right) + N_G \left(4|G|^2 + 4|G|\right) |\chi|^{|G|} \tag{3.54}$$

and

$$\mathcal{R}_{JD}^{+} = N_G \left(4|G|^3\right) + N_G \left(4|G|^2 + 6|G| - 2\right) |\chi|^{|G|}, \tag{3.55}$$

which is clearly linear in the number of groups N_G, but exponentially varying with the group size $|G|$. Substituting (3.54) and (3.55) in (3.52) and (3.53) yields the overall number of complex multiplications and additions, respectively, performed by the ZF based group detection.

The group detection idea has been further explored and studied in [9] and [11]. In [9], an iterative soft signal cancelation and detection technique is presented for overloaded multiple-antenna systems employing BPSK modulation. It combines the idea of group detection with a posteriori probability (APP) extraction where soft symbol estimates of signals within each group are calculated and canceled from the received signal in an iterative manner. Complexity is proportional to the number of groups, but increases exponentially with group size.

In [11], a list-based group-wise iterative signal detection scheme is proposed for overloaded receivers. The use of BPSK and QPSK modulations is considered. The received signal is first preprocessed using a maximum ratio combiner and then fed to an iterative detection stage. Here, symbol groups are formed and passed through branch list estimation (BLE) and global list optimization (GLO) processes to extract a list

of candidate solution vectors. The complexity of this technique grows exponentially with the group sizes formed in the BLE and GLO processes. The detector maintains a near-ML performance at low SNR values for a system with $N_T = 9$ and $N_R = 6$. The computational cost for this system is reported as approximately half that of the ML detector. However, the presence of error-floor in the medium-to-high SNR region results in performance degradation in comparison to the ML detector.

3.7 GENETIC ALGORITHM BASED SOFT BIASED DETECTOR

The genetic algorithm (GA) is a global optimization technique first introduced in [78] and later comprehensively explored in [79–83]. It employs evolutionary processing based on natural selection and natural genetics. The GA based optimization is iterative in nature where each iteration is termed as a generation. In this section, the GA based signal detection technique for overloaded array processing systems presented in [6, 10] is described. The detection process begin by obtaining an initial guess (solution) either randomly or from a suboptimal detection approach. Later in every generation of the algorithm, a new set of candidate solutions called individuals are generated using portions of the fittest individuals obtained in the past[3]. Occasionally, some parts are independently modified with the intention of further improvements in the solution quality[4].

In biased GA detection, the initial population[5] of the GA algorithm is generated by biasing an initial soft estimate from a linear detector. This results in a fitter initial population improving the algorithm convergence. Keeping this in mind, the receiver structure is composed of a linear preprocessor and the GA processing engine. The preprocessor generates initial symbol estimates using a linear ZF detector which are then mapped on to soft bit estimates. The GA processing engine then biases the bit estimates to produce an initial population which has more diversity in the alleles[6] corresponding to the weaker bit estimates. This results in a judicious exploration of

[3]This process is termed as *crossover* in GA literature.
[4]A process called *mutation* in GA based optimization.
[5]Initial population is the population formed for the first generation of the algorithm.
[6]Alleles are the binary values representing a particular symbol from a signal constellation.

the less certain regions of the solution space.

A discussion on the biasing mechanism is presented in the remainder of this section whereas a more detailed stepwise explanation of the GA processing engine is provided in Appendix A. To facilitate the receiver using soft biasing capability, the soft estimate from a ZF filter is first obtained using (3.14) and (3.21) as

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{V}_1\bar{\mathbf{I}}\mathbf{V}_1^H\mathbf{x} + \mathbf{V}_1\bar{\mathbf{\Omega}}\mathbf{V}_2^H\boldsymbol{\nu} \\ &= \check{\mathbf{x}} + \check{\boldsymbol{\nu}},\end{aligned} \tag{3.56}$$

where $\bar{\mathbf{I}} = \bar{\mathbf{\Omega}}\mathbf{\Omega}$. Assuming the transmitted signals are chosen from a QPSK modulation, the real and imaginary parts of the j-th transmitted signal x_j, are denoted $x_{jR} = \mathrm{Re}\,(x_j)$ and $x_{jI} = \mathrm{Im}\,(x_j)$, respectively.

It can be shown using the property of truncated unitary matrices [84, 85] that the real and imaginary components of $\check{\mathbf{x}}$ have mean and variance given by

$$\mu_{\check{x}R} = \frac{N_R}{N_T}\,x_{jR}, \qquad \mu_{\check{x}I} = \frac{N_R}{N_T}\,x_{jI}$$

$$\sigma^2_{\check{x}R} = \sigma^2_{\check{x}I} = \frac{1}{2}\left[\frac{N_R}{N_T} - \left(\frac{N_R}{N_T}\right)^2\right]. \tag{3.57}$$

Also, the real and imaginary parts of $\check{\boldsymbol{\nu}}$ are zero-mean, $\mu_{\check{\nu}R} = \mu_{\check{\nu}I} = 0$, with the approximated variance [10],

$$\sigma^2_{\check{\nu}R} \approx \sigma^2_{\check{\nu}I} = \frac{\sigma^2_\nu}{2}\left(\frac{N_R^2}{N_T^2 - N_R N_T}\right). \tag{3.58}$$

Using (3.57) and (3.58), the probability of x_{jR} and x_{jI} being positive, conditioned on the soft bit estimates $\mathrm{Re}\,(\tilde{x}_j)$ and $\mathrm{Im}\,(\tilde{x}_j)$, respectively, can be found using Bayes' rule as [10]

$$p_{jR} = \left[1 + \exp\left(-\frac{2\,\mathrm{Re}\,(\tilde{x}_j)\,\mu_{\check{x}R}}{\left(\sigma^2_{\check{x}R} + \sigma^2_{\check{\nu}R}\right)}\right)\right]^{-1} \tag{3.59}$$

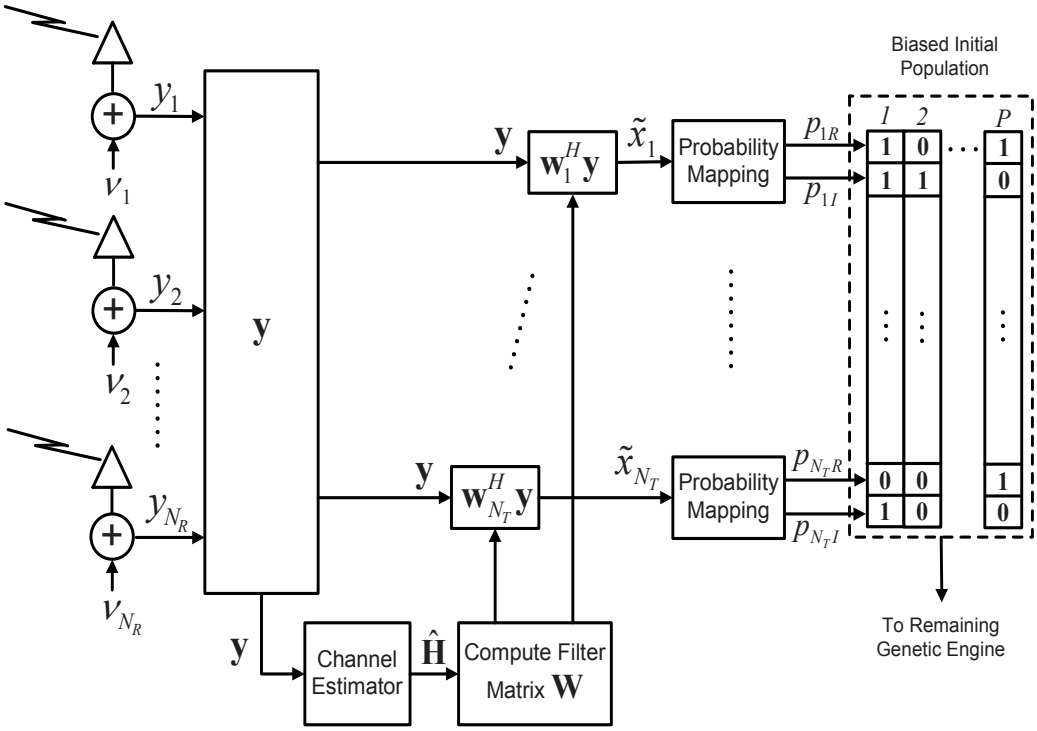

Figure 3.7 Initialization phase of the biased GA detector.

and

$$p_{jI} = \left[1 + \exp\left(-\frac{2\,\mathrm{Im}\left(\tilde{x}_j\right)\mu_{\check{x}I}}{\left(\sigma^2_{\check{x}I} + \sigma^2_{\check{\nu}I}\right)}\right)\right]^{-1} \tag{3.60}$$

Using (3.59) and (3.60), the soft bit estimates corresponding to a transmitted +1 can be stacked into vector as

$$\mathbf{p}_{so} = \left[p_{1R}, \quad p_{1I}, \quad p_{2R}, \quad p_{2I} \quad \cdots \quad p_{N_T R}, \quad p_{N_T I}\right]^T. \tag{3.61}$$

As investigated in [10], the soft values in (3.59) and (3.60) depend on the population size, P_T, and the operating SNR. Therefore, the initial population is biased using the $\mathbf{p}_{so}$ that results from tuning the SNR (through $\sigma^2_{\check{\nu}R}$ in (3.59)) to a value which offers minimum BER. In the case of a system with $N_R = 5$ and $N_T = 6$ and 8, the BER is minimum at a reference SNR of 10 dB and so the estimate in (3.61) is chosen to bias the initial population for this SNR value. Fig. 3.7 depicts the structure of the GA detection based receiver along with the initialization phase which biases the seed

solution produced by a linear receiver.

The GA detector processes a total of G_T generations. In each of these generations, a total of P_T candidate solutions (or individuals) are created according to the procedure outlined in Appendix A. The fitness or quality of each of these solutions is evaluated according to the fitness function

$$\Lambda_p\left(\bar{\mathbf{x}}^p\right) = \left\|\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^p\right\|^2, \qquad p = 1, 2 \ldots P_T, \tag{3.62}$$

where $\bar{\mathbf{x}}^p = \left[\bar{x}_1^p \;\; \bar{x}_2^p \ldots \bar{x}_{N_T}^p\right]^T$ represents the p-th candidate solution of the generation under process. When the termination criteria is met, the solution corresponding to the best fitness value is returned as the detector output[7].

Based on the fact that the GA detector is also composed of a linear filter stage and a genetic processing phase, the complexity of the GA detection in terms of the number of real multiplications and additions can respectively be expressed as

$$\mathcal{R}_{GA}^{\times} = \mathcal{R}_{LD}^{\times} + \mathcal{R}_{GP}^{\times} \tag{3.63}$$

and

$$\mathcal{R}_{GA}^{+} = \mathcal{R}_{LD}^{+} + \mathcal{R}_{GP}^{+}, \tag{3.64}$$

where $\mathcal{R}_{LD}^{\times}$ and $\mathcal{R}_{LD}^{+}$ are the complexities of the linear receiver in terms of the number of real multiplications and additions, respectively. Furthermore, $\mathcal{R}_{GP}^{\times}$ and $\mathcal{R}_{GP}^{+}$ represent the complexity (in terms of real multiplications and additions) of the genetic processing stage. The genetic engine processes G_T generations each having P_T individuals and performs a total of $P_T G_T$ fitness evaluations in (3.62). Since one evaluation of the fitness function results in $(4N_R N_T + 4N_R)$ real multiplications and $(4N_R N_T + 4N_R - 2)$ real additions, the complexity of a genetic processing stage can be expressed as [10]

$$\mathcal{R}_{GP}^{\times} = \left(4N_R N_T + 4N_R\right) P_T G_T \tag{3.65}$$

[7]The reader is referred to Appendix A for more details on individual components of the GA processing engine.

and

$$\mathcal{R}_{GP}^{+} = (4N_R N_T + 4N_R - 2)\, P_T G_T. \tag{3.66}$$

Replacing these values in (3.65) and (3.66) produces the overall complexity of the GA based detector which clearly depends on P_T and G_T.

In [10], it is shown that the GA detector requires $G_T = 12$ generations and a population size of $P_T = 1024$ in order to achieve performance close to that of the ML detector (in low to slightly moderate SNR range) for an overloaded system with $N_T = 8$, $N_R = 5$ employing QPSK modulation. This results in $P_T \times G_T = 1024 \times 12 = 12288$ evaluations of (3.62). Furthermore, the detector starts producing an error floor at a moderate SNR due to the poor performance of the ZF filter.

3.8 ANT-COLONY OPTIMIZATION BASED DETECTOR

Earlier investigations and theoretical findings on behavioral aspects of social insects lead to the development of some powerful metaheuristic optimization techniques. Among many of those, ant colony optimization (ACO) have emerged as a popular combinational optimization technique. It is based on the collective social behavior of a colony of ants. In ACO, the notion of artificial ants along with the artificial *stigmergic model*[8] is designed to tackle a given optimization problem. An artificial ant or simply an ant is modeled as a computational agent which probabilistically creates a solution to the problem. To construct a solution, the metaheuristic uses incremental strategy where solution components[9] are added one after the other to a partially constructed solution. While a comprehensive tutorial on the ACO metaheuristic is available in Appendix B[10], a brief discussion on the application of ACO to near-ML MIMO detection which appeared in [31] is provided in the remainder of this section.

The algorithm presented in [31] formulates the MIMO detection as traveling salesman problem (TSP) where the transmit antennas denote the set of cities and the paths

[8] A model which provides the ants means of communication by modifying their surrounding environment.

[9] Note that each of the solution component is chosen from the entire set of available options.

[10] Interested reader may refer to the appendix for familiarizing themselves with the ACO concepts and terminologies.

between the consecutive cities are represented by the set of all possible transmitted signals from a particular antenna. The objective is to find a minimum-cost path over this TSP representation of the detection problem. The ants start the search process at a particular city and move to the next city by probabilistically selecting a particular path. This results in a partially constructed solution. The search process is completed when all the ants visit all the possible cities. The transition probability for a particular move depends on the associated pheromone strength and the heuristic value. The rules for managing pheromone are governed by the type of ACO algorithm employed whereas the heuristic value depends on the underlying heuristic criteria defined for the detection problem. A well-designed heuristic criteria and the corresponding heuristic value results in superior performance and convergence of the algorithm.

To apply ACO to the MIMO detection problem, the first step is to formulate an effective method for computation of the heuristic values. For this purpose, the real-valued equivalent system model (3.30)-(3.34) which is already developed in Section 3.4 is considered. Using the QR factorization of the real-valued channel matrix, $\check{\mathbf{H}} = \mathbf{QR}$, the ML detection problem is first expressed in the equivalent form as derived in (3.39)

$$\check{\mathbf{x}} = \arg\min_{\check{\mathbf{x}} \in \check{\chi}} \|\bar{\mathbf{y}} - \mathbf{R}\check{\mathbf{x}}\|^2. \tag{3.67}$$

where $\check{\chi}$ is the set of integers which depend on the modulation in use (e.g. $\check{\chi} = \{-3, -1, 1, 3\}$ for 16-QAM modulation). Due to the upper triangular structure of $\mathbf{R}$, the right hand side of (3.67) can also be expressed as

$$\|\bar{\mathbf{y}} - \mathbf{R}\check{\mathbf{x}}\|^2 = \sum_{i=1}^{\check{N}_T} \left| \bar{y}_i - \sum_{j=i}^{\check{N}_T} R_{ij}\ \check{x}_j \right|^2 . \tag{3.68}$$

where $\bar{\mathbf{y}} = \mathbf{Q}^H \check{\mathbf{y}}$ and $\check{N}_T = 2N_T$.

In ACO, A ants are assigned the task of generating A candidate solutions $\bar{\mathbf{x}}_a$, $a = 1, 2 \ldots A$. Each solution represents one way of visiting each city exactly once. Collectively, all the A candidate solutions (representing A different ways of visiting $\check{N}_T$ cities) along with the associated costs (from the first city to the last one) are referred

to as trails. A total of $|\check{\chi}| = \log_2 |\chi|$ paths are available to each ant for moving to the i-th city, where 'path $\check{x}_{ij}$' denotes the path corresponding to signal $\check{x}_j \in \check{\chi}$ (which is selected as i-th component of the solution vector under construction based on the ant's transition probability). Using (3.68), the distance associated with 'path $\check{x}_{ij}$' is defined as

$$d_{ij} = \left| \bar{y}_i - \sum_{l=i+1}^{\check{N}_T} R_{il}\, \check{x}_l - R_{ii} x_j \right|^2 . \tag{3.69}$$

where $i = \check{N}_T \ldots 1$ and $j = 1 \ldots |\check{\chi}|$. The solution construction begins by selecting first component of the solution vector $\check{x}_{\check{N}_T}$. At the selection of i-th signal, the contribution of the previously selected solution components (i.e. $\check{x}_{i+1} \ldots \check{x}_{\check{N}_T}$) forming the middle term of (3.69) are used along with the $|\check{\chi}|$ available choices as the hypothesis for $\check{x}_{i+1}$ represented by the last term of (3.69). Finally, the distance is converted to heuristic value using the log-sigmoid function as

$$\eta_{ij} = \frac{1}{1 + \exp\left(d_{ij}\right)}. \tag{3.70}$$

Using the heuristic value η_{ij} and the pheromone levels τ_{ij} associated with 'path $\check{x}_{ij}$', the ant's transition probability to the i-th city is calculated according to [86]

$$p_{ij} = \frac{(\tau_{ij})^{\alpha} \cdot (\eta_{ij})^{\beta}}{\sum_{l=1}^{|\check{\chi}|} (\tau_{il})^{\alpha} \cdot (\eta_{il})^{\beta}}, \tag{3.71}$$

where α and β are ACO parameters for controlling the significance of heuristic desirability and the pheromone strengths, respectively. Upon the completion of A solutions $\{\check{\mathbf{x}}\}_{a=1}^{A}$ by the ants, the pheromone levels on each of the paths is updated using the ant system (AS) algorithm[11] as

$$\tau_{ij} = (1 - \rho)\, \tau_{ij} + \sum_{a=1}^{A} \Delta\tau_{ij}^{a}, \tag{3.72}$$

where $\rho \in (0, 1]$ represents the pheromone decay parameter and the incremental pheromone

[11] Refer to Section B.4 for details on AS algorithm.

trail is defined using the likelihood value of the generated solutions as

$$\Delta\tau_{ij}^{a} = \begin{cases} \left\|\check{\mathbf{y}} - \check{\mathbf{H}}\check{\mathbf{x}}_a\right\|^2 & \text{if ‘path } \check{x}_{ij}\text{’} \in \check{\mathbf{x}}_a \\ 0 & \text{otherwise.} \end{cases} \tag{3.73}$$

In the modified ACO (MACO) version of the ACO algorithm in [31], each ant is allowed to maintain its own distance, heuristic value and the transition probability to each city. Stepwise details of the algorithm are provided below:

MACO $\left(\check{\mathbf{y}}, \mathbf{R}, \check{N}_T, |\check{\chi}|, A\right)$

1. Initialize the variables $a = 1$, $i = \check{N}_T$ and $bestdist = \infty$.

2. Compute the distances and heuristic values for the a-th ant corresponding to ‘path $\check{x}_{ij}$’ as

$$d_{ij}^{a} = \left| \bar{y}_i - \sum_{l=i+1}^{\check{N}_T} R_{il}\ \check{x}_l^a - R_{ii}x_j \right|^2. \tag{3.74}$$

$$\eta_{ij}^{a} = \frac{1}{1 + \exp\left(d_{ij}^{a}\right)}. \tag{3.75}$$

3. The a-th ant chooses one of the $|\check{\chi}|$ paths to the i-th city according to the transition probability

$$p_{ij}^{a} = \frac{\left(\eta_{ij}^{a}\right)^{\beta}}{\sum_{l=1}^{|\check{\chi}|}\left(\eta_{il}^{a}\right)^{\beta}}, \tag{3.76}$$

4. Decrement i by 1 and repeat steps 1-3 until $i = 1$.

5. Compute the likelihood value of candidate solution generated by the a-th ant

$$\check{\Lambda}_a = \left\|\check{\mathbf{y}} - \check{\mathbf{H}}\check{\mathbf{x}}_a\right\|^2 \tag{3.77}$$

6. Update the *bestdist* according to the rule

$$bestdist = \check{\Lambda}_a \text{ if } \check{\Lambda}_a < bestdist \tag{3.78}$$

7. Increment a if $a < A$, set $i = \check{N}_T$ and return to step 2. Otherwise the solution is found as $\check{\mathbf{x}} \in bestdist$.

The simulation results for a cochannel system with $N_T = 16$ employing 16-QAM modulation and number of receive antennas $N_R = 16$ suggest that the MACO algorithm achieves near-ML performance with considerable complexity savings. At an SNR of 22dB, the MACO algorithm requires $A = 20000$ ants to perform close to ML. Different ACO parameters are set as $\alpha = 0.5$, $\beta = 1, 2$ and $\rho = 0.3$. The drawback of this algorithm, however, is its inability to handle the more challenging case of overload, i.e. when $L_f > 1$. In such a case, the QR decomposition of the channel matrix $\check{\mathbf{H}}$ produces an upper-triangular matrix $\mathbf{R}$ having only the first N_R rows with non-zero entries. Due to this reason, the distance metric in (3.69) is no longer applicable.

3.9 SUMMARY

In this chapter, the signal detection problem is explored and the existing optimum and suboptimum solutions are discussed. It is seen that the complexity of optimal detectors is exponentially varying with the number of transmitted signals which prohibits their implementation in most real-time scenarios where the number of transmitted signals is large and higher order modulations are used.

To overcome the complexity problem inherent in the optimum approaches, a range of suboptimum detection techniques are discussed based on certain assumptions about the underlying system. It is seen that the linear ZF and MMSE [18, 28, 32] receivers provide large savings in the required computational efforts, but their performance is significantly degraded, especially when the system is overloaded [87]. To improve the performance of the linear receivers, an interference cancelation based detection approach can be used which successively estimates each of the transmitted signals and removes its contribution from the received signal before making a decision about the next signal [18, 71]. Since detection ordering is critical to the overall performance achieved by the interference cancelation based detectors, the order in which signals are detected must be carefully determined [72–74]. Overall, the SIC detector offers slightly

better performance than the linear receivers at the cost of increased complexity.

To further improve the performance, researchers have proposed a group detection approach [2, 75–77] and a GA based detector [6, 10]. It is seen that the group detector offers a varying performance-complexity tradeoff by tuning the group size parameter. The complexity is exponential in the group size so it is not suitable for large systems where a larger group size is required for acceptable performance. The GA based detector of [10] biases the initial population by obtaining an initial estimate from a linear detector. Since the approach is based on the ZF filter, it presents an error floor in the moderate to high SNR regime as the system loading factor increases.

Finally, the application of ACO to signal detection for multi-antenna systems is discussed. The algorithm presented in [31] is explored and the associated issues that prohibit its implementation in overloaded scenarios (where $L_f > 1$) is explained in significant detail.

Based on these discussions and observations, there is still a need to develop detection techniques (particularly for overloaded receivers) which can provide near-ML performance with further reduced and affordable complexity to be implemented in real-time scenarios. This problem is addressed by this thesis and our solutions are presented in the subsequent chapters.

Chapter 4

DETECTION OF COCHANNEL SIGNALS USING ANT COLONY OPTIMIZATION

4.1 INTRODUCTION

Since the search space for the optimal MAP or ML detector is discrete with a solution having integer components, it is therefore an integer least-squares optimization problem [30, 41]. Furthermore, it also belongs to the class of NP-hard combinatorial optimization problems [24, 30, 88].

An alternative approach is to consider metaheuristic techniques in order to achieve near optimal performance at significantly reduced complexity. As explained in Appendix B, ACO is a metaheuristic algorithm [89] based on the behavior of a colony of ants searching for food. Real ants randomly explore the space around their nest in search of food. Once an ant finds the food source, it starts the journey back to the nest thereby depositing pheromone along the path. This deposition of the pheromone is a mechanism for communication between the ants. It is used by ants to inform each other about the path they follow to the food source (*stigmergy*). With the passage of time, the ants following the shorter route make more trips to the food source, thus contributing to increasing pheromone level along that path. As a result, in the steady state, most of the ants trace the shortest path with a high likelihood.

An ant is modeled as a computational agent which explores the entire solution space and builds a candidate solution to a given problem [13]. It does so by adding components of the solution (in a sequential manner) to a partially constructed solution. Each solution component, which belongs to a set containing all the possibilities, is probabilis-

tically selected based on a priori available problem-specific heuristic information and the artificial pheromone trails resulting from the search experience of the ants. Fitness evaluation is performed after the solution is constructed to determine its suitability. Pheromone values are updated based on the fitness value of each solution before the start of the next iteration.

During the initial stages of the ACO search process, the pheromone concentration along different paths is almost the same, and therefore, ACO uniformly explores the search space mainly based on the heuristic information (in our case soft symbol estimates). Thus, the algorithm shows explorative behavior during this period. In later phases, the ants keep depositing the pheromone on ever more attractive paths[1] and evaporating the pheromone from less suitable paths. To formulate a better solution, during this phase, ACO uses the pheromone more dominantly than the heuristic information in making a decision to choose a particular path. Note that in the steady state, the pheromone contains the information about the best solution found so far. Thus the algorithm exploits previous knowledge about good solutions and hence tends to converge to a more refined and higher-quality solution. Furthermore, pheromone deposition and evaporation are the operations through which the search process is controlled in ACO.

This chapter starts with specifying the system model in Section 4.2. A soft iterative ACO based detection algorithm for binary modulation is presented in Section 4.3. The heuristic framework of the developed receiver and its integration with the ACO are explained. A discussion on computational complexity of the proposed scheme is provided along with the simulation results. In Section 4.4, a sub-optimum ACO assisted detector for non-binary modulations is developed. The heuristic criteria used for non-binary modulations is specified and the ACO search process is discussed in detail. Expressions for the computational complexity are developed and simulation results are presented. At the end, a chapter summary is provided in Section 4.5.

[1] A path is said to be attractive if it offers smaller fitness value than the available alternatives.

4.2 SYSTEM MODEL

We consider a cochannel communication system where N_T signals are impinging on a receiver antenna array having N_R elements. Each transmitted signal x_j is assumed to be chosen from a $|\chi|$-ary modulation χ. The complex channel gain between the j-th transmitter and the i-th receiver antenna element is denoted as h_{ji} and assumed to be i.i.d. Rayleigh distributed random variable with mean zero and variance of 0.5 per dimension. The signals collected from all the receive antennas can be arranged in a vector form and expressed using an input-output relation

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\nu}, \tag{4.1}$$

where $\mathbf{y} \in \mathbb{C}^{N_R}$ represents the received signal vector and $\boldsymbol{\nu} \in \mathbb{C}^{N_R}$ is the AWGN noise vector. The elements of $\boldsymbol{\nu}$ are assumed to be temporally white complex Gaussian with $E\left[\boldsymbol{\nu}\boldsymbol{\nu}^H\right] = \sigma_\nu^2 \mathbf{I}_{N_T}$. Furthermore, $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ represents a matrix of channel gains interconnecting the transmitters and the receiver antennas.

4.3 ACO ASSISTED DETECTOR FOR BINARY SIGNALS

ACO metaheuristic is employed in a number of important classes of optimization problems namely the routing problems[2], assignment problems[3], subset problems[4], machine learning problems[5] and the network routing problems[6] [86, 90].

In context of signal detection, ACO has already been used to perform MUD in CDMA systems [50–52]. Recently, it has been employed for near optimum detection in MIMO systems [31]. The problem however with this approach is that it depends on the QR factorization of the channel matrix and therefore cannot easily handle the singularity problem seen in the overloaded scenario. To overcome this problem, we developed an

[2]This includes the classical traveling salesman problem (TSP) and the sequential ordering.

[3]The quadratic assignment, graph coloring and generalized assignment problems fall under this category.

[4]Some of the examples include multiple knapsack, set covering and weight constrained graph tree partition problems.

[5]Examples are classification rules, baysian networks and the fuzzy systems.

[6]This includes the optical network routing, connection-oriented and connectionless network routing.

ACO based detection method for binary phase shift keying (BPSK) modulated signal transmissions. This is described in detail below.

4.3.1 The Heuristic Framework

We now present our soft decision based ACO technique for near-ML detection of the cochannel signals assuming BPSK transmissions [87]. Note that our proposed receiver structure has the ability to perform consistently even when the system is operating under overload. Furthermore, it offers near ML performance with significant reduction in the computational complexity.

Recall the k-th BPSK modulated symbol is denoted $x_k \in \pm 1$. The soft estimate of the k-th transmitted signal can be expressed as [91]

$$\tilde{x}_k = \ln\left(\frac{\sum_{\mathbf{x}\in\chi_{+1}^k} \bar{p}(\mathbf{y}|\mathbf{x},\mathbf{H})}{\sum_{\mathbf{x}\in\chi_{-1}^k} \bar{p}(\mathbf{y}|\mathbf{x},\mathbf{H})}\right), \tag{4.2}$$

where $x_k \in \chi$, $\chi_b^k \subseteq \chi^{N_T}$ represents the set of all possible solutions in which the k-th transmitted signal, $x_k = b$ ($b = \pm 1$), and is defined as

$$\chi_b^k = \left\{\mathbf{x} = (x_1 \dots\ x_k \dots\ x_{N_T})^T \mid x_j \in \chi,\ x_k = b\right\}. \tag{4.3}$$

The conditional PDF of $\mathbf{y}$ is already obtained (Section 3.2) as

$$\bar{p}(\mathbf{y}|\mathbf{x},\mathbf{H}) = \frac{1}{(\pi\sigma_\nu^2)^{N_T}}\ \exp\left(-\frac{\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{\sigma_\nu^2}\right). \tag{4.4}$$

Therefore, substituting (4.4) into (4.2) and solving produces the soft estimate of the k-th transmitted signal as

$$\tilde{x}_k = \ln\left[\sum_{\mathbf{x}\in\chi_{+1}^k} \exp\left(\frac{-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{2\sigma_\nu^2}\right)\right] - \ln\left[\sum_{\mathbf{x}\in\chi_{-1}^k} \exp\left(\frac{-\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2}{2\sigma_\nu^2}\right)\right] \tag{4.5}$$

Here, the evaluation of the exponential function for all the $|\chi|^{N_T-1}$ terms in the summations imposes excessive computational efforts for large values of N_T. To avoid this,

only the most dominant term can be considered using the so called max-log MAP approximation [92] for updating the soft estimation of the transmitted signals. Hence (4.2) can be approximated by

$$\tilde{x}_k \approx \frac{1}{2\sigma_\nu^2}\left[\min_{\mathbf{x}\in\chi_{-1}^k}\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2 - \min_{\mathbf{x}\in\chi_{+1}^k}\|\mathbf{y}-\mathbf{H}\mathbf{x}\|^2\right] \tag{4.6}$$

In our soft ACO detection algorithm [87], a group of artificial ants performs a sequence of operations iteratively to construct a list of candidate solution vectors $C = \{\mathbf{c}_a\}_{a=1}^{A}$ which are contained within a subset of the overall ML solution space, where A is the total number of ants. This list is then divided into two sublists C_{+1}^k and C_{-1}^k based on the value of the k-th element in $\mathbf{c}_a$ (where C_b^k is the list containing all the vectors in which the k^{th} element is equal to b). Now the sets χ_{-1}^k and χ_{+1}^k are replaced with the lists C_{-1}^k and C_{+1}^k respectively in (4.5) to generate soft information to be used in the next iteration. The iterative procedure continues until the termination criteria has reached.

A block diagram of the proposed ACO assisted receiver for BPSK signalling[7] is shown in Fig. 4.1. An MMSE based preprocessor produces an initial soft estimate and the corresponding hard decision. The hard decision is used to create a moving-pattern table while the soft estimate is employed in the calculation of the heuristic value. For the purpose of updating the soft signal estimates, the ACO search process generates a list of likely candidate solutions. To build a candidate solution, every ant performs a circular tour through a moving-pattern table.

After the solution construction process is over, each ant deposits its solution to a candidate solution list which is then used to update the soft estimates of the transmitted signal. As a last step before going to the next iteration, the pheromone update is performed according to the elite ant strategy[8]. The developed receiver is capable of handling all type of system load factors, L_F. It is assumed that the channel estimator is producing a perfect estimate of the CSI i.e. $\hat{\mathbf{H}} = \mathbf{H}$. The internal specifications and working of the channel estimation process is outside the scope of this work.

[7] For BPSK modulation, the set $\chi = \{-1, +1\}$ and the modulation alphabet size $|\chi| = 2$.
[8] This is done in order to keep the search experience of the ants up to date.

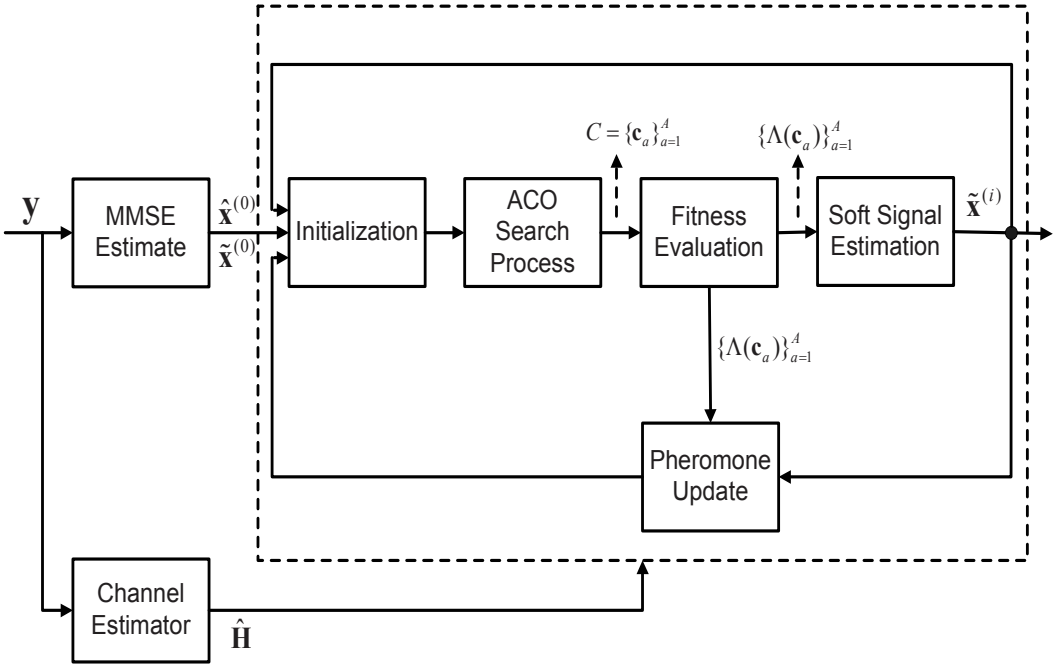

Figure 4.1 Structure of the soft ACO based receiver

Finally, a complete sequence of steps performed by the algorithm to compute the solution, $\hat{\mathbf{x}} = [\hat{x}_1 \dots \hat{x}_{N_T}]$, are shown in Fig. 4.2 while each individual phase is described next. Inputs to the detector are, the received signal $\mathbf{y}$, channel matrix $\mathbf{H}$, modulation alphabet size $|\chi|$, noise variance σ_ν^2, the number of ACO iterations I and number of ants A.

4.3.2 Preprocessing

During preprocessing, an initial soft estimate, $\tilde{\mathbf{x}}^{(0)} = [\tilde{x}_1 \tilde{x}_2 \dots \tilde{x}_{N_T}]$, and the corresponding hard decision, $\hat{\mathbf{x}}^{(0)} = [\hat{x}_1 \hat{x}_2 \dots \hat{x}_{N_T}]$, are obtained from the linear MMSE detector. To compute an estimate of the k-th transmitted signal, the received signal is passed through a linear MMSE filter represented by the matrix $\mathbf{W}$ as

$$\begin{aligned} \tilde{x}_k &= \mathbf{w}_k^H \mathbf{y} \\ &= \mathbf{w}_k^H \, \mathbf{h}_k x_k + \sum_{\substack{l=1 \\ l \neq k}}^{N_T} \mathbf{w}_k^H \, \mathbf{h}_l x_l + \mathbf{w}_k^H \boldsymbol{\nu} \\ &= \tilde{w}_k x_k + \tilde{\nu}_k, \end{aligned} \tag{4.7}$$

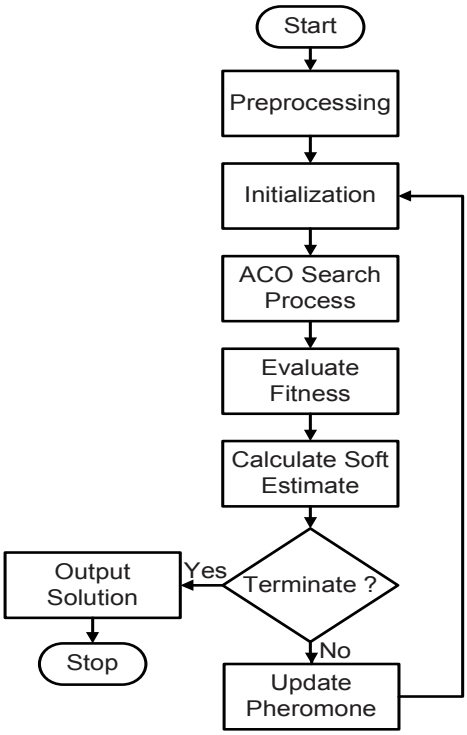

Figure 4.2 Flowchart of the proposed ACO based receiver for binary signalling.

where $\tilde{w}_k = \mathbf{w}_k^H \mathbf{h}_k$ is a scaling factor while $\tilde{\nu}_j = \sum_{\substack{l=1 \\ l \neq j}}^{N_T} \mathbf{w}_k^H \mathbf{h}_l x_l + \mathbf{w}_k^H \boldsymbol{\nu}$ represents the combination of CCI and the transformed noise. The variable $\tilde{\nu}_k$ is well approximated by a Gaussian distribution [93] with mean zero and variance $\sigma_{\tilde{\nu},k}^2 = \tilde{w}_k (1 - \tilde{w}_k)$. Hence, the PDF of $\tilde{x}_k$ can be expressed as [94]

$$\bar{p}(\tilde{x}_k | x_k, \tilde{w}_k) = \frac{1}{\pi \sigma_{\tilde{\nu},k}^2} \exp\left(- \frac{|\tilde{x}_k - \tilde{w}_k x_k|^2}{\sigma_{\tilde{\nu},k}^2} \right), \ \forall x_k \in \chi. \tag{4.8}$$

Finally, the initial soft estimate of the k-th transmitted signal is defined by the likelihood value [91]

$$\tilde{L}_k = \ln \frac{\bar{p}(\tilde{x}_k | x_k = +1, \tilde{w}_k)}{\bar{p}(\tilde{x}_k | x_k = -1, \tilde{w}_k)}, \tag{4.9}$$

This can be further simplified and expressed using (4.8) as

$$\begin{aligned}\tilde{L}_k &= \ln \frac{\exp\left(-\frac{1}{\sigma^2_{\tilde{\nu},k}} |\tilde{x}_k - \tilde{w}_k x_k)|^2_{x_k=+1}\right)}{\exp\left(-\frac{1}{\sigma^2_{\tilde{\nu},k}} |\tilde{x}_k - \tilde{w}_k x_k|^2_{x_k=-1}\right)} \\ &= \frac{1}{\sigma^2_{\tilde{\nu},k}} \left(|\tilde{x}_k - \tilde{w}_k x_k|^2_{x_k=-1} - |\tilde{x}_k - \tilde{w}_k x_k|^2_{x_k=+1}\right). \end{aligned} \tag{4.10}$$

The ants utilize the soft estimate in (4.10) to calculate the heuristic value corresponding to each transmitted signal. From (4.10), a hard output can also be obtained as

$$\hat{x}_k = Q_\chi\left(\tilde{L}_k\right), \qquad j = 1, 2 \ldots N_T \tag{4.11}$$

where $Q_\chi(.)$ is defined as

$$Q_\chi\left(\tilde{L}_k\right) = \arg\min_{x \in \chi} \left|\tilde{L}_k - x\right|^2. \tag{4.12}$$

The detected signal $\hat{x}_k$ is used in creating the moving-pattern table during the initialization phase.

4.3.3 Iterative Solution Construction and Refinement

Initialization

As a first step towards the iterative solution construction and refinement, the following parameters are initialized by the detector to carry out the ACO search process [87].

1. A moving-pattern table, $\mathbf{\Phi}^{(0)}_{2\times N_T} = \{\phi_{ij}\}$, which consists of $\hat{\mathbf{x}}^{(0)}$ as its first row and its complement as the second row[9]. An ant starts walking within this table at a randomly chosen column index and selects a single element from that column (i.e. either from first row or second row of $\mathbf{\Phi}$) based on the associated amount of pheromone and the heuristic value. The ants have to take a cyclic tour within

[9]For BPSK, the complement can be obtained by just flipping the bits at the corresponding locations in $\hat{\mathbf{x}}^{(0)}$.

the moving-pattern table and select one element from each column in order to construct a solution.

2. A pheromone table, $\boldsymbol{\tau}_{2\times N_T} = \{\tau_{ij}\} = \mathbf{1}$ (where $\mathbf{1}$ is an all ones matrix of the same dimension as $\boldsymbol{\tau}$), to which each ant contributes the experience about its search process in the form of depositing and evaporating the pheromone. Furthermore, it is periodically reinitialized for the purposes of a thorough exploration of the search space and to avoid an early convergence to a local optimum solution.

3. An ant trail table, $\mathbf{T}^{(a)}_{2\times N_T} = \left\{T^a_{ij}\right\}$ for $a = 1, \ldots, A$, $i = 1, 2$ and $j = 1, \ldots, N_T$. It keeps record of the visited locations in the moving-pattern table, $\boldsymbol{\Phi}^{(0)}_{2\times N_T}$. This information is used by the ants to deposit the pheromone in the pheromone table $\boldsymbol{\tau}$ during the pheromone update stage.

A graph based illustration of the moving-pattern table is presented in the next section along with its link to both the pheromone table and the ant trail table.

ACO Search Process

The moving-pattern table can be modeled as a multigraph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of N_T vertices or nodes and $\mathcal{E}$ is a set of $(|\chi|\, N_T)$ edges. Each node corresponds to a component of the solution vector and the $|\chi| = 2$ edges between any two consecutive nodes provide links for ants to access them. Ants are allowed to walk on this graph from one node to another along the edges. Each of the edges has an associated heuristic value and pheromone strength using which the ant calculates a transition probability. According to this probability, an ant chooses a particular component of the solution vector from the signal set χ and moves on to the next node. In this way, ants follow an incremental solution construction strategy where a single component of the solution vector is selected at a time. To build a complete length N_T candidate solution vector, an ant takes N_T steps selecting one solution component during each step.

Entries of the pheromone table, $\boldsymbol{\tau}$, reflect the pheromone strength on the edges $\mathcal{E}$ of the graph $\mathcal{G}$. Also the ant trail table, $\mathbf{T}$, keeps track of the set of edges followed by each ant during the solution construction. This information is used by the algorithm to

update the pheromone on the edges $\mathcal{E}$ of the graph later during the pheromone update phase.

The ACO search process is carried out as outlined in [87] to generate the candidate solutions list $C = \{\mathbf{c}_a^{(i)}\}_{a=1}^{A}$ containing A solution vectors generated by the A ants. The ants walk through the moving-pattern table (or equivalently over the problem graph) and probabilistically select one of the two components from each of its column based on the associated pheromone strength and the heuristic value. The heuristic value for the k-th transmitted signal calculated by the a-th ant is given by

$$\eta_k^a = \frac{1}{2 + |\tilde{x}^{(i-1)}(k)| + \sum_{l \in V^a} |\tilde{x}^{(i-1)}(l)|}, \tag{4.13}$$

where V^a represents the short-term memory of the a-th ant. It is a set containing the indices of the locations where the ant has chosen an element of the solution vector from second row (or equivalently from the complement of the current best solution[10]) of the moving-pattern table, $\mathbf{\Phi}$. Note that the definition of heuristic value in (4.13) constrains the solution constructed by the a-th ant in the neighborhood of the current best solution.

Based on this heuristic value, the probability of selection for the k-th transmitted signal by the a-th ant is given as

$$p_k^a = \tau_{2k}\, \eta_k^a. \tag{4.14}$$

Here, τ is the pheromone table which is already defined and initialized in the initialization step. More precisely, (4.14) specifies the probability of selecting the complement of the hard decision from the previous iteration.

A partial graphical view of the search process is shown in Fig. 4.3 where the moving-pattern table is modeled as a multigraph with ants walking over it to build solutions. Following this interpretation, (4.14) represent the probability of transition of the a-th ant from $(k-1)$-th node to the k-th node as shown in Fig. 4.3. Clearly, the transition probability depends on both, the heuristic value and, the pheromone strength

[10]Current best solution, $\hat{\mathbf{x}}$, is the MMSE solution in the first iteration and the solution obtained from (4.22) during later iterations.

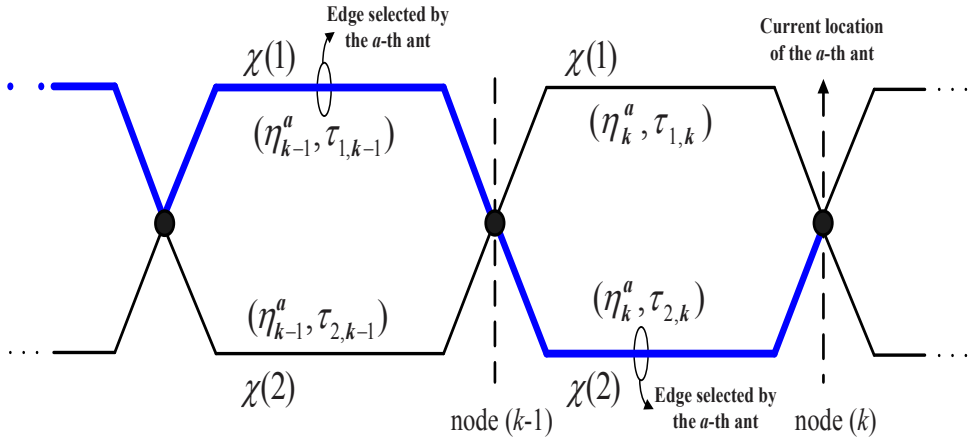

Figure 4.3 Problem graph and an illustration of solution construction by the ants.

on the edges joining $(k-1)$-th node to the k-th node. The heuristic value in (4.13) is calculated from the soft estimates which in turn are calculated based on the MAP rule (4.2)-(4.6). The pheromone strength, which represents the past search experience of the ant-colony, actually biases the soft values in the region where more promising solutions were discovered in the past.

Fitness Evaluation

The fitness value of all the solutions contained in list C is evaluated next using a fitness function. The fitness function is defined using the ML Euclidean distance metric (5.5)

$$\Lambda(\mathbf{u}) = \|\mathbf{y} - \mathbf{H}\mathbf{u}\|^2, \tag{4.15}$$

where $\mathbf{u}$ is the candidate solution vector in consideration. It is important to note that the fitness function produces a scalar value which acts as a quality indicator of the corresponding solution. The fitness values of the generated solutions are therefore calculated as

$$\Lambda(\mathbf{c}_a) = \|\mathbf{y} - \mathbf{H}\mathbf{c}_a\|^2, \qquad a = 1, 2 \ldots A \tag{4.16}$$

where $\mathbf{c}_a \in C$ and $C = \{\mathbf{c}_1, \mathbf{c}_2 \ldots \mathbf{c}_A\}$.

Soft Signal Estimation

The soft estimate of the k-th transmitted symbol $\tilde{x}_k \in \tilde{\mathbf{x}}$ is updated using the candidate solutions list $C = \{\mathbf{c}_a\}_{a=1}^{A}$ and the max-log MAP approximation in (4.6) as

$$\tilde{x}_k \approx \frac{1}{2\sigma_\nu^2}[\|\mathbf{y} - \mathbf{H}\mathbf{c}_{-1}\|^2 - \|\mathbf{y} - \mathbf{H}\mathbf{c}_{+1}\|^2], \tag{4.17}$$

where $\mathbf{c}_b = \arg\min_{\mathbf{c}\in\, C_b^k} \|\mathbf{y} - \mathbf{H}\mathbf{c}\|^2$ and C_b^k is a sublist of vectors from the list $C = \{\mathbf{c}_a\}_{a=1}^{A}$ (i.e. $C_b^k \subseteq C$) defined as

$$C_b^k = \left\{\mathbf{c}_a = (c_{a1} \dots\ c_{ak} \dots\ c_{aN_T})^T \mid \mathbf{c}_a \in C,\ c_{ak} = b\right\},\ a = 1 \dots A \tag{4.18}$$

where $b = -1$ or $+1$.

Pheromone Update

In pheromone update step, the pheromone table $\boldsymbol{\tau}$ is modified to reflect up to date search experience of the ants. Pheromone evaporation is performed first and is followed by depositing new pheromone.

- During pheromone evaporation, a fraction of pheromone strength associated with all the possible edges is removed first as

$$\boldsymbol{\tau} = (1 - \rho)\ \boldsymbol{\tau}, \tag{4.19}$$

  where $\rho = \frac{1}{2}$ for the AS algorithm[11].

- The pheromone is deposited based on the *elite ant strategy* [95]. In elite ant strategy, a set of ants, termed *elite ants*, is formed based on the fitness value of the solutions generated by the ants. The selected ants are then only allowed to participate in the pheromone update activity. We define the elite ants to be those which have produced a better quality solution than the best solution in terms of the fitness value. The fitness value of the best solution, $\hat{\mathbf{x}}^{(i-1)} = Q_\chi\left(\tilde{\mathbf{x}}^{(i-1)}\right)$ (here

[11] More details on the AS algorithm can be found in Appendix B.

$Q_\chi(.)$ again denote the constellation quantizer), is denoted $\Lambda' = \Lambda(\hat{\mathbf{x}}^{(i-1)})$ and is obtained via (4.15). Using Λ', the set of elite ants is formed as

$$C_E = \left\{ \mathbf{c}_a = (c_{a1} \dots \; c_{ak} \dots \; c_{aN_T})^T \, | \mathbf{c}_a \in C, \; \Lambda' \geq \Lambda(\mathbf{c}_a) \right\}, \; a = 1 \dots A. \quad (4.20)$$

The pheromone is then increased on the edges $\mathcal{E}$ of the graph $\mathcal{G}$ belonging to only the elite ants. To do so, the algorithm employs the ant trail table $\mathbf{T}$ which contains the information about the locations visited by the ants within the moving-pattern table (or equivalently the complete path followed by the ant on the graph to build its solution). As such, $\mathbf{T}$ represents the memory of the ants. The pheromone is deposited as follows

$$\tau\left(T^a_{1k}, T^a_{2k}\right) = \tau\left(T^a_{1k}, T^a_{2k}\right) + \Delta\tau, \qquad \forall \, \mathbf{c}_a \in C_E \quad (4.21)$$

where $\Delta\tau = \frac{N_T}{\Lambda(\mathbf{c}_a)}$. Also, T^a_{1k} is the row index whereas T^a_{2k} is the column index of the k-th selected signal by the a-th ant inside the moving-pattern table.

An important point here is to note that the convergence process ceases after a certain number of iterations initially. This is due to the reason that all the ants start producing similar solutions giving rise to a stagnation situation [86]. As a solution to this problem, it is discovered that by periodically reinitializing the pheromone table, $\boldsymbol{\tau}$, the premature convergence of the algorithm to a suboptimum solution can be avoided.

Output Solution

After the completion of I iterations, a hard decision $\hat{\mathbf{x}}$ is obtained from $\tilde{\mathbf{x}}$ as

$$\hat{x}_j = Q_\chi\left(\tilde{x}_j\right), \qquad j = 1, 2 \dots N_T \quad (4.22)$$

where $Q_\chi(.)$ is defined as

$$Q_\chi\left(\tilde{x}_j\right) = \arg\min_{x \in \chi} \left|\tilde{x}_j - x\right|^2. \quad (4.23)$$

4.3.4 Computational Complexity

The complexity of the proposed receiver depends on the number of fitness evaluations performed during the ACO based iterative signal estimation process. The approximate complexity in terms of ML metric evaluations, (5.5) or equivalently (4.15), is given as AI metric calculations[12]. This is much less than the complexity of ML detection which requires 2^{N_T} metric evaluations using (5.5). As an example, with $N_T = 8$ transmitted signals, $2^{16} = 65536$ metric evaluations are required for the exhaustive ML search. On the other hand, it takes $AI = (500)(28) = 14000$ metric evaluations by ACO to reach nearly the same performance in the critically loaded systems.

The complexity of the developed receiver in terms of the number of real multiplications and additions can be expressed respectively as

$$\mathcal{R}^{\times}_{ACO-I} = \mathcal{R}^{\times}_{PP} + \mathcal{R}^{\times}_{IP} \tag{4.24}$$

and

$$\mathcal{R}^{+}_{ACO-I} = \mathcal{R}^{+}_{PP} + \mathcal{R}^{+}_{IP}, \tag{4.25}$$

where $\mathcal{R}^{\times}_{PP} = \mathcal{R}^{\times}_{LD}$ and $\mathcal{R}^{+}_{PP} = \mathcal{R}^{+}_{LD}$ are the complexities of the preprocessing stage in terms of the number of real multiplications and additions which are already obtained in (3.18) and (3.19), respectively[13]. The terms $\mathcal{R}^{\times}_{IP}$ and $\mathcal{R}^{+}_{IP}$ represent the complexity (in terms of real multiplications and additions) of the ACO assisted iterative soft processing part of the algorithm.

As explained, the total number fitness evaluations performed by the ACO assisted receiver is AI. With BPSK modulation in use, the fitness value calculation of a single candidate solution requires $2N_RN_T + 2N_R$ real multiplications and $2N_RN_T + N_R - 1$ real additions. Therefore, the complexity terms $\mathcal{R}^{\times}_{IP}$ and $\mathcal{R}^{+}_{IP}$ can be calculated as

$$\mathcal{R}^{\times}_{IP} = (2N_RN_T + 2N_R)\, AI \tag{4.26}$$

[12]The computational costs associated with ACO operations like calculation of heuristic values, transition probability and the pheromone update are significantly smaller than those for the metric calculations.

[13]This is due to the fact that the preprocessing is based on the linear MMSE detector which is discussed in Section 3.3.

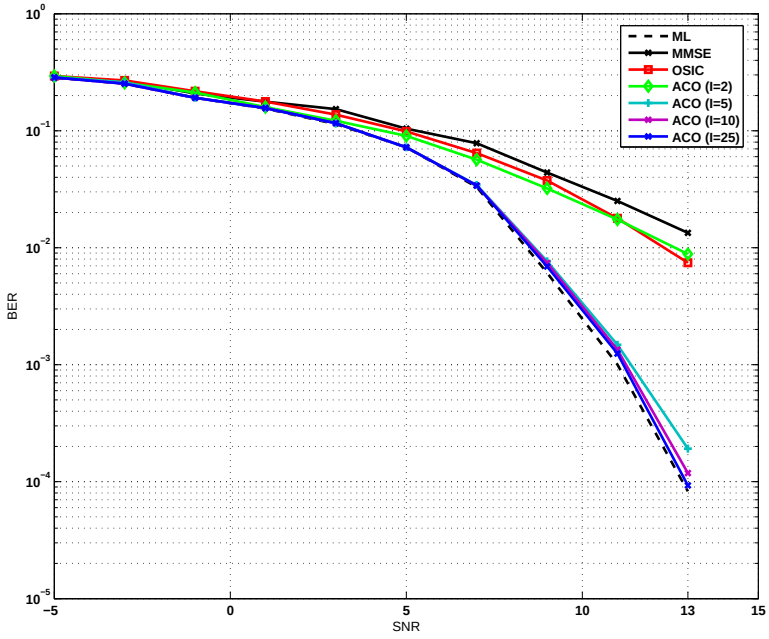

Figure 4.4 BER performance of a critically loaded system with $N_T = N_R = 16$ and $A = 500$.

and

$$\mathcal{R}_{IP}^{+} = (2N_R N_T + N_R - 1)\, AI, \tag{4.27}$$

Substituting these values of $\mathcal{R}_{IP}^{\times}$ and $\mathcal{R}_{IP}^{+}$ in (4.24) and (4.25) produces the overall complexity of the ACO based detector.

4.3.5 Simulation Results

In this section, we present the simulation results for signal detection in a frequency-flat fading channel. We consider first a critically loaded cochannel system, $L_F = 1$, with $N_T = 16$ transmitted signals and $N_R = 16$ receive diversity branches. We assume perfect CSI is available at the receiver. A table of parameters for our soft ACO based detection is presented in Table 4.1.

Fig. 4.4 depicts the BER performance of the ACO-based soft detector where we have used $A = 500$ ants to search for the optimum solution over the entire solution space χ^{N_T}. Different curves in this figure represent the performance of different detection schemes

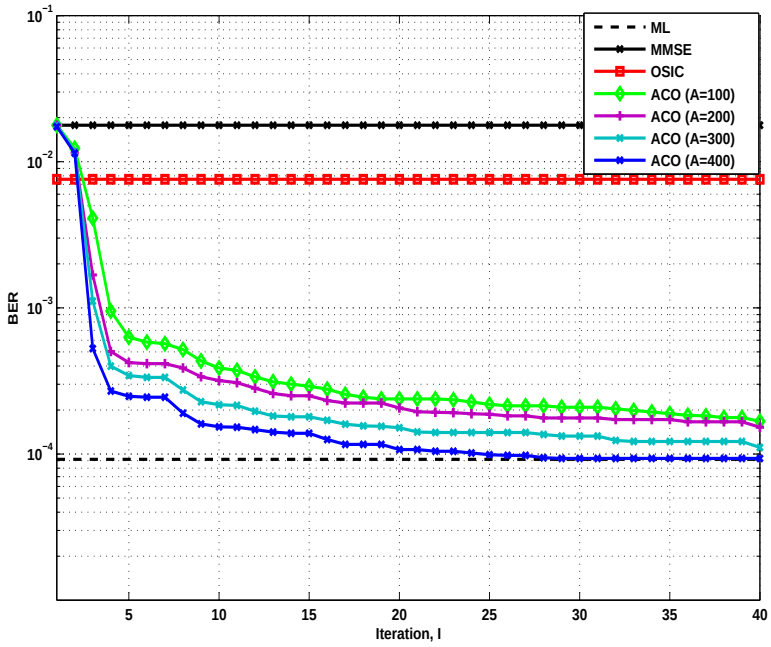

Figure 4.5 BER versus I, at 13dB for a critically loaded system with $N_T = N_R = 16$ and $A = 100 - 400$.

Table 4.1 ACO Detection Parameters (Modulation is BPSK for all the cases)

Figure	N_T	N_R	A	I
Fig. 4.4	16	16	500	2, 5, 10, 25
Fig. 4.5	16	16	100-400	1-40
Fig. 4.6	16	12	500	2, 5, 10, 40
Fig. 4.7	16	12	100-500	1-40
Fig. 4.8	16	10	500	2, 5, 10, 40

(ML, MMSE, ordered SIC (OSIC), and ACO) with different numbers of iterations ($I = 1, 2, 3, 4, 5, 10, 25$).

For the same system, we present the BER performance at 13 dB with varying number of iterations in Fig.4.5. It can be seen, the performance of the proposed ACO based detection converges to that of ML detection within a finite number of iterations ($I = 28$ in the case of $A = 400$). Also in Fig. 4.5, it is important to note that the algorithm attains a stagnation situation [86] after a certain number of iterations (for example, after 7 iterations initially). Due to this, all the ants start producing the same solutions resulting in no further performance improvement. As a solution, it is found that by

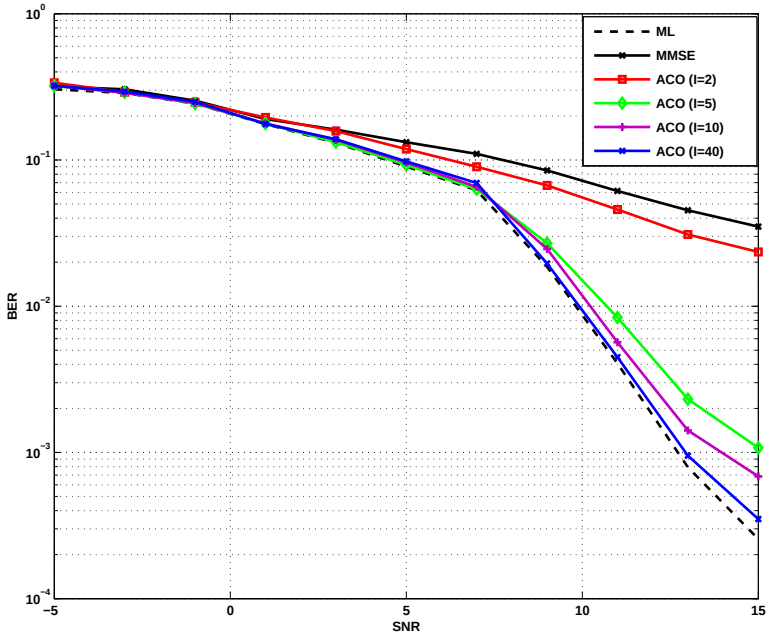

Figure 4.6 BER performance of an overloaded system with $N_T = 16$, $N_R = 12$ and $A = 500$.

periodically reinitializing the pheromone table, $\boldsymbol{\tau}$, the stagnation situation can be exited. In this case, the pheromone table is reinitialized to, $\boldsymbol{\tau} = \mathbf{1}$ (where $\mathbf{1}$ is an all ones matrix of the same dimension as $\boldsymbol{\tau}$) after 7 iterations initially and then after every 3 iterations.

Similarly, BER performance comparison of different detection schemes in overloaded conditions is shown in Fig. 4.6. Overloaded conditions are challenging as $N_T > N_R$, which results in an underdetermined system. Here, we consider $N_T = 16$ and $N_R = 12$ which implies an overload factor of $L_F = 1.33$. Again, it is evident that the performance of the ACO based scheme approaches that of ML detection with an increasing number of ACO iterations. In Fig. 4.7, it can again be seen that at an SNR of 13 dB, the overloaded soft ACO detector is converging to the performance of ML detection using $A = 500$ ants. Again, the stagnation situation is clearly evident here which periodically limits the algorithm convergence. As discussed earlier, the pheromone is reinitialized to recommence the convergence process and to improve the system performance.

Finally, in Fig. 4.8, the BER performance comparison of ML, MMSE and ACO for

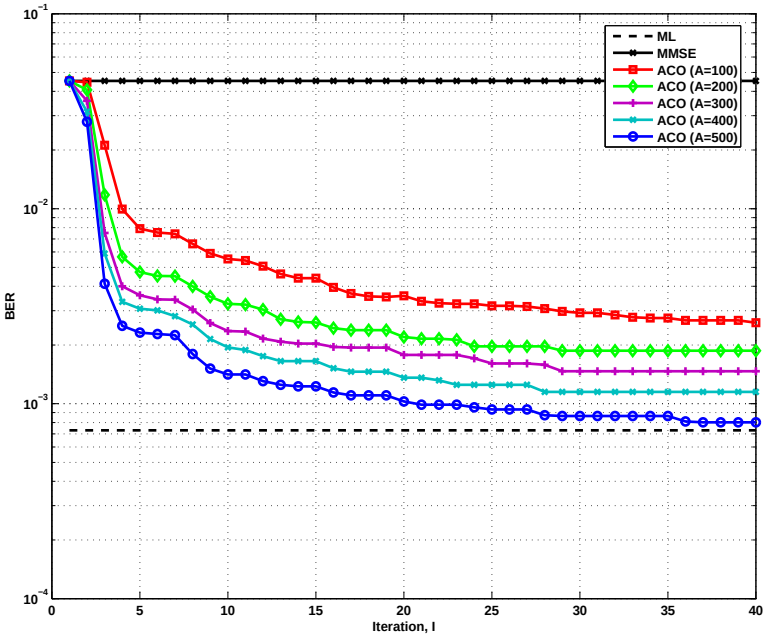

Figure 4.7 BER versus I at 13dB for an overloaded system with $N_T = 16, N_R = 12$ and $A = 100 - 500$.

$N_T = 16$ and $N_R = 10$ is shown. Note that the system loading factor in this case increases to $L_F = 1.60$ which implies an overload of 60%. We observe an error floor below a BER of 10^{-3} for $A = 500$ and $I = 40$. Error control coding could be used to further improve performance in this case.

An estimate of complexity figures associated with each of the above scenario is presented in Table 4.2. The number of transmitted signals $N_T = 16$ and the modulation is BPSK in all these cases. For reference purposes, a percentage complexity comparison with ML detector is also shown. For the system with $N_R = 16$, it can be seen that the ACO detector achieves ML performance at the expense of a computational cost which is only 19% of the ML detector. This implies 81% complexity savings in comparison to the ML detector. Similarly, the ACO detector offers 69.48% savings in computational costs when $N_R = 12$ and $N_R = 10$.

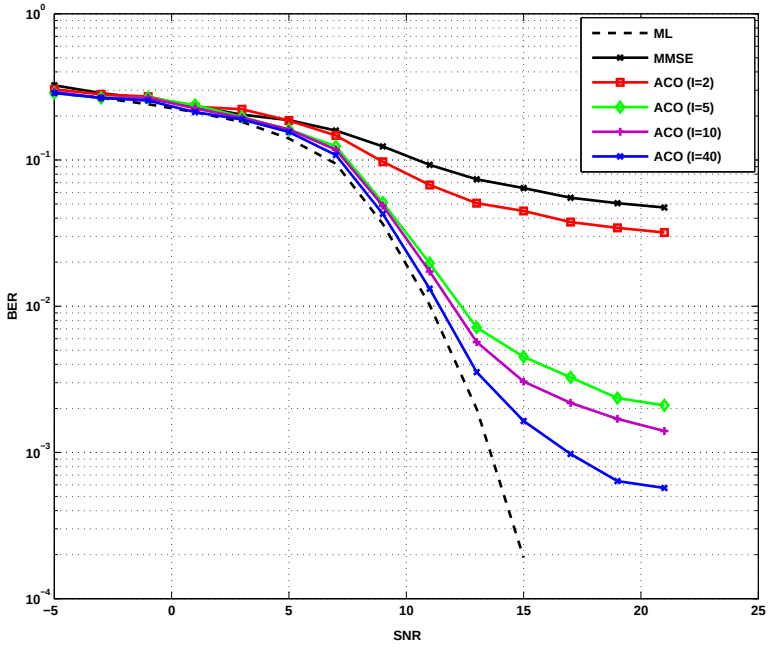

Figure 4.8 BER performance of an overloaded system with $N_T = 16$, $N_R = 10$ and $A = 500$.

Table 4.2 Number of Real Multiplications ($N_T = 16$ in all cases)

	ACO Detector			ML Detector	% ML
N_R	A	I	$\mathcal{R}^{\times}$	$\mathcal{R}^{\times}$	
16	500	2	$5.44E5$	3.57E7	1.53%
16	500	5	$1.36E6$	3.57E7	3.81%
16	500	10	$2.72E6$	3.57E7	7.63%
16	500	25	$6.80E6$	3.57E7	19.07%
12	500	2	$4.08E5$	2.67E7	1.53%
12	500	5	$1.02E6$	2.67E7	3.81%
12	500	10	$2.04E6$	2.67E7	7.63%
12	500	40	$8.16E6$	2.67E7	30.52%
10	500	2	$3.40E5$	2.23E7	1.53%
10	500	5	$8.50E5$	2.23E7	3.81%
10	500	10	$1.70E6$	2.23E7	7.63%
10	500	40	$6.80E6$	2.23E7	30.52%

4.4 ACO ASSISTED DETECTOR FOR NON-BINARY SIGNALS

In the last section, we developed ACO assisted detector for the cochannel systems employing binary signals for transmission. For better utilization of system bandwidth

and to achieve higher data rates, a higher order signalling scheme is used in most of the practical communication systems. Therefore, we develop ACO assisted detector for non-binary modulation in this section. The proposed algorithm relies on the MMSE detector for the calculation of the heuristic value whereas MMAS algorithm[14] is used to define the solution construction mechanism.

Although the ACO based algorithm presented in this section can only handle a lightly overloaded system offering a degraded performance but the results provide motivation to investigate novel techniques in this direction to discover new algorithms having potential of achieving near-ML performance. One such technique is presented in the next chapter.

4.4.1 The Heuristic Framework

In MUD, the success of ACO based algorithms depends on the way the problem-specific heuristic information is defined and utilized. A well-defined heuristic criteria and the corresponding heuristic value results in superior performance and convergence of the algorithm. In this section, we develop ACO assisted detection algorithms which rely upon the linear MMSE based detector for defining and processing the heuristic values.

Recalling from Section 3.3, the MMSE estimate of the i-th transmitted signal is formed by passing the received signal through a linear filter as follows

$$\begin{aligned}\tilde{x}_i = \mathbf{w}_i^H \mathbf{y} &= \mathbf{w}_i^H \, \mathbf{h}_i x_i + \sum_{\substack{l=1 \\ l \neq i}}^{N_T} \mathbf{w}_i^H \, \mathbf{h}_l x_l + \mathbf{w}_k^H \boldsymbol{\nu} \\ &= \tilde{w}_i x_i + \tilde{\nu}_i, \end{aligned} \tag{4.28}$$

where $\tilde{w}_i = \mathbf{w}_i^H \, \mathbf{h}_i$ and the quantity $\tilde{\nu}_i = \sum_{\substack{l=1 \\ l \neq j}}^{N_T} \mathbf{w}_i^H \, \mathbf{h}_l x_l + \mathbf{w}_i^H \boldsymbol{\nu}$ can be approximated by a Gaussian distribution with mean zero and variance $\sigma_{\tilde{\nu},i}^2 = \tilde{w}_i \left(1 - \tilde{w}_i\right)$ [93].

The PDF of $\tilde{x}_i$ can be expressed as [94]

$$\bar{p}\left[\tilde{x}_i | x_i = \chi(l)\right] = \frac{1}{\pi \sigma_{\tilde{\nu},i}^2} \exp\left(-\frac{\left|\tilde{x}_i - \tilde{w}_i x_i\right|^2}{\sigma_{\tilde{\nu},i}^2}\right), \quad l = 1, 2 \ldots |\chi| \,. \tag{4.29}$$

[14] Interested reader is referred to Appendix B for details on MMAS ACO algorithm.

For QPSK modulated signals, $\chi = \{+1+j, +1-j, -1+j, -1-j\}$, the evaluation of (4.29) produces a value of $\bar{p}\left[\tilde{x}_i | x_i = \chi(l)\right]$ corresponding to each of the possible symbols in χ. Therefore, the PDF in (4.29) is used by the proposed ACO assisted detection algorithm to define the heuristic value.

4.4.2 Solution Construction using ACO

A total of A ants are assigned the task of traversing the problem graph and building A candidate solution vectors. The ants consider the solution construction process as a random walk over a multigraph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of N_T vertices or nodes and $\mathcal{E}$ is a set of $(|\chi|\, N_T)$ edges or links. Each node represents a component of the solution vector and the $|\chi|$ edges between two consecutive nodes provide links for ants to travel across the graph. Each of the ants is assigned a starting node position at the beginning of the algorithm[15]. It then performs a circular tour on the graph moving from one node to the other to form a length-N_T solution vector.

The heuristic value associated with the l-th edge between node $(i-1)$ and node i (which is a potential candidate for the i-th component of the solution vector) is calculated using (4.29) as

$$\eta_{il} = \frac{1}{\xi_{il}}, \quad l = 1, 2 \ldots |\chi|\,, \tag{4.30}$$

where $\xi_{il} = \bar{p}\left[\tilde{x}_i | x_i = \chi(l)\right]$ is already available from (4.29).

Ants move from node $(i-1)$ to node i, according to the transition probability

$$p_{il} = \frac{(\tau_{il})^{\alpha} \cdot (\eta_{il})^{\beta}}{\sum_{k=1}^{|\chi|} (\tau_{ik})^{\alpha} \cdot (\eta_{ik})^{\beta}}, \quad l = 1, 2 \ldots |\chi|\,, \tag{4.31}$$

where τ_{il} and η_{ik} are the pheromone strength and the heuristic value associated with the edges involved.

Using this incremental solution construction strategy, each ant builds a complete length-N_T candidate solution vector in N_T moves where a single component of the solution vector is selected in one such move. The output of the search process is a list $\bar{\mathcal{L}} =$

[15]This is usually done by generating a length-A vector with elements randomly (following a uniform distribution) picked from the set of all the possible indices $\{1, 2 \ldots N_T\}$.

$\{\mathbf{x}^a\}_{a=1}^{A}$ containing $\left|\bar{\mathcal{L}}\right| = A$ candidate solution vectors. The next stage is the evaluation of fitness of each of the generated solutions in the list $\bar{\mathcal{L}}$. The fitness value produced by the fitness function acts as a quality indicator of the hypothesized solution. For this purpose, the ML Euclidean distance metric in (5.5) is defined as the fitness function and the fitness value is computed accordingly as

$$\Lambda_a = \left\|\mathbf{y} - \mathbf{H}\mathbf{x}_a\right\|^2, \quad a = 1, 2 \ldots A. \tag{4.32}$$

The vector with the minimum fitness value is designated as the iteration's best solution, $\mathbf{x}^\dagger$, and is calculated as

$$\mathbf{x}^\dagger = \arg\min_{\mathbf{x}\in\bar{\mathcal{L}}} \left\|\mathbf{y} - \mathbf{H}\mathbf{x}\right\|^2. \tag{4.33}$$

Likewise, the global best solution, $\mathbf{x}^\ddagger$, is defined as the best solution in terms of fitness value since the start of the algorithm. Now, if the algorithm is performing its first search iteration then, $\mathbf{x}^\ddagger = \mathbf{x}^\dagger$, otherwise it is calculated as

$$\mathbf{x}^\ddagger = \begin{cases} \mathbf{x}^\ddagger, & \text{if } \Lambda^\ddagger <= \Lambda^\dagger \\ \mathbf{x}^\dagger, & \text{if } \Lambda^\ddagger > \Lambda^\dagger. \end{cases} \tag{4.34}$$

Both the iteration best and the global best solutions are used in the pheromone update process.

4.4.3 Pheromone Update

Before moving on to the next ACO iteration / stage, pheromone update is performed on the edges of the problem graph to reflect the up to date search experience of the ants. Using MMAS algorithm, the pheromone is updated in two steps: pheromone evaporation and pheromone deposit. First, the pheromone evaporation is performed as [86]

$$\tau_{mn} = (1 - \rho)\, \tau_{mn}, \ \forall\, (m, n), \tag{4.35}$$

where ρ is the pheromone evaporation rate. It is assigned a small value to ensure a thorough exploration of the search space. We use $\rho = 0.02$ as suggested in [86]. Second,

the pheromone is deposited on the edges corresponding to iteration best solution, $\mathbf{x}^{\dagger}$, in proportion to its fitness value, $\Lambda^{\dagger} = \left\|\mathbf{y} - \mathbf{H}\mathbf{x}^{\dagger}\right\|^2$, as follows [86]

$$\tau_{mn} = \tau_{mn} + \Delta\tau_{mn}^{\dagger}, \tag{4.36}$$

where

$$\Delta\tau_{mn}^{\dagger} = \frac{1}{\Lambda^{\dagger}} \tag{4.37}$$

To avoid stagnation situation, the pheromone strength on each of the edges is allowed to assume values only inside a particular interval $[\tau_{min}, \tau_{max}]$. The maximum possible pheromone value τ_{max} is defined using ρ and (4.33) as [86]

$$\tau_{max} = \frac{1}{\rho \cdot \Lambda^{\dagger}}. \tag{4.38}$$

Note that the initial values of the pheromone on all the edges are set equal to τ_{max}, i.e. $\boldsymbol{\tau}_0 = \tau_{max} \cdot \mathbf{1}$, where $\mathbf{1}$ is an all ones matrix with the same dimensions as $\boldsymbol{\tau}$. Using the maximum pheromone value τ_{max} stated in (4.38) and an ACO parameter a_0 [86], the minimum pheromone value τ_{min} can be computed as

$$\tau_{min} = \frac{\tau_{max}}{a_0}. \tag{4.39}$$

The parameter a_0 is defined as

$$a_0 = \frac{(N_T - 1)\ \sqrt[N_T]{p_{best}}}{1 - \sqrt[N_T]{p_{best}}}, \tag{4.40}$$

where p_{best} is the probability of generating a best solution when the solution has converged. A recommended value for p_{best} is 0.05 [96].

The solution construction process and the pheromone update process are repeatedly performed by the algorithm for a total of I ACO iterations. At the end, the global best solution, $\mathbf{x}^{\ddagger}$, is declared as the detected signal vector by the ACO based detector.

4.4.4 Computational Complexity

From the point of view of complexity, the detection process can be divided into two main parts namely, the computation of individual symbol probabilities from (4.29) and the evaluation of fitness values in (4.32). Since, the symbol probabilities are based on the MMSE estimate so the associated complexity can be stated in terms of the complexity of MMSE detector. Keeping this in view, the complexity of the ACO assisted detector in terms of the number of real multiplications and additions can respectively be expressed as

$$\mathcal{R}_{ACO}^{\times} = \mathcal{R}_{LD}^{\times} + \mathcal{R}_{FE}^{\times} \tag{4.41}$$

and

$$\mathcal{R}_{ACO}^{+} = \mathcal{R}_{LD}^{+} + \mathcal{R}_{FE}^{+}, \tag{4.42}$$

where $\mathcal{R}_{LD}^{\times}$ and $\mathcal{R}_{LD}^{+}$ are the complexities of the linear receiver in terms of the number of real multiplications and additions, respectively. These are already obtained in Section 3.3 as

$$\mathcal{R}_{LD}^{\times} = 4N_T^3 + 8N_T^2 N_R + 4N_R N_T \tag{4.43}$$

and

$$\mathcal{R}_{LD}^{+} = 4N_T^3 + N_T^2 \left(8N_R - 2\right) + 2N_R N_T - N_T. \tag{4.44}$$

The terms $\mathcal{R}_{FE}^{\times}$ and $\mathcal{R}_{FE}^{+}$ represent the complexity of the fitness evaluation stage in terms of real multiplications and additions. During the fitness evaluation step, the ML Euclidean distance is evaluated for a total of A candidate solutions in every ACO iteration. One such evaluation of the fitness function results in $(4N_R N_T + 4N_R)$ real multiplications and $(4N_R N_T + 4N_R - 2)$ real additions. Therefore, the total complexity of performing AI fitness evaluations results in the complexity

$$\mathcal{R}_{FE}^{\times} = \left(4N_R N_T + 4N_R\right) AI \tag{4.45}$$

and

$$\mathcal{R}_{FE}^{+} = \left(4N_R N_T + 4N_R - 2\right) AI. \tag{4.46}$$

Substituting (4.43)-(4.46) back in (4.41) and (4.42) yield the overall complexity of the ACO assisted detector.

4.4.5 Simulation Results

We now discuss the performance of the proposed ACO assisted detector with the help of computer simulations. We consider a frequency-flat fading channel and perfect CSI available at the receiver. All the transmitted signals are assumed to be quaternary phase-shift keying (QPSK) modulated. Since the internal specifications of ACO assisted detector rely on hard decision statistics, therefore most of the performance gains are obtained within one ACO iteration. Hence, $I = 1$ is assumed for all the scenarios. Table 4.3 lists the parameter settings for the simulation results in this section.

Table 4.3 Detector parameters used in simulations

Figure	N_T	N_R	A	I	Modulation
Fig. 4.9	8	8	1000, 2000, 5000, 7500, 10000, 12500	1	QPSK
Fig. 4.10	8	8	1000, 2000, 5000, 7500, 10000, 12500	1	QPSK
Fig. 4.11	8	6	1000, 2000, 5000, 7500, 10000, 12500	1	QPSK
Fig. 4.12	8	6	1000, 2000, 5000, 7500, 10000, 12500	1	QPSK
Fig. 4.13	8	5	1000, 2000, 5000, 7500, 10000, 12500, 15000	1	QPSK
Fig. 4.14	8	5	1000, 2000, 5000, 7500, 10000, 12500, 15000	1	QPSK

The BER performance of a critically loaded system with $N_T = N_R = 8$ having $L_F = 1.00$ is shown in Fig. 4.9. All the transmitted signals are QPSK modulated. Clearly, the ACO detector offers near-ML performance in low-to-medium SNR region. The performance, however, starts deteriorating at high SNR values. The performance of ACO detector against varying number of ants at an SNR of $\gamma = 10$ dB is shown in Fig. 4.10. Here, MMSE and ML detector performances are also shown for the purpose of comparisons. In general, the performance of ACO assisted detector improves as the number of ants are increased, but this comes at the cost of a higher computational requirements as indicated by Table 4.4.

Fig. 4.11 shows the performance of a cochannel system with $N_T = 8$ transmitted signals and a receiver employing $N_R = 6$ antennas. This corresponds to a load factor of $L_F =$

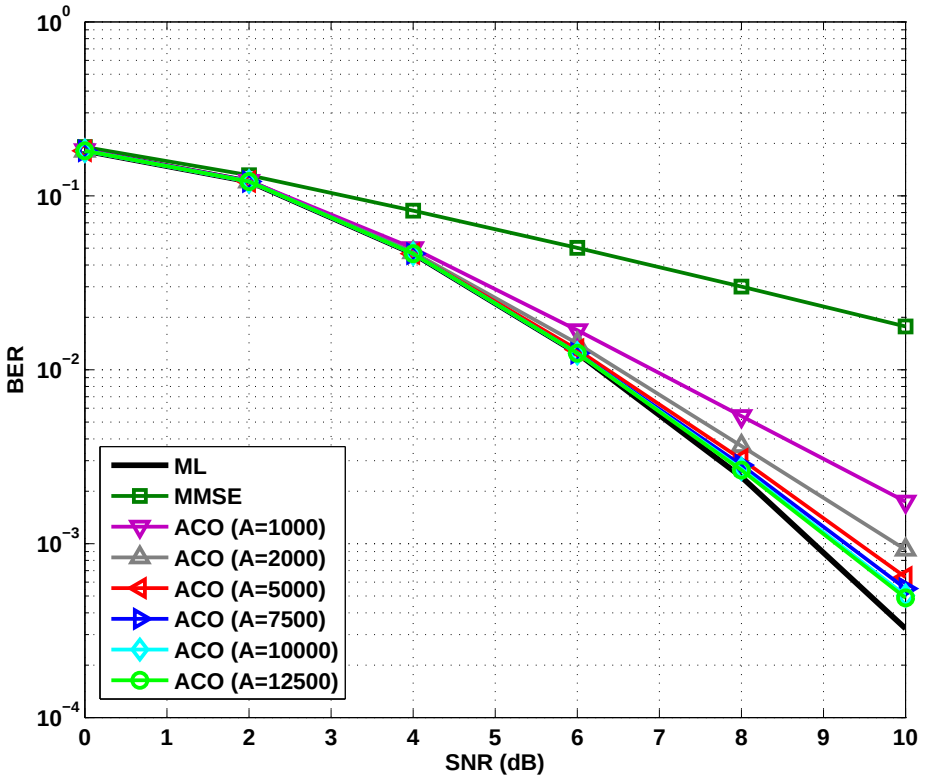

Figure 4.9 BER performance for a critically loaded system with $N_T = N_R = 8$.

Table 4.4 Complexity Comparison ($N_T = 8$, $N_R = 8$)

No. of Real Multiplications				% ML
ML	MMSE	ACO		
$1.89E7$	$6.40E3$	$A = 1000$	$2.88E5$	1.53%
$1.89E7$	$6.40E3$	$A = 2000$	$5.76E5$	3.05%
$1.89E7$	$6.40E3$	$A = 5000$	$1.44E6$	7.63%
$1.89E7$	$6.40E3$	$A = 7500$	$2.16E6$	11.44%
$1.89E7$	$6.40E3$	$A = 10000$	$2.88E6$	15.26%
$1.89E7$	$6.40E3$	$A = 12500$	$3.60E6$	19.07%

1.33. The results show that the ACO detector with 1000 ants is capable of achieving a superior BER performance than MMSE detector. The performance further improves if more ants are used. However, the performance gap between ML and ACO detectors increases in high SNR regime. In Fig. 4.12, performance of ACO based detector is plotted with a varying number of ants, A, at an SNR of $\gamma = 12$ dB. The performance

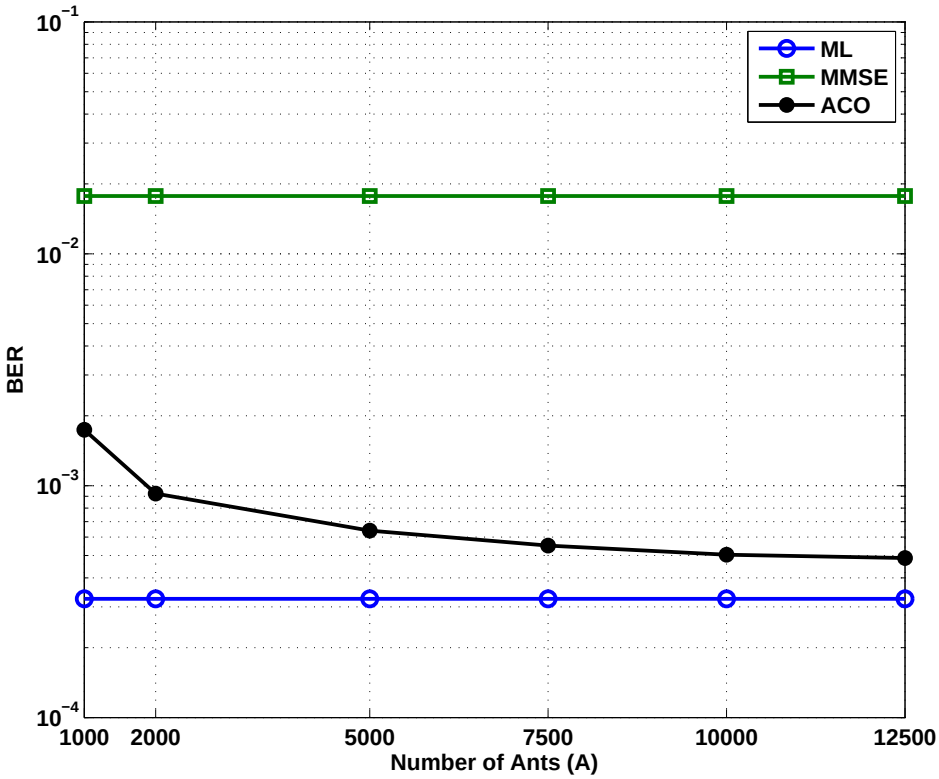

Figure 4.10 BER performance of $N_T = N_R = 8$ cochannel system at $\gamma = 10$ dB.

of MMSE and ACO detectors are also plotted for comparison purposes. It is evident that the performance improves as A increases, but at the cost of more computational expenses. This can be seen in Table 4.5 which lists the complexity associated with the algorithm. A percentage complexity comparison between the proposed ACO assisted detector and that of the ML detector is also provided.

Table 4.5 Complexity Comparison ($N_T = 8$, $N_R = 6$)

No. of Real Multiplications				% ML
ML	MMSE	ACO		
$1.42E7$	$5.31E3$	$A = 1000$	$2.16E5$	1.53%
$1.42E7$	$5.31E3$	$A = 2000$	$4.32E5$	3.05%
$1.42E7$	$5.31E3$	$A = 5000$	$1.08E6$	7.63%
$1.42E7$	$5.31E3$	$A = 7500$	$1.62E6$	11.44%
$1.42E7$	$5.31E3$	$A = 10000$	$2.16E6$	15.26%
$1.42E7$	$5.31E3$	$A = 12500$	$2.70E6$	19.07%

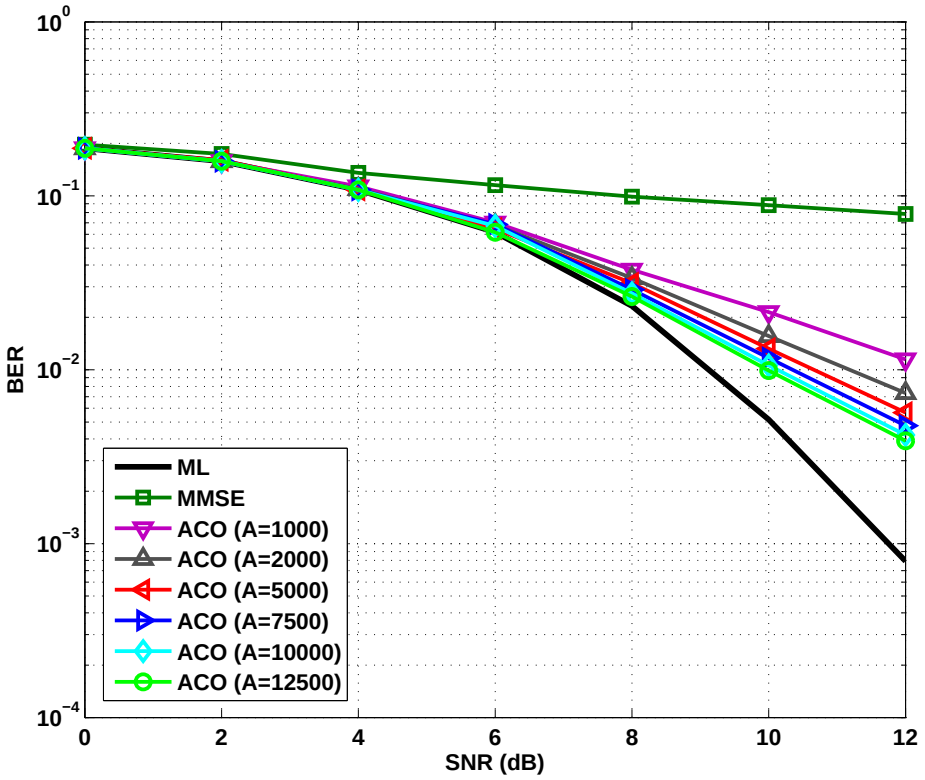

Figure 4.11 BER performance for an overloaded system with $N_T = 8, N_R = 6$.

To consider a more challenging scenario, the number of receive antennas are decreased to $N_R = 5$ whereas the number of transmitted signals are remained unchanged, $N_T = 8$. This translates into a load factor of $L_F = 1.60$. Fig. 4.13 compares the BER performance of different detectors. Although, the ACO detector show better results in comparison to the MMSE detector, an error floor starts appearing at high SNR values. This is due to the fact that the heuristic information defined for the ACO algorithm is based on the output of the MMSE detector. A better heuristic criteria may lead to better system performance[16]. A comparison of the performance between ML, MMSE and the proposed ACO assisted detector (with varying number of ants, A) is presented in Fig. 4.14. It can be seen that the addition of more ants produce better results but with increased computational complexity. Table 4.6 presents the complexity comparison of ML, MMSE and the ACO assisted detector.

[16]This is evident from the technique presented in the next chapter.

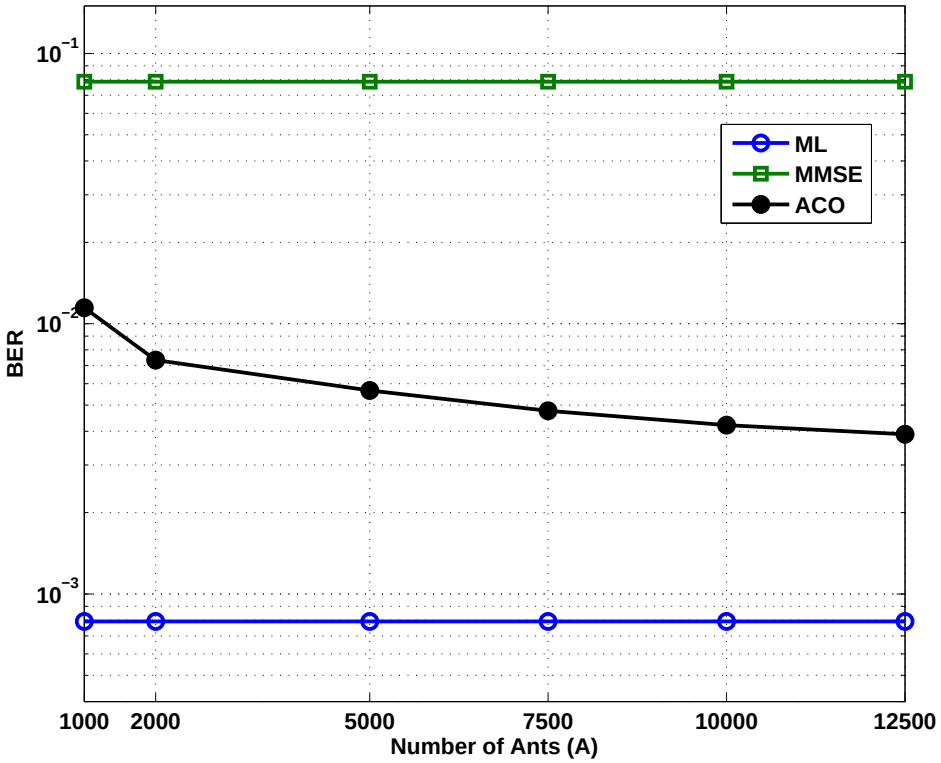

Figure 4.12 BER performance of $N_T = 8, N_R = 6$ cochannel system at $\gamma = 12$ dB.

Table 4.6 Complexity Comparison ($N_T = 8$, $N_R = 5$)

No. of Real Multiplications				% ML
ML	MMSE	ACO		
$1.18E7$	$4.77E3$	$A = 1000$	$2.16E5$	1.53%
$1.18E7$	$4.77E3$	$A = 2000$	$4.32E5$	3.05%
$1.18E7$	$4.77E3$	$A = 5000$	$1.08E6$	7.63%
$1.18E7$	$4.77E3$	$A = 7500$	$1.62E6$	11.44%
$1.18E7$	$4.77E3$	$A = 10000$	$2.16E6$	15.26%
$1.18E7$	$4.77E3$	$A = 12500$	$2.70E6$	19.07%
$1.18E7$	$4.77E3$	$A = 15000$	$2.70E6$	22.89%

Although the performance of the proposed detector in this section is not ML-approaching, the results are still encouraging and provide motivations for further investigations in this direction. Furthermore, it is important to note that the performance of the proposed ACO detector can be further improved by combining an error control coding scheme and designing an effective soft information processing decoder at the receiver.

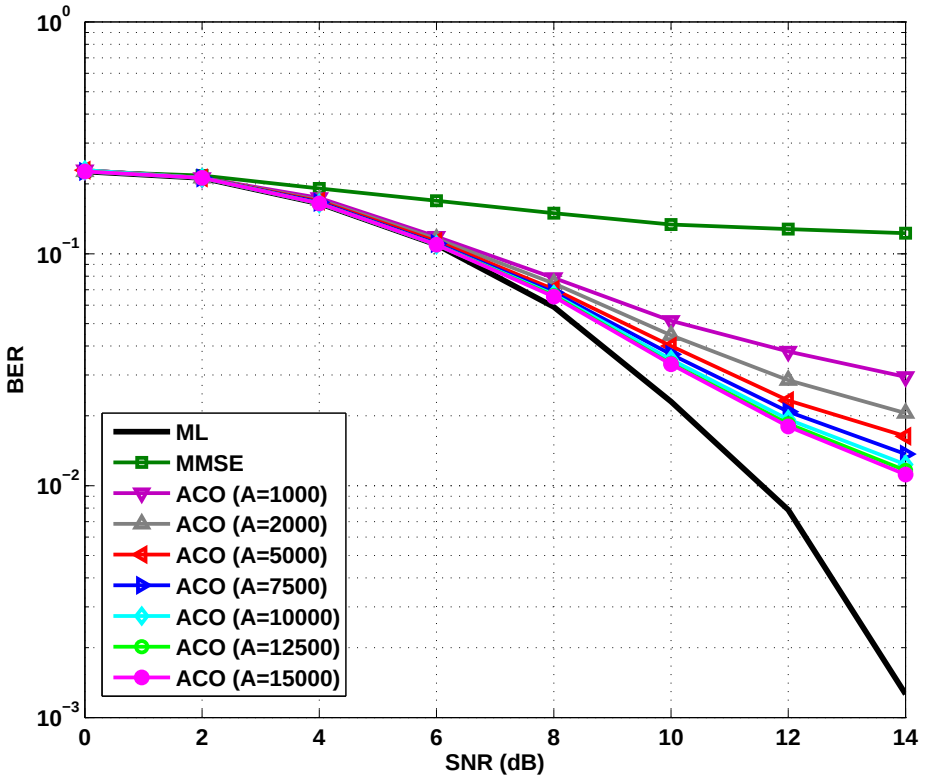

Figure 4.13 BER performance for an overloaded system with $N_T = 8, N_R = 5$.

4.5 SUMMARY

In this chapter, we consider a cochannel system where N_T signals are transmitted over a frequency-flat fading channel. The receiver employs N_R antennas to collect the faded and superimposed copies of the transmitted signals. We developed two reduced complexity detectors based on the ACO algorithm: a soft iterative processing detection algorithm for binary signals [87] and an ACO assisted detector for non-binary modulations.

The soft iterative processing detection algorithm for binary signals is composed of two stages: preprocessing and iterative solution refinement. During preprocessing, the receiver obtains an initial soft estimate and a hard decision from the linear MMSE detector. The hard decision is used to create a moving-pattern table while the soft estimate is employed in the calculation of the heuristic value. In the second phase,

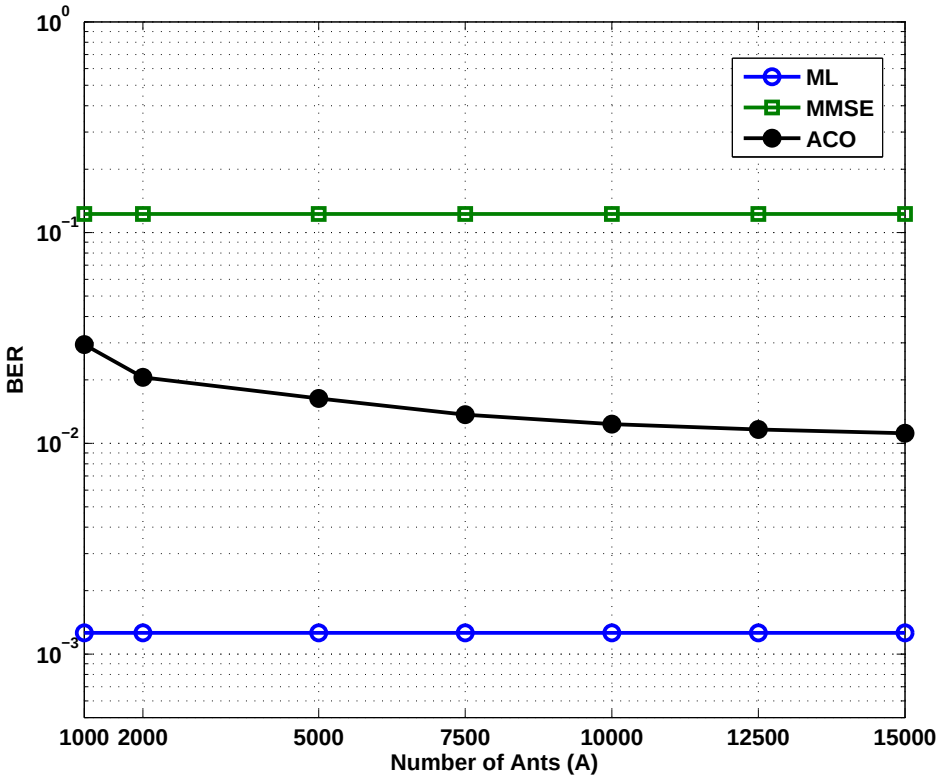

Figure 4.14 BER performance of $N_T = 8, N_R = 5$ cochannel system at $\gamma = 14$ dB.

ACO search process is iteratively executed to generate a list of likely candidate solution vectors. The ants walk through the moving-pattern table (or equivalently the problem graph) and probabilistically select one of the two components from each of its column (which correspond to an edge in the problem graph) based on the associated pheromone strength and the heuristic value. The search process benefits from the search experience of the ants stored in the pheromone table as well as the heuristic values generated using the updated soft signal estimates. At the end of the search process, each ant deposits its solution to a candidate solution list (containing a total of A candidate solution vectors). The list is used to update the soft estimate of the transmitted signals. The iterative process continues to perform a total of I iterations where a more refined and better quality solution emerges as a result. The detector outputs the final solution after the completion of the last iteration. The receiver structure is designed such that it continues operating under overloaded conditions. As such, its performance is

investigated for three different system load factors ($L_F = 1$, $L_F = 1.33$ and $L_F = 1.60$). It is observed that the proposed technique achieves near-ML performance in critically loaded cases ($L_F = 1$) with significantly reduced complexity. For the challenging case of overload ($L_F = 1.33$ and $L_F = 1.60$), it still offers performance close to ML at low to moderate SNR. The receiver however exhibits an error floor at high SNRs with increasing values of system load factor $L_F > 1$.

The application of ACO to non-binary modulations have resulted in a sub-optimum detector capable of handling lightly overloaded situations. The solution construction process is modeled as a shortest-path finding problem over a multigraph. The ants build candidate solutions by performing random walk on this graph. The algorithm employs the PDF of the MMSE estimate as the heuristic criteria and computes the heuristic information from it. ML Euclidean distance metric is used as a fitness function to evaluate the quality of solution formed during the search process. The performance of the developed approach is validated through computer simulations. Two different loading conditions are considered for this purpose namely, $L_F = 1.33$ and $L_F = 1.60$. For the case of lightly loaded conditions $L_F = 1.33$, the proposed detector offers sub-optimum but acceptable performance. However, the performance deteriorates as the load factor is increased to $L_F = 1.60$ and an error-floor starts appearing at high SNR. This is due to the fact that the heuristic information depends on the PDF of the MMSE estimate which itself suffers from error-floor in the high SNR regime. A better heuristic criteria may lead to better system performance which will become evident from the technique presented in the next chapter.

Chapter 5

NEAR-ML DETECTION OF COCHANNEL SIGNALS USING ANT COLONY OPTIMIZATION

5.1 INTRODUCTION

In this chapter, we investigate a low-complexity near-ML detection technique for overloaded cochannel communication systems, where there are fewer antennas at the receiver than there are transmitted signals. It is well-known that optimal performance is achieved using a joint ML receiver comes at the cost of complexity, which grows exponentially with the number of transmitted signals. Many problem-specific heuristic [2–5, 8, 9] and metaheuristic [6, 10, 11] reduced complexity techniques have been developed to achieve near-ML performance, but they either require large computational efforts or tend to degrade performance (especially at high SNR).

In the previous chapter, we proposed an iterative soft detection technique based on ACO which provided near-ML performance for significantly reduced complexity. However, this approach is only valid for BPSK modulated signals. Furthermore, an error-floor at high SNRs is also observed when the system load factor is increased beyond 1, i.e. for $L_F > 1$. The main reason for this is the degraded performance of the MMSE detector which is used to provide an initial estimate to the ACO assisted iterative detection stage.

Here, we describe an integrated framework comprising of an ACO metaheuristic and a recursively defined ML search criteria to handle multilevel modulations. The resulting receiver treats detection as finding a minimum-cost path over a weighted multigraph. Artificial ants are allowed to walk on the graph in order to probabilistically build

solutions using a modified ML metric. The receiver framework is independent of the system loading condition. It achieves near-ML performance with significant savings in the required computational efforts even in overloaded scenarios. As an example, for $N_t = 8$ QPSK/4-QAM transmitted signals and a receiver equipped with $N_r = 5$ receive antennas, the ACO based receiver offers approximately 94% complexity savings in comparison to the ML detector [23] and 69% complexity reduction compared to the GA based approach [10].

The rest of this chapter is organized as follows. The system model and the receiver structure is explained in Section 5.2. The detection framework of the proposed algorithm is developed in Section 5.3. In Section 5.4, graph and tree based representations of the detection problem are discussed in conjunction with an ACO algorithm. Section 5.5 presents a method to calculate the transition probability and the pheromone update strategy. The extension of the proposed detection technique beyond QPSK is explained in Section 5.6. An upper-bound on computational complexity of the proposed scheme is derived in Section 5.7. Simulation results are presented in Section 5.8 along with a discussion and comparison of complexity for different detection schemes. Finally, Section 5.9 presents a chapter summary to conclude the discussion.

5.2 SYSTEM MODEL AND RECEIVER STRUCTURE

We consider the multiple-antenna system shown in Fig. 5.1, where N_T cochannel signals arrive at a receiver equipped with N_R antennas after transmission through a Rayleigh fading channel. Each transmitted signal, $x_j \in \chi$, is chosen from a $|\chi|$-ary complex signal constellation. The complex fading gain between the j-th transmitted signal and the i-th receive antenna element is denoted h_{ij}. After passing the received signals through a matched filter followed by symbol-rate sampling, a vector of received signals, $\mathbf{y} \in \mathbb{C}^{N_R}$, can be formed which is related to the input vector $\mathbf{x} \in \mathbb{C}^{N_T}$ as

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \boldsymbol{\nu}, \tag{5.1}$$

where $\boldsymbol{\nu} \in \mathbb{C}^{N_R}$ is additive temporally white complex Gaussian noise vector with $E\left[\boldsymbol{\nu}\boldsymbol{\nu}^H\right] = \sigma_\nu^2 \mathbf{I}_{N_T}$. The matrix $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ contains complex channel gains interconnecting the transmitters and the receiver antenna array. The elements, h_{ij}, of $\mathbf{H}$ are assumed to be iid with mean zero and a variance of 0.5 per dimension.

The overall receiver structure is shown in Fig. 5.1. The received signal is first passed through a preprocessing stage after which the ACO search process is applied to construct a near optimal solution vector. We assume the channel estimator produces perfect CSI i.e. $\hat{\mathbf{H}} = \mathbf{H}$. The internal specifications and working of the channel estimation process are outside the scope of this work.

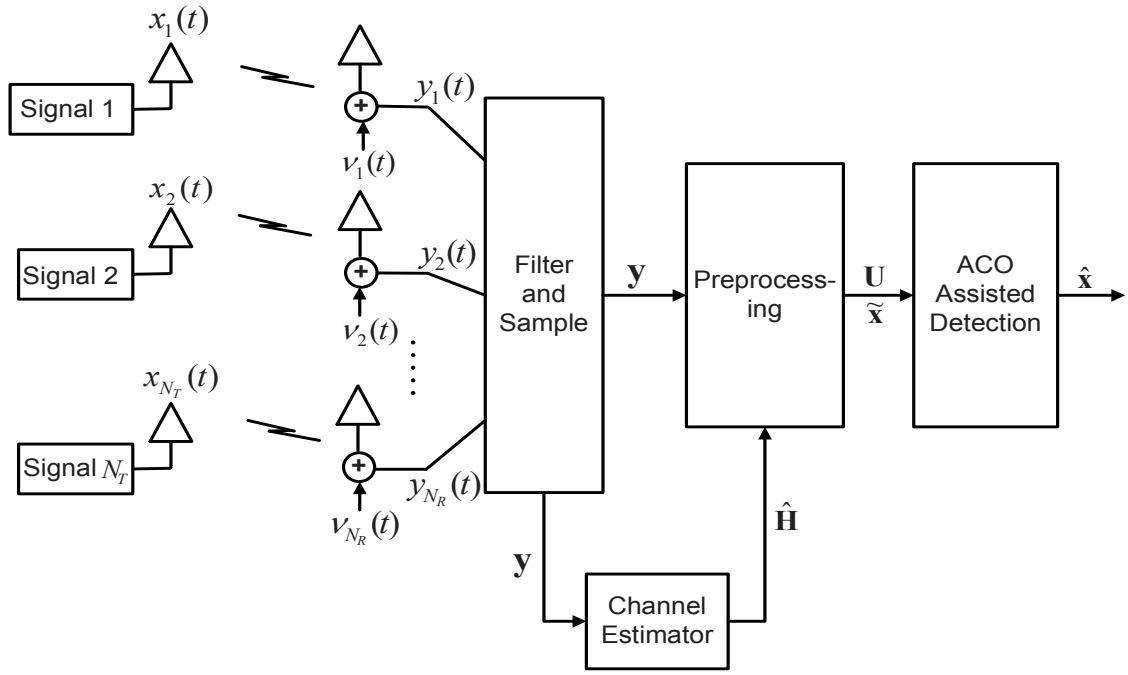

Figure 5.1 Receiver structure for the proposed detector.

5.3 THE HEURISTIC FRAMEWORK

The motivation for using ACO algorithms in MUD [50–52] is based on its successful application to some well-known NP-hard optimization problems including the traveling salesman problem (TSP) and the quadratic assignment problem [86]. Recently [31], ACO has been used for near ML detection in critically loaded MIMO systems.

However, the approach in [31] cannot handle overload because of its dependence on QR factorization of the equivalent real-valued channel matrix. Hence, we develop an algorithmic framework for overloaded conditions. An attractive feature of the proposed algorithm is its direct applicability to multi-level signal modulations. The resulting detection technique achieves ML performance when the system is critically loaded or overloaded with significant complexity savings.

We begin by specifying a theoretical framework for constant modulus constellations in order to apply the ACO algorithm. Using the Cholesky factorization, $\mathbf{U}^H\mathbf{U} = \left(\mathbf{H}^H\mathbf{H} - \sigma_\nu^2\mathbf{I}\right)$ and the linear MMSE estimate in (3.14) with the MMSE filter matrix given in (3.27), the modified ML metric is defined as

$$\begin{aligned}\Lambda'(\mathbf{x}) &= \|\mathbf{U}\left(\mathbf{x} - \tilde{\mathbf{x}}\right)\|^2, \\ &= \left(\mathbf{x} - \tilde{\mathbf{x}}\right)^H \mathbf{U}^H\mathbf{U}\left(\mathbf{x} - \tilde{\mathbf{x}}\right).\end{aligned} \tag{5.2}$$

Upon substituting $\mathbf{U}^H\mathbf{U} = \left(\mathbf{H}^H\mathbf{H} - \sigma_\nu^2\mathbf{I}\right)$ and $\tilde{\mathbf{x}} = \left(\mathbf{H}^H\mathbf{H} + \sigma_\nu^2\mathbf{I}_{N_t}\right)^{-1}\mathbf{H}^H\mathbf{y}$ into (5.2) and solving, we get

$$\begin{aligned}\Lambda'(\mathbf{x}) &= \left(\mathbf{x} - \tilde{\mathbf{x}}\right)^H \left(\mathbf{H}^H\mathbf{H} - \sigma_\nu^2\mathbf{I}\right)\left(\mathbf{x} - \tilde{\mathbf{x}}\right) \\ &= \mathbf{x}^H\mathbf{H}^H\mathbf{H}\mathbf{x} - \mathbf{y}^H\mathbf{H}\mathbf{x} - \mathbf{x}^H\mathbf{H}^H\mathbf{y} + \sigma_\nu^2\mathbf{x}^H\mathbf{x} + \mathbf{y}^H\mathbf{H}\left(\mathbf{H}^H\mathbf{H} + \sigma_\nu^2\mathbf{I}\right)^{-1}\mathbf{H}^H\mathbf{y}\end{aligned} \tag{5.3}$$

Adding and subtracting $\mathbf{y}^H\mathbf{y}$ on the right hand side of (5.3) and collecting the terms produces

$$\Lambda'(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \sigma_\nu^2\mathbf{x}^H\mathbf{x} + \mathbf{y}^H\left(\mathbf{H}\left(\mathbf{H}^H\mathbf{H} + \sigma_\nu^2\mathbf{I}\right)^{-1}\mathbf{H}^H - \mathbf{I}\right)\mathbf{y}. \tag{5.4}$$

Clearly, the last term in (5.4) does not depend on $\mathbf{x}$ and the middle term is a constant since $\mathbf{x}^H\mathbf{x} = \textit{constant}$, for constant modulus constellations. So, only the first term, $\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2$, depends on $\mathbf{x}$ and hence minimizing (5.2) is equivalent to minimizing the

ML Euclidean distance metric

$$\Lambda(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 . \tag{5.5}$$

We can therefore write

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \chi^{N_T}} \Lambda(\mathbf{x}) = \arg \min_{\mathbf{x} \in \chi^{N_T}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 . \tag{5.6}$$

The modified ML metric in (5.2) can be stated in equivalent form as

$$\begin{aligned} \Lambda'(\mathbf{x}) &= (\mathbf{x} - \tilde{\mathbf{x}})^H \mathbf{U}^H \mathbf{U} (\mathbf{x} - \tilde{\mathbf{x}}) \\ &= \sum_{i=1}^{N_T} \left| \sum_{k=1}^{N_T} u_{ik} (x_k - \tilde{x}_k) \right|^2 . \end{aligned} \tag{5.7}$$

This can be computed using a recursive definition employing the componentwise ML metric as [97]

$$\lambda_i = \lambda_{i+1} + \left| \sum_{k=i+1}^{N_T} u_{ik} (x_k - \tilde{x}_k) + u_{ii} (x_i - \tilde{x}_i) \right|^2 , \quad \text{for } i = (N_T - 1) \ldots 2, 1, \tag{5.8}$$

and

$$\lambda_{N_T} = |u_{N_T N_T} (x_{N_T} - \tilde{x}_{N_T})|^2 . \tag{5.9}$$

Recall, the ML detector in (3.9) searches a list $\mathcal{L}$ containing all $|\chi|^{N_T}$ possible combinations[1] of the vector to be estimated (i.e. all combinations of $\mathbf{x}$ where the individual components x_k are chosen from χ) and decides in favor of the one that yields the smallest value of (5.7). Using ACO, we will consider only a subset, $\bar{\mathcal{L}}$, of the most promising vectors from $\mathcal{L}$ as candidate solution vectors for consideration under the modified ML rule of (5.7). The ants are modeled as intelligent computational agents responsible for sensing more appropriate search regions containing better candidate solutions and thus assisting in forming the sublists $\bar{\mathcal{L}}$. To do so, the ants form a complete candidate solution using a problem-specific heuristic value and the pheromone levels associated

[1] ML detector actually treats all $|\chi|^{N_T}$ possible vectors as candidate solution vectors.

with the paths on which they travel.

5.4 SOLUTION CONSTRUCTION USING ACO

In our proposed ACO based approach, the detection problem is viewed as finding a minimum weight path over a weighted multigraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of $(N_T + 1)$ vertices or nodes and $\mathcal{E}$ is a set of $(|\chi|\, N_T)$ edges. Each ant is allowed to traverse the graph from one node to another along one of the $|\chi|$ edges.

5.4.1 Graph-Based Search Model

Detection starts by dispatching A ants to invoke the ACO search process and builds a solution in every search stage/iteration. At the end of the search process, a list $\bar{\mathcal{L}} \subseteq \mathcal{L}$ is generated containing a maximum of A possible candidate solution vectors. The a-th vector in the list can be expressed as $\mathbf{x}^a = \left[x_1^a\ x_2^a \dots x_{N_T}^a\right]$. Each ant treats the construction of the length N_T candidate solution vector as a random walk through $(N_T + 1)$ nodes. A total of $|\chi|$ edges or paths exist between any two consecutive nodes each uniquely representing one member of the signal constellation. As a first step towards solution construction, the a-th ant chooses an appropriate candidate symbol in place of $x_{N_T}^{(a)}$. Selection is based on the transition probability associated with the path, which depends on the path-specific heuristic value and the pheromone level[2]. The ant then moves to the next node, $(N_T - 1)$, to decide the next candidate symbol, $x_{N_T-1}^{(a)}$, thus forming a partial candidate solution vector $[x_{N_T-1}^{(a)}\ \ x_{N_T}^{(a)}]$. The process continues until the ant arrives at the last node, where it selects $x_1^{(a)}$, and thus the entire transmitted signal vector $\mathbf{x}^{(a)}$ is formed.

The process is illustrated in Fig. 5.2 for 4-QAM modulation which implies $|\chi| = 4$ and $\chi = \{-1 - j, -1 + j, 1 - j, 1 + j\}$. At node $(i + 1)$, the ant has $|\chi| = 4$ paths available to travel to the next node, node-i. Associated with each of these paths is its heuristic desirability η_{il}^a $(l = 1, 2 \dots |\chi|\,)$ calculated by the a-th ant, and the pheromone level, τ_{il}. Together these two quantities guide the ants in choosing one of the $|\chi|$ paths and hence in performing a complete walk to construct a length N_T candidate solution.

[2]Computation of the heuristic value and pheromone level is explained in detail in the next section.

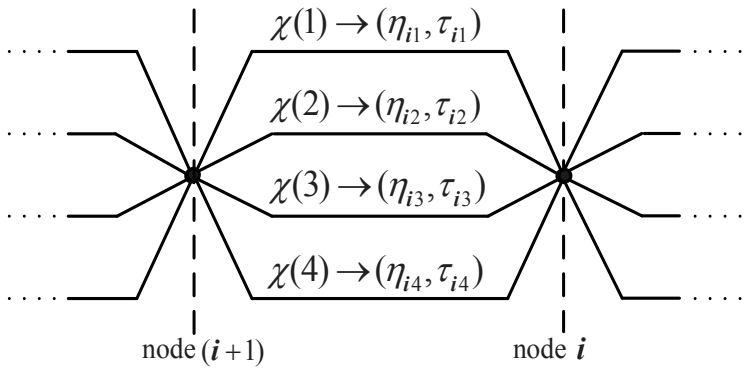

Figure 5.2 ACO detection process for 4-QAM modulation.

While the pheromone-levels are controlled differently for each variant of the ACO algorithm, the heuristic desirability, η_{il}^a, is calculated as

$$\eta_{il}^a = \frac{1}{\lambda_{il}^a}, \tag{5.10}$$

where λ_{il}^a is the partial Euclidean distance from node N_T to node i with only the l-th path from node $(i+1)$ to node i included[3]. It is also referred to as the fitness value of the l-th path from node $(i+1)$ to node i and is computed using (5.8) as

$$\lambda_{il}^a = \lambda_{i+1}^a + \left|\psi_i + u_{ii}\left(x_l - \tilde{x}_i\right)\right|^2, \quad l = 1 \ldots |\chi|,$$

where

$$\psi_i = \sum_{k=i+1}^{N_t} u_{ik}\left(x_k^{(a)} - \tilde{x}_k\right), \quad i = (N_T - 1) \ldots 2, 1 \text{ and } a = 1 \ldots A, \tag{5.11}$$

and λ_{i+1}^a is the partial ML metric (PMM) stored by the a-th ant after a specific path to node i is chosen. It takes into account the contribution from already detected symbols by the a-th ant, $x_{i+1}^a \ldots x_{N_T}^a$. The term, $u_{ii}(x_l - \tilde{x}_i)$, in (5.11) denotes the current symbol, which has not yet been decided. Also, the desirability for the path associated with symbol, x_{N_t}, which is detected first, is calculated using (5.9) as

$$\lambda_{N_T l}^a = \left|u_{N_T N_T}\left(x_l - \tilde{x}_{N_T}\right)\right|^2, \quad l = 1, 2 \ldots |\chi|. \tag{5.12}$$

[3]Note that x_{N_t} is the first symbol added to the candidate solution vector construction step. Likewise, x_{N_T-1} is added second and so on down to x_{i+1}.

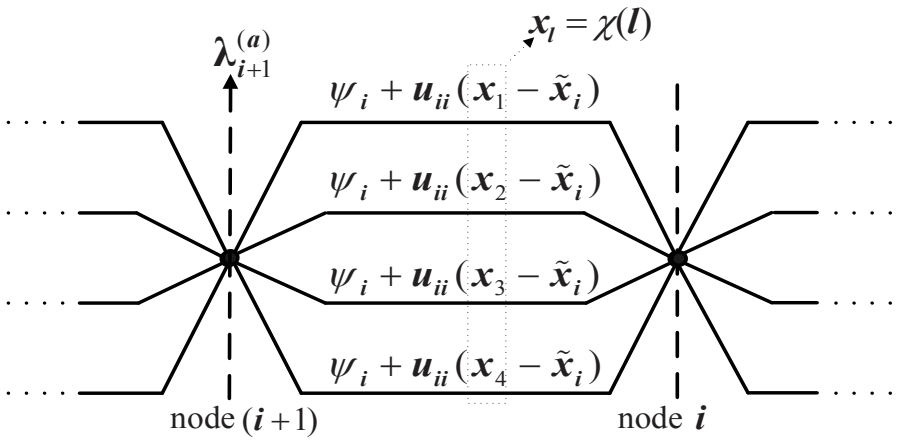

Figure 5.3 Heuristic value calculation using PMM.

Fig. 5.3 depicts the calculation of λ_{il}^a for the l-th path to node i. When the ant selects the final component of the solution vector, $x_1^{(a)}$, the PMM λ_1^a actually represents the complete ML metric (CMM) corresponding to that vector and therefore it is called the CMM value.

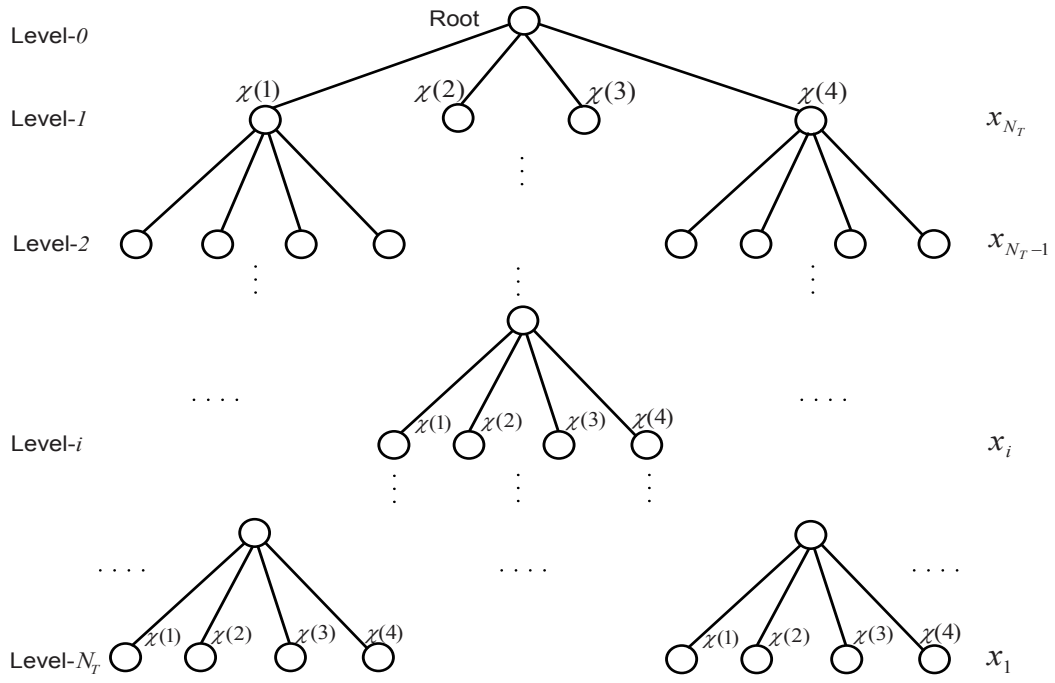

Figure 5.4 Tree-based model for the detection problem with QPSK / 4QAM in use.

5.4.2 Tree-Based Search Model

The detection problem can also be interpreted as a tree search problem with a total of $(N_T + 1)$ levels (the topmost level-0 only containing the root node) and $|\chi|$ branches originating from each node. A sample graph of the problem is depicted in Fig. 5.4. In this tree, i-th level corresponds to all the possibilities for the signal x_i. Solution construction is visualized as artificial ants performing a random walk over this tree starting from the root node down to the N_T-th level of the search tree.

Starting from the root node, the a-th ant probabilistically selects one out of $|\chi|$ available immediate descendant nodes. The transition probability of the a-th ant to the l-th child node, $p^a_{N_T l}$ (where $l = 1, 2 \ldots |\chi|$), is calculated based on the pheromone strength, $\tau_{N_T l}$, on the branch joining the two nodes and the associated heuristic value, $\eta_{N_T l}$. The heuristic value $\eta_{N_T l}$ for the first element of the candidate solution is calculated using (5.12) and (5.10) and a method for calculating p^a_{il} is explained in the next section.

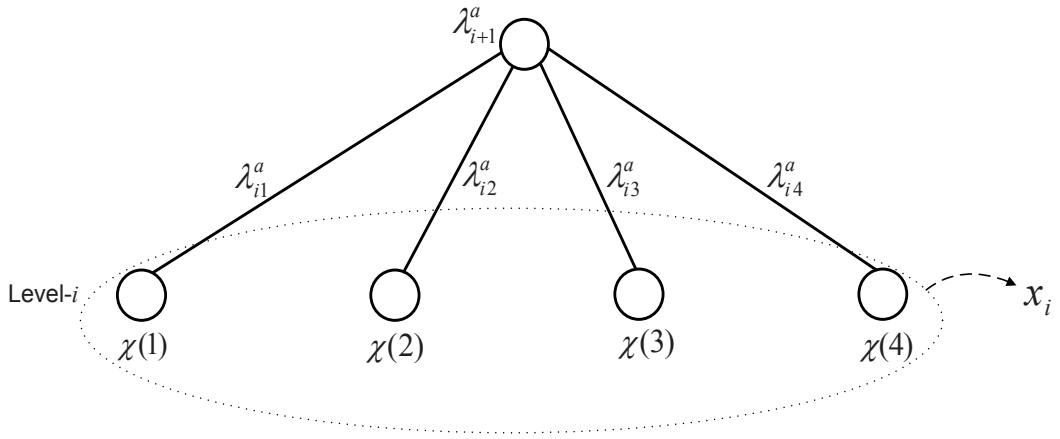

Figure 5.5 PMM computation for the $|\chi| = 4$ nodes at the i-th level for QPSK / 4QAM.

After selecting the N_T-th element of the partial solution vector $\mathbf{x}^a_{N_T} = \left[x^a_{N_T}\right]^T$, the a-th ant resides at one of the nodes at level-1 of the search tree. The next transition to a node at level-2 results in selection of the $(N_T - 1)$-th element of the solution vector. This produces an updated partial candidate solution vector $\mathbf{x}^a_{N_T-1} = \left[x^a_{N_T-1}\ x^a_{N_T}\right]^T$, where $x^a_i \in \chi$ represents a probable solution for the signal emitted by the i-th source. To evaluate the transition probability $p^a_{(N_T-1)l}$, the heuristic value is calculated using

(5.11) and the PMM value of $\lambda_{N_T}^a$ which is already evaluated while calculating the heuristic value of the N_T-th element. The PMM evaluation by the a-th ant at the i-th level of the search tree is depicted in Fig 5.5 while the child node selection based on the pheromone strength and the heuristic value is shown in Fig. 5.6.

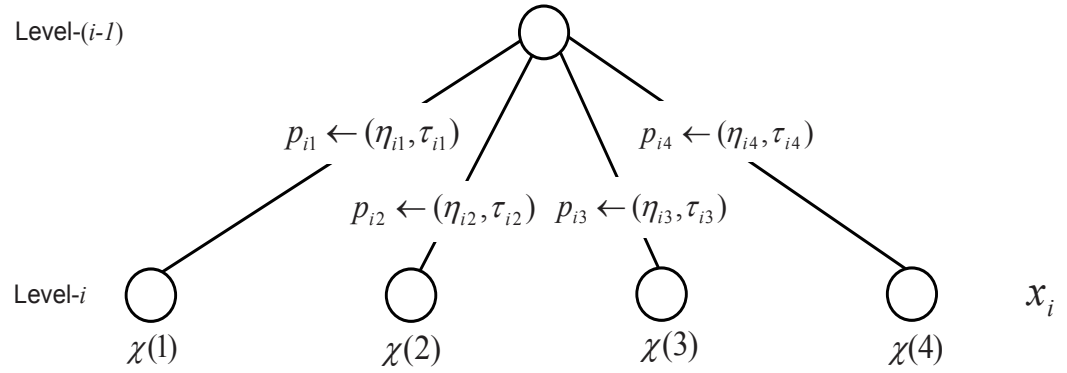

Figure 5.6 Transition probability calculation using the pheromone and the heuristic value for QPSK / 4QAM.

The solution construction process by the a-th ant continues in a similar manner until it arrives at the leaf node corresponding to the signal emitted by the first transmitter. At this point the entire candidate solution vector has been built by the a-th ant, i.e. $\mathbf{x}^a = \left[x_1^a \; x_2^a \; \ldots \; x_{N_T}^a\right]^T$.

Note that the computational complexity of our algorithm is already much less due to the recursive nature of the algorithm and the way the solution is constructed using only a limited number of ants (i.e. $A \ll |\chi|^{N_T}$). Furthermore, each ant is allowed to construct an entire solution vector before passing control to the next ant. By doing this, the ants producing bad solutions are pruned out at early stages, which results in a significant reduction in the complexity. To elaborate this point further, we state the following property regarding the PMM value of the a-th ant during the solution construction process.

Property-I: The PMM value of the a-th ant at the i-th node in the search tree is a non-decreasing function of the partial candidate solution vector.

$$\ldots \geq \lambda_i^a\left(\mathbf{x}_i^a\right) \geq \lambda_{i+1}^a\left(\mathbf{x}_{i+1}^a\right) \geq \ldots \tag{5.13}$$

where $\mathbf{x}_i^a = \begin{bmatrix} x_i^a & \mathbf{x}_{i+1}^a \end{bmatrix}$ for $i = N_T, \ldots, 1$ and $\mathbf{x}_{N_T+1}^a = \phi$ (here ϕ denote an empty set). Based on this property, the ant pruning strategy can be implemented as follows:

- At the start of the algorithm, when the first ant completes its turn, it copies its CMM value into a global variable ϱ which serves as a reference value for the ants to follow. This can be stated as

$$\varrho = \lambda_1^1. \tag{5.14}$$

- All the succeeding ants compare their PMM value (at every level of the search tree) with ϱ. If any of the ant's PMM values happen to be greater than this value, the ant ceases its execution, and passes control to the next ant. Consider the a-th ant at the i-th level of the tree during the solution construction process. At this place, its PMM value λ_i^a is compared with the ϱ as

```
if λ_i^a ≤ ϱ then
    process the child nodes;
else
    abort construction process;
end
```
(5.15)

- If the a-th ant manages to construct a solution vector with a CMM value less than ϱ, an update is performed on ϱ and its new value is set equal to the CMM value of the current ant.

$$\varrho = \lambda_1^a. \tag{5.16}$$

- The next ant then starts its solution construction process with the same branch and bound strategy.

The search process continues until all the ants finish their execution ($a = 1, 2 \ \ldots \ A$).

5.5 TRANSITION PROBABILITY AND PHEROMONE UPDATE

We now develop the transition probability calculation and the pheromone update strategy for the algorithm. This is performed according to the Max-Min Ant System (MMAS) algorithm[4].

5.5.1 Transition Probability

The algorithm starts with the initialization of the $(|\chi| \times N_T)$ pheromone array, $\boldsymbol{\tau}$. With $i = N_T$, the a-th ant calculates the heuristic desirability values $\eta^a_{N_T l}$ associated with the first signal to be detected using (5.12). The ant then chooses the l-th path to node-N_T (i.e. it selects $x^{(a)}_{N_T} = x_l$) with probability

$$p^a_{N_T l} = \frac{(\tau_{N_T l})^\alpha \cdot \left(\eta^a_{N_T l}\right)^\beta}{\sum_{k=1}^{|\chi|} (\tau_{N_T k})^\alpha \cdot \left(\eta^a_{N_T k}\right)^\beta}, \tag{5.17}$$

where $\tau_{N_T l}$ is the pheromone strength on the l-th path and the parameters α and β are ACO parameters for controlling the heuristic desirability and the pheromone strengths, respectively. The PMM of the chosen path becomes the current PMM, $\lambda^{(a)}_{N_T}$, which is used in heuristic value calculation of the next signal. During the i-th step, the ant calculates the heuristic value η^a_{il} associated with the i-th component of the candidate solution vector using (5.11). It then chooses the l-th path to node-i with probability

$$p^a_{il} = \frac{(\tau_{il})^\alpha \cdot (\eta^a_{il})^\beta}{\sum_{k=1}^{|\chi|} (\tau_{ik})^\alpha \cdot \left(\eta^a_{ik}\right)^\beta}. \tag{5.18}$$

To illustrate the computation of the heuristic value using PMM during the search process, an example of an $N_t = N_r = 4$ system is shown in Fig. 5.7.

5.5.2 Pheromone Update

After solution construction is complete, each successful ant deposits a length N_t candidate solution vector into the candidate list $\bar{\mathcal{L}}$ along with its associated CMM value.

[4]Appendix B presents a detailed account of MMAS algorithm for the interested readers.

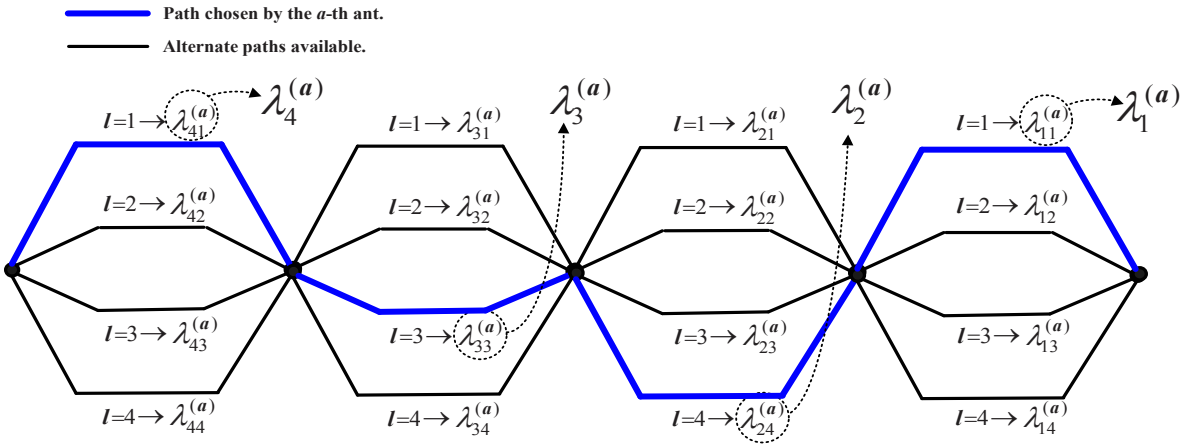

Figure 5.7 PMM generation on a graph during the detection process for a $N_T = N_R = 4$ system.

This list consists of a maximum of $\left|\bar{\mathcal{L}}\right| = A$ candidate vectors. The vector with the minimum CMM value is designated as the iteration's best solution, $\mathbf{x}^{\dagger}$, and is kept along with its CMM value $\Lambda^{\dagger}$ which is calculated as

$$\Lambda^{\dagger} = \min_{a} \lambda_1^{(a)}. \tag{5.19}$$

We now describe how the global best solution, $\mathbf{x}^{\ddagger}$, is updated based on $\mathbf{x}^{\dagger}$. If the algorithm is performing its first search iteration then $\mathbf{x}^{\ddagger} = \mathbf{x}^{\dagger}$ and the CMM of $\mathbf{x}^{\ddagger}$ is

$$\Lambda^{\ddagger} = \Lambda^{\dagger}. \tag{5.20}$$

In later iterations, the CMM values of the current global best solution and iteration best solution are compared and the one with the smallest value is declared to be the new global best solution

$$\mathbf{x}^{\ddagger} = \begin{cases} \mathbf{x}^{\ddagger}, & \text{if } \Lambda^{\ddagger} <= \Lambda^{\dagger} \\ \mathbf{x}^{\dagger}, & \text{if } \Lambda^{\ddagger} > \Lambda^{\dagger} \end{cases} \tag{5.21}$$

and

$$\Lambda^{\ddagger} = \min(\Lambda^{\ddagger}, \Lambda^{\dagger}). \tag{5.22}$$

Next, the pheromone update phase starts with the intention of modifying $\boldsymbol{\tau}$. This

is done in a manner that reflects the experience of the ants about the search space. Generally, the paths corresponding to the regions of the search space offering more promising candidate solutions are promoted by incrementing their pheromone strength. In the ACO literature, there are various algorithm-specific strategies for modifying the pheromone along the paths traversed by the ants. In MMAS, the pheromone is updated by either the ant that produces the iteration best solution or the ant corresponding to the global best solution. We use the iteration best strategy to update $\boldsymbol{\tau}$ which is carried out in two steps. In the first step, pheromone evaporation is performed as [86]

$$\tau_{mn} = (1-\rho)\,\tau_{mn},\ \ \forall\,(m,n), \tag{5.23}$$

where ρ is the pheromone evaporation rate. In MMAS, ρ is assigned a small value so that the entire search space is thoroughly explored to find better solutions. We use $\rho = 0.02$ as suggested in [86]. In the second step, the pheromone is deposited along the edges corresponding to the iteration best solution, $\mathbf{x}^{\dagger}$. This is done in proportion to the fitness value of $\mathbf{x}^{\dagger}$ obtained in (5.19) as [86]

$$\tau_{mn} = \tau_{mn} + \Delta\tau^{\dagger}_{mn}, \tag{5.24}$$

where

$$\Delta\tau^{\dagger}_{mn} = \frac{1}{\Lambda^{\dagger}},\quad (m,n) \in \mathbf{x}^{\dagger}. \tag{5.25}$$

Moreover in MMAS, it is ensured that the pheromone values corresponding to each path are kept inside a particular interval $[\tau_{min}, \tau_{max}]$. This implies τ_{mn} can assume a minimum value of τ_{min}, $(\forall m,\ n)$ and it cannot exceed the maximum value of τ_{max}, $(\forall m,\ n)$ at all times. The maximum possible pheromone value τ_{max} is defined using ρ and (5.19) as [86]

$$\tau_{max} = \frac{1}{\rho \cdot \Lambda^{\dagger}}. \tag{5.26}$$

It is important to note that the initial values of the pheromone on all the paths are set equal to τ_{max}, i.e. $\boldsymbol{\tau}_0 = \tau_{max} \cdot \mathbf{1}$, where $\mathbf{1}$ is an all ones matrix with the same dimensions as $\boldsymbol{\tau}$. Using the maximum pheromone value τ_{max} stated in (5.26) and another MMAS

algorithm parameter a_0 [86], the minimum pheromone value τ_{min} can be computed as

$$\tau_{min} = \frac{\tau_{max}}{a_0}. \tag{5.27}$$

The parameter a_0 is defined as

$$a_0 = \frac{(N_T - 1)\ \sqrt[N_T]{p_{best}}}{1 - \sqrt[N_T]{p_{best}}}, \tag{5.28}$$

where p_{best} is the probability of generating a best solution when the solution has converged. A suitable value for p_{best} is recommended as 0.05 in [96].

To justify the use of the iteration best strategy for pheromone updating, note that if the update is performed by the ant corresponding to the global best solution, then higher pheromone values will be accumulated along the corresponding edges and the edges in the neighborhood. This will bias the search process within the first few search stages / iterations by restricting the focus to a subspace (in the neighborhood of the global best solution) of the ML search space. Also, when the global best solution is not the same as the optimum solution, the detector can possibly get stuck in a local minimum. On the other hand, when using the iteration best strategy, there is a likelihood of a different iteration best solution each time. Hence, more edges will contain larger pheromone levels and the search will be less directed to a specific region for a few more initial iterations. This can help the detector avoid convergence to a local minimum.

Finally, for exploring the regions of the search space corresponding to a small probability of being searched, the pheromone trails are occasionally reinitialized in the MMAS algorithm. This reinitialization takes place only if the algorithm is unable to produce an improved solution after a given number of search stages / iterations. The maximum pheromone level, τ_{max}, is used to reinitialize $\boldsymbol{\tau}$ as follows

$$\boldsymbol{\tau} = \tau_{max} \cdot \mathbf{1}. \tag{5.29}$$

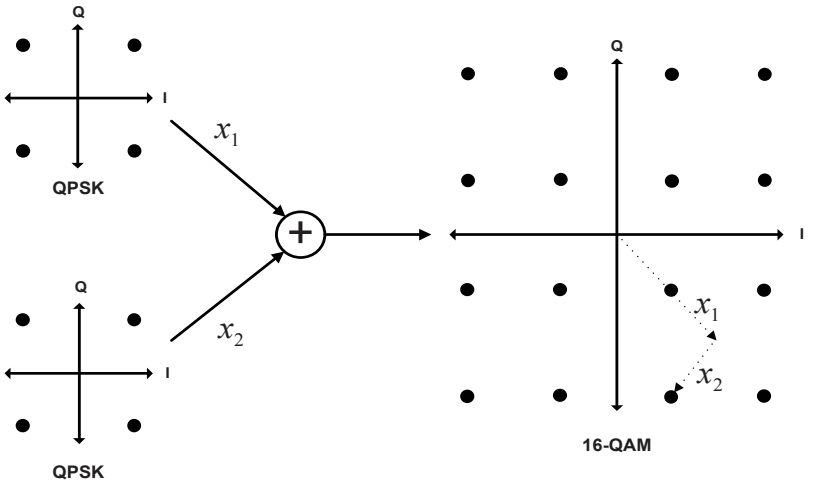

Figure 5.8 Construction of a 16-QAM symbol through addition of two QPSK symbols.

5.6 HIGHER ORDER MODULATIONS: THE $|\chi|$-ARY QAM

To improve spectral efficiency and achieve higher data rates, the transmitted signal can be chosen from a higher order modulation. We now explain the application of the proposed ACO assisted detection technique when the transmitted signals are chosen from a $|\chi|$-ary QAM, where $|\chi| = 2^b$ and b is an even number greater than 2. Following the results presented in [98], any $|\chi|$-ary square QAM signal x can be expressed as a linear combination of $\frac{b}{2}$ QPSK constellations as

$$x = \sum_{l=1}^{\frac{b}{2}-1} \left(\frac{2^l}{\sqrt{2}} \right) x_l, \tag{5.30}$$

where $x_l \in \chi_q$ and χ_q is a QPSK signal constellation. As an example, the construction of a 16-QAM symbol using two QPSK symbols is illustrated in Fig. 5.8.

Using this idea, a cochannel system with N_T transmitted signals, each chosen from a $|\chi|$-ary modulation, can be decomposed into an equivalent system having $\frac{b}{2}N_T$ transmitted signals each belonging to a QPSK modulation. For the purpose of demonstration, when $|\chi| = 16$ and $b = 4$, the 16-QAM transmitted signal vector $\mathbf{x}$ can be written as [99, 100]

$$\mathbf{x} = \sqrt{2}\mathbf{x}_1 + \frac{1}{\sqrt{2}}\mathbf{x}_2, \tag{5.31}$$

where $\mathbf{x}_1 \in \chi_q^{N_T}$ and $\mathbf{x}_2 \in \chi_q^{N_T}$. Substituting (5.31) back into (5.1), the received signal vector becomes

$$\begin{aligned}
\mathbf{y} &= \mathbf{H}\,\mathbf{x} + \boldsymbol{\nu} \\
&= \mathbf{H}\,\left(\sqrt{2}\mathbf{x}_1 + \frac{1}{\sqrt{2}}\mathbf{x}_2\right) + \boldsymbol{\nu} \\
&= \sqrt{2}\mathbf{H}\,\mathbf{x}_1 + \frac{1}{\sqrt{2}}\mathbf{H}\,\mathbf{x}_2 + \boldsymbol{\nu} \\
&= \begin{bmatrix} \sqrt{2}\mathbf{H} & \frac{1}{\sqrt{2}}\mathbf{H} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \boldsymbol{\nu} \\
&= \bar{\mathbf{H}}\bar{\mathbf{x}} + \boldsymbol{\nu}
\end{aligned} \tag{5.32}$$

where the matrix $\bar{\mathbf{H}} \in \mathbb{C}^{N_R \times 2N_T}$ and the vector $\bar{\mathbf{x}} \in \chi_q^{2N_T}$ are defined as

$$\bar{\mathbf{H}} = \begin{bmatrix} \sqrt{2}\mathbf{H} & \frac{1}{\sqrt{2}}\mathbf{H} \end{bmatrix} \text{ and } \bar{\mathbf{x}} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \tag{5.33}$$

respectively. After performing the Cholesky decomposition on $\bar{\mathbf{H}}$ and defining the modified ML criteria as in (5.7), the problem can be represented as a tree with a total of $(2N_T + 1)$ levels. Hence, the proposed detection algorithm in Section 5.5 can now be easily applied to (5.32) with 16-QAM signals replaced with the QPSK signals.

5.7 COMPUTATIONAL COMPLEXITY OF PROPOSED ALGORITHM

We now discuss the computational complexity of the algorithms. It is important to mention that the complexity expressions derived in this section are approximate. Exact algorithmic complexity also depends on factors such as the number of comparisons, memory and the processing capabilities. Before proceeding with the derivation, general rules about the conversion from complex number figures to real number figures are introduced.

5.7.1 Complex-to-Real Conversions

We begin with the complex addition (CA) operation when two complex numbers z_1 and z_2 are added. Suppose $z_1 = x_1 + jy_1$ and $z_2 = x_2 + jy_2$, then from knowledge of elementary complex numbers

$$z_1 + z_2 = (x_1 + jy_1) + (x_2 + jy_2) = (x_1 + x_2) + j(y_1 + y_2). \tag{5.34}$$

Clearly, one complex addition is equivalent to two real additions. Secondly, a complex multiplication (CM) can be written as

$$z_1\, z_2 = (x_1 + jy_1)\,(x_2 + jy_2) = (x_1x_2 - y_1y_2) + j(x_1y_2 + x_2y_1). \tag{5.35}$$

Thus a complex multiplication requires four real multiplications and two real additions according to (5.35). Also, multiplication of a real scalar number c and a complex number $z = x + jy$ results in two real multiplications. Finally, the absolute value squared of z can be expressed as

$$|z|^2 = x^2 + y^2 = x\, x + y\, y. \tag{5.36}$$

This shows that the absolute value operation on a complex number costs two real multiplications and one real addition. To summarize the results in this section Table 5.1 present a summary of the conversions of complex operations to real operations. It is assumed that z_1, z_2 and z are complex while c is a real scalar number.

Table 5.1 Conversion from complex to real operations

	Operands	Operation	Real Multiplications	Real Additions
1	z_1, z_2	Multiplication (z_1z_2)	4	2
	c, z_2	Multiplication (cz_2)	2	-
2	z_1, z_2	Addition ($z_1 + z_2$)	-	2
3	z	Absolute Value ($\|z\|^2$)	2	1

5.7.2 Upper Bound on Complexity

In ACO, each ant has to recursively calculate the PMM metric using (5.11) and (5.12), which represent most of the computational expenses. Note that due to the adaptive ant strategy employed by the ACO algorithm, exact closed form expressions are not possible for the number of additions and multiplications. Instead, the expressions derived in this section serve as an upper-bound on the complexity of the proposed algorithm. The average complexity figures reported in the results section are obtained from computer simulations and are lower than the values resulting from these bounds. This shows the benefit of terminating the ants as soon as it is clear they cannot produce a better solution.

Complexity for QPSK / 4-QAM Modulation

In this section, all the transmitted signals are assumed to be QPSK modulated and the complexity expressions are developed accordingly. A discussion on the higher order modulations is presented in the next section.

To have a clear understanding of the computational complexity, we express the PMM calculation for the construction of an entire candidate solution vector as

$$\lambda_{N_T l} = \left| u_{N_T N_T} \left(x_l - \tilde{x}_{N_T} \right) \right|^2 . \tag{5.37}$$

$$\lambda_{il} = \lambda_{i+1} + \left| \psi_i + u_{ii} \left(x_l - \tilde{x}_i \right) \right|^2 , \quad i = N_T - 1, N_T - 2 \ldots 1 \tag{5.38}$$

where

$$\psi_i = \sum_{j=i+1}^{N_T} u_{ij} \left(x_j - \tilde{x}_j \right) , \quad i = N_T - 1, N_T - 2 \ldots 1. \tag{5.39}$$

Recalling that the entries on the main diagonal of the matrix $\mathbf{U}$ are real, a closer look at (5.37) reveals that it requires $4\left|\chi\right|$ real multiplications and $3\left|\chi\right|$ real additions (4 real multiplications and 3 real additions each time x_l takes a value from χ).

For the computation of PMM at the i-th recursive step using (5.38), it is assumed that the $N_T - i$ factors $(x_j - \tilde{x}_j)$ appearing in the summation in (5.39) are already computed in the last $N_T - i$ recursive steps and can therefore be used. Also the whole term ψ_i

remains constant for all the hypothesized values of x_l in (5.38). Therefore, i-th recursive computation of the PMM in (5.38) requires $4|\chi| + 4(N_T - i)$ real multiplications and $6|\chi| + 2(N_T - i) + 2(N_T - i - 1)$ real additions. Adding all these gives the complexity of the calculations performed by a single ant to construct the entire candidate solution vector, namely $4N_T|\chi| + 2N_T(N_T - 1)$ real multiplications and $3(2N_T + 1)|\chi| + 2(N_T - 1)^2$ real additions. Since there are A ants searching for the solution in a total of I ACO iterations, the overall complexity becomes

$$\mathcal{R}^{+}_{ACO} = \left[3(2N_T + 1)|\chi| + 2(N_T - 1)^2\right] AI \tag{5.40}$$

$$\mathcal{R}^{\times}_{ACO} = \left[4N_T|\chi| + 2N_T(N_T - 1)\right] AI, \tag{5.41}$$

where $\mathcal{R}^{+}_{ACO}$ and $\mathcal{R}^{\times}_{ACO}$ denote the number of real addition and multiplication operations associated with the ACO detection, respectively.

Complexity for Higher Order QAM

As explained in Section 5.6, a multi-antenna system where the N_T transmitted signals are chosen from a $|\chi| = 2^b$-ary QAM, can be represented by an equivalent system with $\frac{b}{2}N_T$ QPSK transmitted signals. The equivalence is demonstrated in (5.33) for a system employing 16-QAM signalling. Using (5.33), the problem can be represented by a tree with $(2N_T + 1)$ levels (since $b = 4$ for 16-QAM modulations). The topmost level contains only the root node. Each of the nodes has four child nodes except for the nodes at the bottom most level called the leaf nodes. An ant begins the search process at the root node and travels down towards the leaf nodes thus selecting a particular component of the solution vector at each level based on its transition probability[5].

For a general $|\chi| = 2^b$-ary QAM, each ant has to walk through a total of $\left(\frac{b}{2}N_T + 1\right)$ levels on the QPSK based tree to build a complete candidate solution. From the point

[5]Explained in detail in Section 5.4.

of view of complexity, this translates into the following expressions

$$\bar{\mathcal{R}}^{+}_{ACO} = \left[3\left(bN_T+1\right)|\chi| + \frac{1}{2}(bN_T-2)^2\right] AI \tag{5.42}$$

$$\bar{\mathcal{R}}^{\times}_{ACO} = \left[2bN_T|\chi| + \frac{b}{2}N_T\left(bN_T-2\right)\right] AI, \tag{5.43}$$

where $\bar{\mathcal{R}}^{+}_{ACO}$ and $\bar{\mathcal{R}}^{\times}_{ACO}$ respectively denote the required number of real addition and multiplication operations.

Table 5.2 Number of Real Multiplications for various Detection Techniques

	Algorithm	No. of Real Multiplications
1	ML	$\mathcal{R}^{\times}_{ML} = \left(4N_RN_T + 4N_R\right)|\chi|^{N_T}$
2	MMSE	$\mathcal{R}^{\times}_{MMSE} = 4N_T\left(N_T^2 + 2N_RN_T + N_R\right)$
3	ACO (QPSK Modulation)	$\mathcal{R}^{\times}_{ACO} = \left[4N_T|\chi| + 2N_T\left(N_T-1\right)\right] AI$
4	ACO ($\|\chi\|$-ary Modulation)	$\bar{\mathcal{R}}^{\times}_{ACO} = \left[2bN_T|\chi| + \frac{b}{2}N_T\left(bN_T-2\right)\right] AI$

Table 5.2 presents a summary of the approximate computational complexity in terms of number of real multiplications (RM) for the MMSE, ML and the ACO detection algorithms. Similarly, the complexity in terms of real additions (RA) is provided in Table 5.3.

Table 5.3 Number of Real Additions for various Detection Techniques

	Algorithm	No. of Real Additions
1	ML	$\mathcal{R}^{+}_{ML} = \left(4N_RN_T + 4N_R - 2\right)|\chi|^{N_T}$
2	MMSE	$\mathcal{R}^{+}_{MMSE} = 4N_T\left(N_T^2 + 2N_RN_T + N_R\right)$
3	ACO (QPSK Modulation)	$\mathcal{R}^{+}_{ACO} = \left[3\left(2N_T+1\right)|\chi| + 2(N_T-1)^2\right] AI$
4	ACO ($\|\chi\|$-ary Modulation)	$\bar{\mathcal{R}}^{+}_{ACO} = \left[3\left(bN_T+1\right)|\chi| + \frac{1}{2}(bN_T-2)^2\right] AI$

Note that the a-th ant aborts the solution construction process if its PMM value at the i-th level of the search tree, $\lambda_i^a > \varrho$. In this case, the PMM computations from level-i down to level-1 are not performed, which results in computational savings. This actually forms the basis of complexity reduction in our developed scheme.

5.8 SIMULATION RESULTS

We now present simulation results of the detection schemes operating under frequency-flat fading channel conditions. Perfect CSI is assumed at the receiver. Due to hard

symbol detection based recursive specifications of the proposed ACO assisted detector, $I = 1$, is used in all the considered cases. Simulation parameters for the results presented in this section are summarized in Table 5.4.

Table 5.4 Detector parameters used in simulations

	N_T	N_R	A	I	Modulation, $\|\chi\|$
Fig. 5.9	8	5	1000, 2000, 5000, 10000	1	4-QAM
Fig. 5.10	8	5	1000, 2000, 5000, 10000	1	4-QAM
Fig. 5.11	8	4	1000, 2000, 5000, 10000	1	4-QAM
Fig. 5.12	8	4	1000, 2000, 5000, 10000	1	4-QAM
Fig. 5.13	9	6	5000, 15000, 30000, 60000	1	4-QAM
Fig. 5.14	9	6	5000, 15000, 30000, 60000	1	4-QAM
Fig. 5.15	4	3	1000, 2000, 5000, 8000, 10000	1	16-QAM
Fig. 5.16	4	2	1000, 2000, 5000, 8000, 10000	1	16-QAM

First, we consider an overloaded cochannel system with $N_T = 8$ transmitted signals and $N_R = 5$ receive antennas having a load factor of $L_F = 1.60$. All the transmitted signals are assumed to be QPSK modulated. Fig. 5.9 depicts the BER performance comparison of ML, MMSE, and ACO. Under the assumed overloaded situation, the MMSE detector shows severe performance degradation in comparison to the ML detector and is therefore unusable. On the other hand, it is clearly evident that the proposed ACO based detection technique is capable of achieving ML performance with a sufficient number of ants. The number of ants can be varied to provide a range of performance-complexity tradeoffs. In Fig. 5.10, BER performance is shown for a varying number of ants, A, at an SNR of $\gamma = 16dB$. It can be seen that the ACO detector achieves the performance of ML detection by employing $A = 10000$ ants when only one iteration is used.

Table 5.5 Complexity Comparison ($N_T = 8, N_R = 5, |\chi| = 4$)

Average No. of Real Multiplications				% Comparison with ML and GA	
ACO		ML	GA [10]	ML	GA [10]
$A = 1000$	$1.13E5$	$1.18E7$	$2.21E6$	0.96%	5.12%
$A = 2000$	$2.23E5$	$1.18E7$	$2.21E6$	1.89%	10.08%
$A = 5000$	$3.73E5$	$1.18E7$	$2.21E6$	3.16%	16.84%
$A = 10000$	$6.77E5$	$1.18E7$	$2.21E6$	5.73%	30.59%

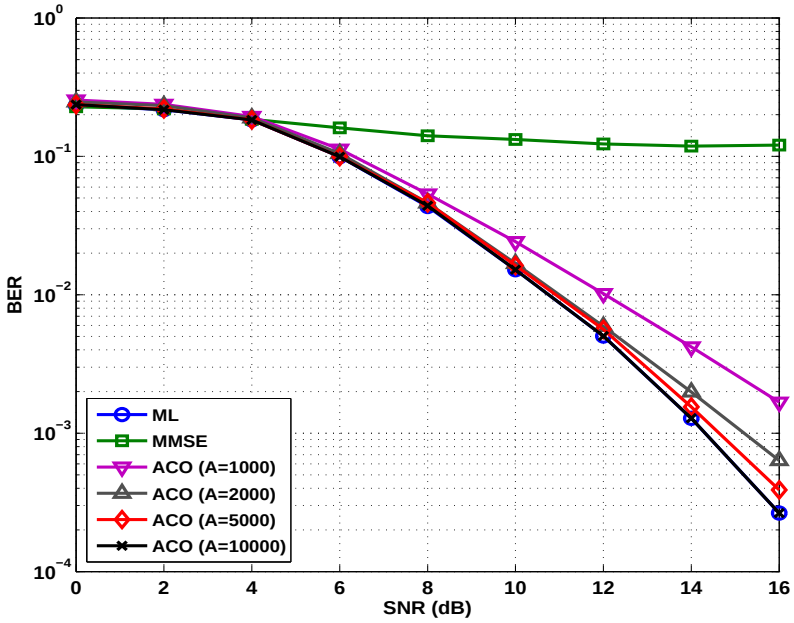

Figure 5.9 Performance of $N_T = 8$, $N_R = 5$ cochannel system employing QPSK modulation.

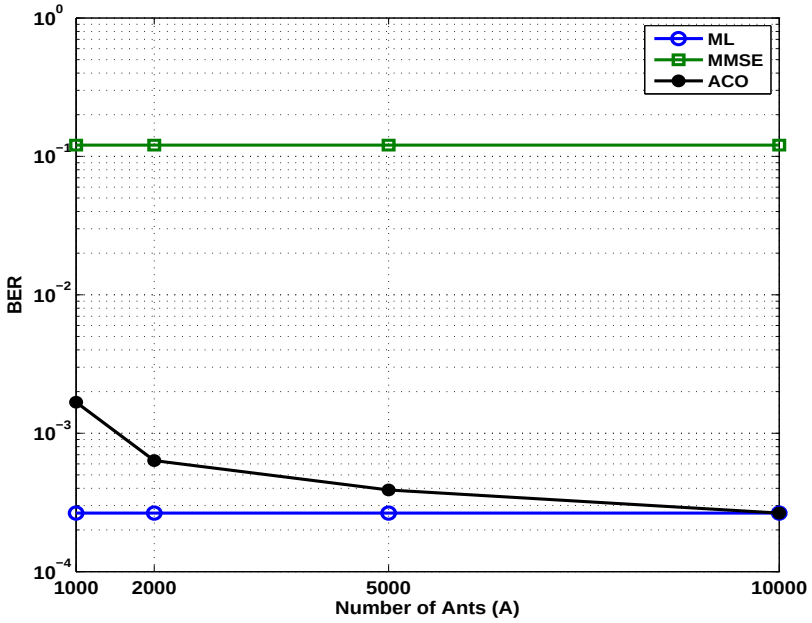

Figure 5.10 Performance of $N_T = 8$, $N_R = 5$ cochannel system at $\gamma = 16dB$.

Table 5.5 provides approximate complexity figures in terms of the number of real multiplications, $\mathcal{R}^{\times}$, along with the percentage comparison of the proposed ACO detection

technique with ML and the existing GA based approach in [10]. Again, it can be seen that ACO detection achieves the performance of ML detection with 94.27% complexity savings. Furthermore, it is shown in [10] that the GA requires $G = 12$ generations and a population size of $P = 1024$ in order to achieve performance close to that of ML detection. This results in a $P \times G = 1024 \times 12 = 12288$ evaluations of the so called fitness function in (5.5). One such evaluation requires $(4N_R N_T + 4N_R - 2)$ real additions and $(4N_R N_T + 4N_R)$ real multiplications. Hence, the GA approach results in a total of $(4N_R N_T + 4N_R - 2)PG$ real additions and $(4N_R N_T + 4N_R)PG$ real multiplications, respectively. Table 5.5 also shows that ACO assisted detection for overloaded systems outperforms GA based detection in terms of complexity with 69.41% complexity savings in comparison to the GA approach proposed in [10].

Fig. 5.11 presents the results for a more heavily overloaded system with $N_T = 8$ cochannel signals chosen from a QPSK modulation. The receiver at the base-station is assumed to be employing $N_R = 4$ antennas which increases the system load factor to $L_F = 2.00$. Here again it can be observed that the performance of the ACO algorithm approaches that of ML with a sufficient number of ants. In Fig. 5.12, performance of the same system at an SNR of $18dB$ with a varying number of ants is shown. The detection converges to that of ML as the number of ants is increased.

Another set of results in Fig. 5.13 and Fig. 5.14 show the performance of a cochannel system with $N_T = 9$ transmitted signals and a receiver employing $N_R = 6$ receive antennas. The load factor in this case is $L_F = 1.50$. Again, each transmitter uses QPSK modulation for the purpose of signalling. It can be seen that the performance of our proposed algorithm approaches that of ML as the number of ants is increased. Note that more ants are required for this larger system to achieve ML performance, namely $A = 60000$. Although, as shown there are a range of complexity performance tradeoff options available.

In Fig. 5.15, the error-rate performance of an overloaded system with $N_T = 4$ transmitted signals is depicted. All the signals are chosen from a 16-QAM modulation. The receiver employs $N_R = 3$ antennas for reception. The ACO assisted detector achieves the performance of ML using $A = 10000$ ants. A more severe overloaded system with

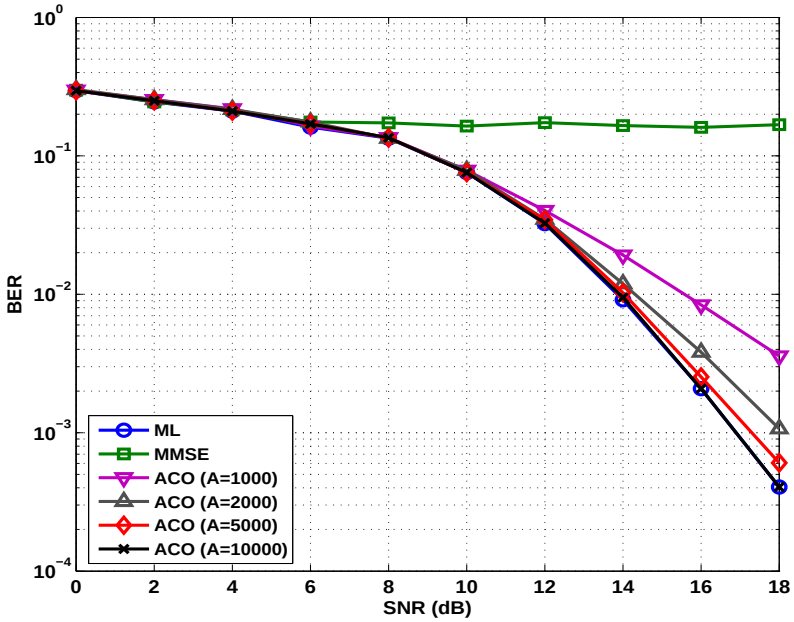

Figure 5.11 Performance of $N_T = 8$, $N_R = 4$ cochannel system employing QPSK modulation.

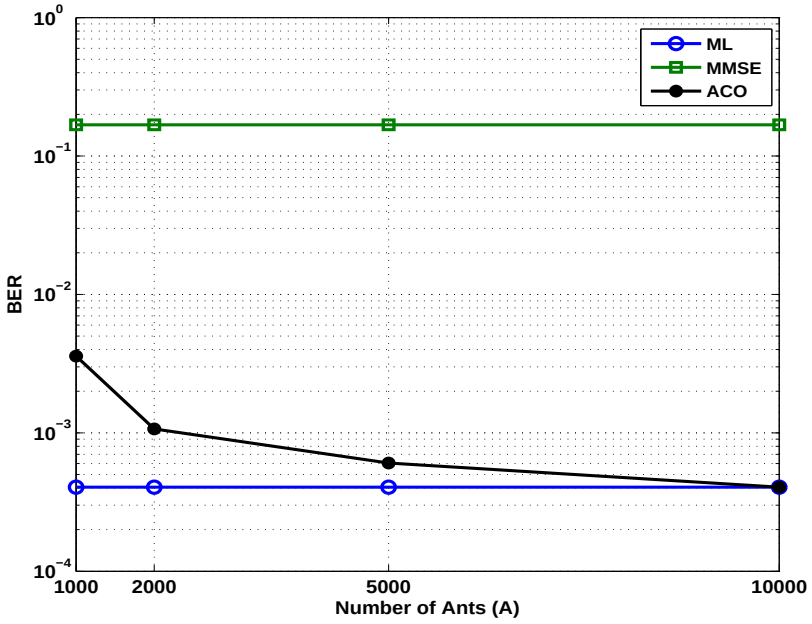

Figure 5.12 Performance of $N_T = 8$, $N_R = 4$ cochannel system at $\gamma = 18dB$.

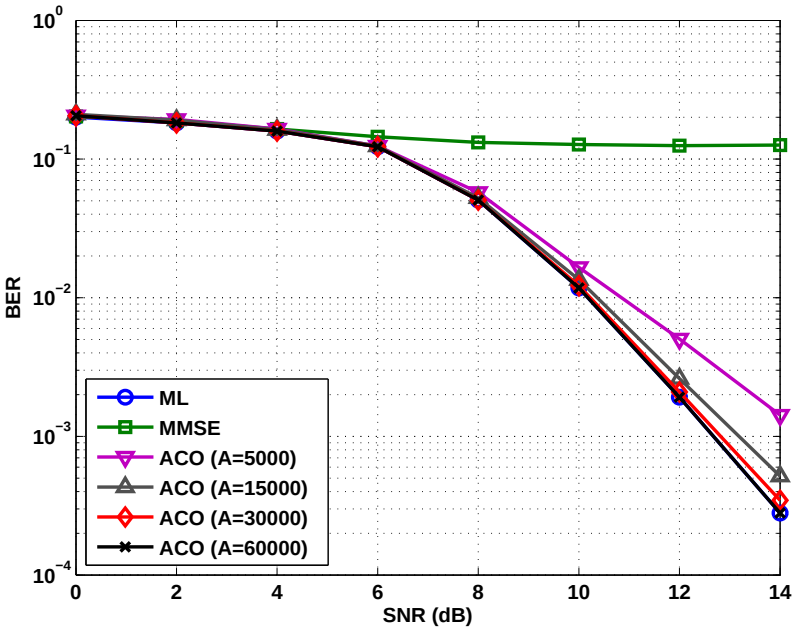

Figure 5.13 Performance of $N_T = 9$, $N_R = 6$ cochannel system employing QPSK modulation.

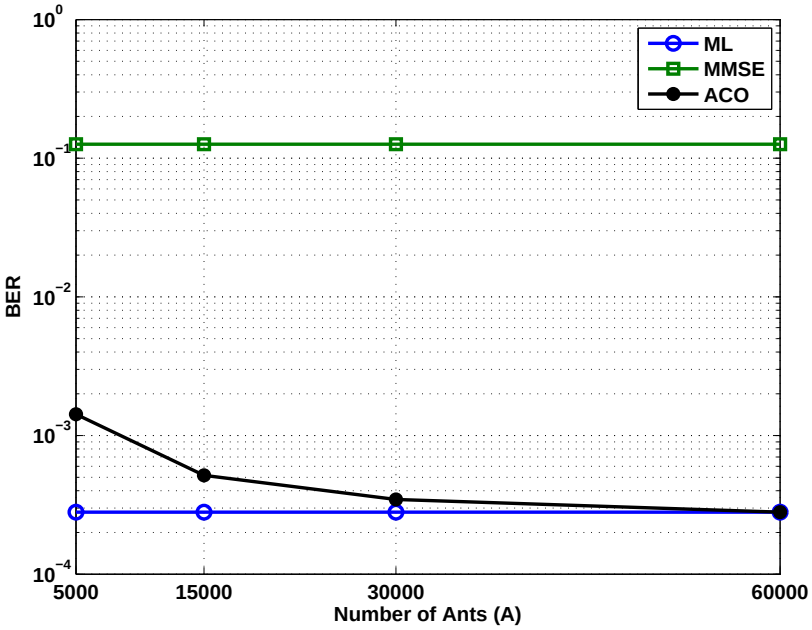

Figure 5.14 Performance of $N_T = 9$, $N_R = 6$ cochannel system at $\gamma = 14dB$.

$N_R = 2$ transmitted signals having a load factor of $L_F = 2.00$ is shown in Fig. 5.16 validating the performance of the developed detector. Percentage complexity comparison between the proposed scheme and the ML detector in terms of required number of real multiplications is provided in Table 5.6.

Table 5.6 Number of Real Multiplications ($N_T = 4$ in all cases)

	ACO Detector		ML Detector	% ML
N_R	A	$\mathcal{R}^{\times}$	$\mathcal{R}^{\times}$	
3	1000	$3.56E4$	$3.93E6$	0.91%
3	2000	$8.17E4$	$3.93E6$	2.08%
3	5000	$1.24E5$	$3.93E6$	3.16%
3	8000	$1.66E5$	$3.93E6$	4.22%
3	10000	$2.96E5$	$3.93E6$	7.52%
2	1000	$2.51E4$	$2.62E6$	0.96%
2	2000	$5.86E4$	$2.62E6$	2.23%
2	5000	$1.12E5$	$2.62E6$	4.28%
2	8000	$1.75E5$	$2.62E6$	6.69%
2	10000	$2.18E5$	$2.62E6$	8.30%

5.9 SUMMARY

In this chapter, an integrated framework comprising an ACO metaheuristic and a recursively defined ML search criteria is developed to tackle the signal detection problem for a multi-antenna cochannel communication system. Transmitted signals are considered to be chosen from a multilevel modulation. The receiver is capable of handling overloaded environments and achieves near-ML performance with significantly reduced complexity.

The proposed receiver consists of a preprocessing stage and a solution construction phase. The preprocessing stage obtains an MMSE estimate and performs a Cholesky factorization of the channel matrix. The solution construction phase is then started. It employs the ACO algorithm and the PMM metric to extract a list of most likely candidate solution vectors. Graph and tree based models of the proposed algorithm are presented and the ants adaptation strategy is discussed which plays a key role in the complexity reduction of the proposed algorithm. Due to the branch and bound

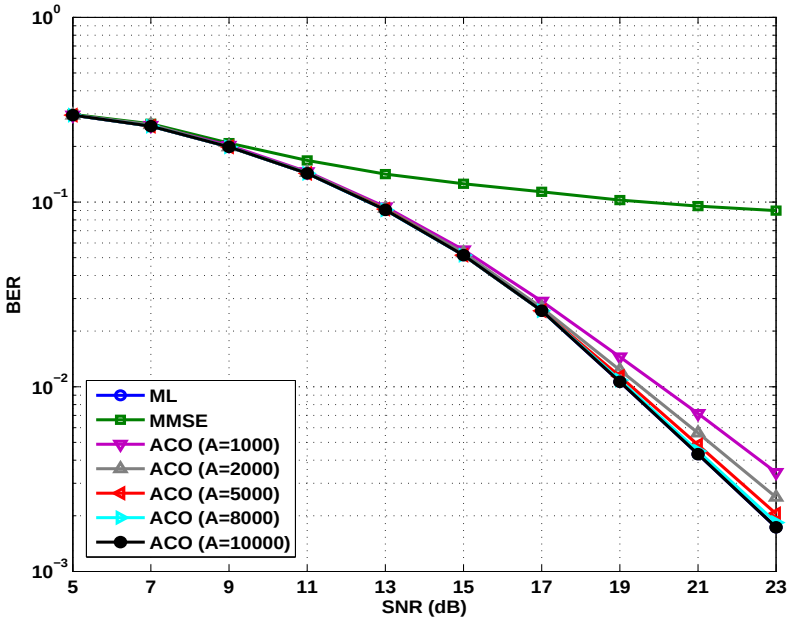

Figure 5.15 Performance of $N_T = 4$, $N_R = 3$ cochannel system employing 16-QAM modulation.

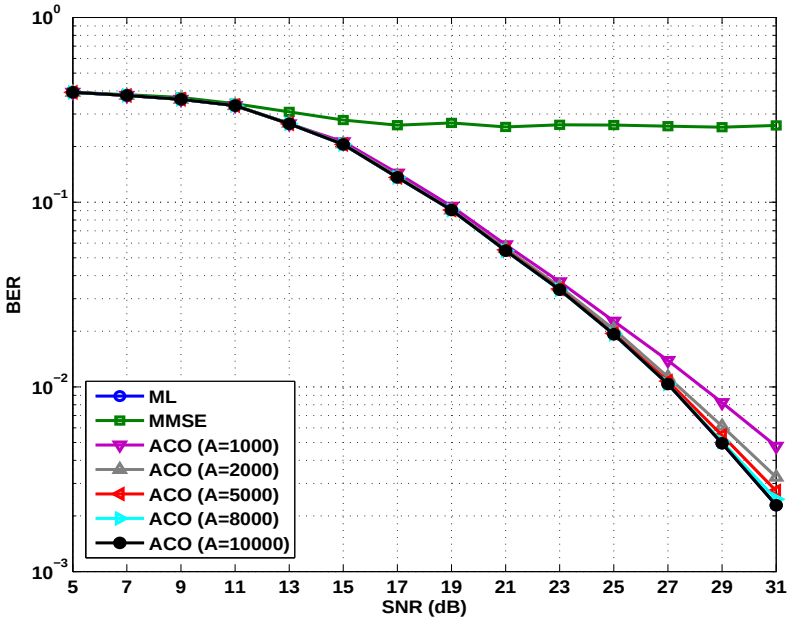

Figure 5.16 Performance of $N_T = 4$, $N_R = 2$ cochannel system employing 16-QAM modulation.

nature of the algorithm, an exact expression for the complexity cannot be determined. Instead, an upper bound on computational complexity is developed.

The performance is investigated for three different system settings, (i) $N_T = 8, = N_R = 5$ corresponding to $L_F = 1.60$ (ii) $N_T = 8, = N_R = 4$ corresponding to $L_F = 2.00$ and (iii) $N_T = 9, = N_R = 6$ corresponding to $L_F = 1.50$, all employing QPSK / 4-QAM signal constellations. It is observed that the proposed technique achieves near-ML performance in all these cases with significantly reduced computation.

A complete procedure to apply the proposed ACO detector to other square QAM constellations is also described. The performance results obtained through computer simulations for two systems with different load factors (i.e. $L_F = 1.60$ and $L_F = 2.00$) and employing 16-QAM signalling suggest that the detector is capable of achieving near-ML performance with a significantly reduced complexity.

Chapter 6

CONCLUSION AND FUTURE RESEARCH

6.1 INTRODUCTION

In this thesis, we have used ACO to develop low-complexity near-optimal detection techniques for multi-antenna cochannel communication systems. A scenario where multiple signals are simultaneously transmitted over Rayleigh fading channels is considered and the receiver is assumed to be operating in an overloaded environment.

The motivation behind the use of ACO algorithms is its successful application to multiuser detection in CDMA systems, which resulted in several low-complexity detection techniques offering performance close to that of optimum detection [50–52, 101, 102]. Recently, ACO has been used for near ML detection for MIMO systems [31] where the received MIMO signal is transformed into an equivalent real-valued representation[1] to formulate the ACO detection algorithm specifically for critically loaded and underloaded MIMO antenna system. However, the application of ACO to signal detection for overloaded conditions is of particular interest and requires special considerations.

This chapter is organized as follows: Section 6.2 concludes our work on low-complexity detection approaches and Section 6.3 outlines the proposed future research directions.

6.2 CONCLUSION

An iterative receiver structure for the detection of BPSK modulated cochannel signals transmitted over frequency-flat Rayleigh fading channels is developed in Chapter 4.

[1] The real-valued channel matrix is decomposed using QR factorization and an equivalent detection problem is formed.

The receiver is composed of two stages, namely preprocessing and iterative solution construction. In the preprocessing stage, an initial estimate of the transmitted signal is calculated from the MMSE detector which is used to build an ant moving-pattern table and to calculate the heuristic information concerning the problem. The iterative solution construction phase employs the ACO algorithm to extract a list of most likely candidate solution vectors during each ACO iteration. The ants walk through the moving-pattern table and probabilistically select one of the two components from each of its columns based on the associated pheromone strength and the heuristic value. Furthermore, the refined solution in each iteration is used to construct an updated moving-pattern table which is used by the ants to build candidate solutions for the next iteration. The calculation of new heuristic values is also based on the updated soft signal estimates. The performance of the proposed receiver suggests that it achieves performance approximately the same as ML detection in critically loaded cases ($L_F = 1$) with significantly reduced complexity. For the challenging case of overload, it still offers performance close to ML at low to moderate SNR. An error floor at high SNRs with increasing values of system load factor $L_F > 1$ is observed.

In Chapter 5, we developed an integrated framework comprising of ACO algorithm and a modified ML search criteria for signal detection in multi-antenna cochannel communication systems. The ML detection metric in Chapter 3 is transformed into an equivalent metric for fast recursive processing which also has an inherent tree representation. The modified computationally efficient ML metric is then integrated with the ACO algorithm for finding a near-optimal solution. The proposed receiver consists of preprocessing and solution construction stages. The preprocessing stage obtains an MMSE estimate and performs a Cholesky factorization of the channel matrix. The solution construction phase then uses the ACO algorithm and the PMM metric to search for better quality solutions. An upper bound on computational complexity is also developed. The simulation results for a $N_T = 8$, $N_R = 5$ system employing a QPSK / 4-QAM modulation shows that the proposed receiver is capable of approaching ML performance with complexity less than 10% of ML detector.

6.3 FUTURE RESEARCH DIRECTIONS

The low-complexity signal detection techniques based on ACO algorithms developed during the course of this research have motivated us to consider some more challenging and computationally expensive problems which are listed below.

6.3.1 Frequency-Selective Channels

During the literature study, we observed that ACO is not yet applied to the maximum likelihood sequence detection (MLSD) problem present in time-dispersive multipath channels. This motivates us to consider extending our developed detection scheme to more challenging case of frequency-selective channels in the future.

The optimum detection algorithm for frequency-selective systems is presented in [103, 104]. It finds the most likely sequence using dynamic programming approaches. The complexity for this scheme is linear in block length, but varies exponentially with the length of the channel and the number of bits per symbol. This suggest that the complexity is high for large channel length and high-level modulations (which are employed in most of the modern communication systems to gain high spectral efficiencies).

Another straight forward extension of the developed ACO algorithms is to consider the more computationally expensive problems when multiple cochannel signals are transmitted in the presence of ISI. The channel between each transmitted signal and the receive antenna is composed of multiple taps instead of a single complex gain. Complexity in this case grows exponentially with (i) number of transmitted signals, (ii) length of the channel and (iii) the number of bits per symbol. To make the problem more challenging, the receiver can be assumed to be operating under an overloaded setup.

6.3.2 Coded Systems

We propose to combine the developed ACO detector with forward error correction to formulate joint detection and decoding algorithms analogous to the list sphere decoder.

Comparisons with the list sphere decoder will be of most interest [105–108]. Furthermore, we intend to apply ACO for reducing the decoding complexity of capacity approaching low density parity check (LDPC) error correction codes. It is a well-known fact that maximum-likelihood decoding of LDPC codes is an NP-complete problem [109], which necessitates development of reduced complexity decoding approaches in this area.

6.3.3 Joint Channel Estimation and Data Detection

Multi-antenna wireless communication systems that only estimate the channel before actual information transmission begins, rely on the assumption that the channel remains constant for a considerable duration of time. On the contrary, if the channel fading process exhibit rapid variations, the training and channel tracking based approaches may become infeasible from a practical implementation perspective. An alternate approach in such scenarios is to perform joint channel estimation and signal detection. The optimal approach in this case has an exponential complexity in the problem dimension and thus requires low-complexity near-optimum alternatives. We therefore propose to apply the ACO algorithms to develop such reduced complexity approaches.

6.3.4 Antenna Selection

In certain MIMO systems, the number of receive chains N_C are less than the number of receive antennas N_R at the receiver ($N_C < N_R$). For such systems, the receiver scans the N_R receiver antennas and only selects a subset N_C of them for further processing. The optimal receive antenna selection requires the receiver to find a sub-matrix $\mathbf{H}_C$ (containing only N_C rows) of the $N_R \times N_T$ channel matrix $\mathbf{H}$ such that the channel capacity is maximized. A brute force solution is to evaluate the channel capacity expression for all the $N_C \times N_T$ sub-matrices of $\mathbf{H}$ and select the one which produces the maximum output. The complexity of this technique is however a bottleneck toward its practical implementation. Antenna selection is also employed in distributed and

cooperative systems to extend the performance of conventional systems. We therefore list antenna selection problems for future research.

6.3.5 Cooperative Diversity Systems

Finally, the idea of cooperative diversity can be employed to convert a severely overloaded system into a reasonably overloaded or a critically loaded system with the help of relays and then standard ACO based MUD algorithms can be applied. This idea is novel in itself as no effort is made in this direction yet.

Appendix A

THE GENETIC ALGORITHM FOR SIGNAL DETECTION IN OVERLOADED COCHANNEL SYSTEMS

A.1 INTRODUCTION

Due to a generic framework, Genetic Algorithm (GAs) are capable of solving a wide range of complex problems. Some of the applications include data mining [110], routing [111–113], artificial neural networks [114, 115] and multiuser detection in CDMA systems [43, 46, 116].

As outlined in [83], a typical GA must have the following components:

- An *initialization* step for creating some guessed solutions to the problem. Each guess is termed an *individual* while the collection of guesses is called a *population*.
- A mean for evaluating the goodness of the individuals / solutions produced during the optimization process, which is called the *fitness evaluation metric / function*.
- A method for mixing the portions of the good solutions to create new and better solutions, called the *crossover*.
- An operation to retain diversity in the produced solutions, referred to as *mutation*.
- A method to preserve the best known solutions, called *elitism*.

A.2 THE GENETIC ENGINE

Structurally, a GA employs an evolutionary processing engine which repeatedly carries out a series of steps for the given optimization problem until either convergence is

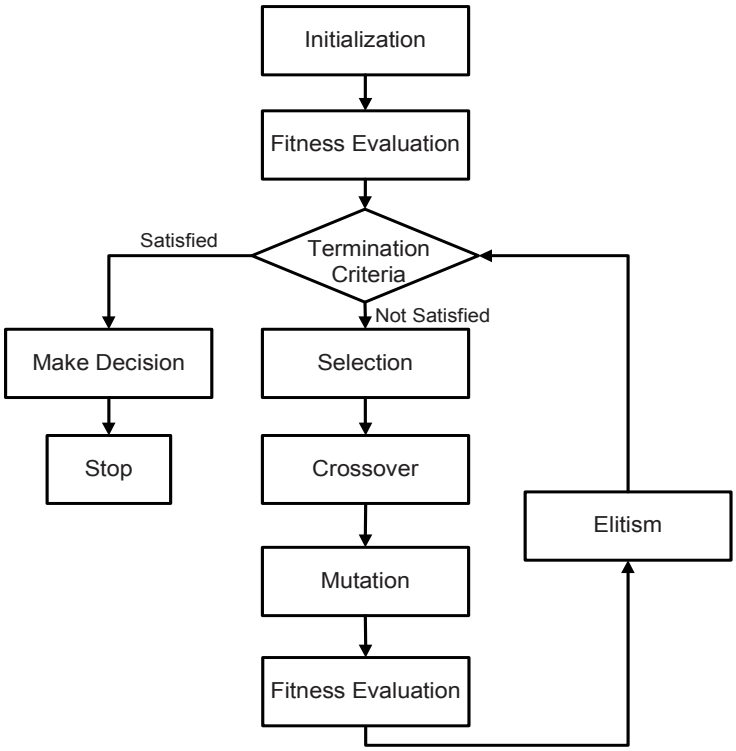

Figure A.1 Genetic Processing Engine.

reached or a termination criteria is met. A flowchart of the GA detector is shown in Fig. A.1 and each of the components is now briefly discussed.

A.2.1 Initialization

The algorithm begins by initializing a *population*, $\mathcal{P} = \left\{\bar{\mathbf{x}}^1, \bar{\mathbf{x}}^2 \ldots \bar{\mathbf{x}}^{P_T}\right\}$, comprised of P_T individuals each of which is a vector of length N_T, $\bar{\mathbf{x}}^p = \left[\bar{x}_1^p, \bar{x}_2^p \ldots \bar{x}_{N_T}^p\right]^T$, representing a candidate solution to the detection problem. The j-th component of the p-th individual, $\bar{x}_j^p$, is the estimate of the j-th transmitted signal[1] and can be expressed as a binary string of length bN_T called *alleles*, where $b = \log_2 |\chi|$, is the number of bits per symbol.

The initialized population can either be biased or unbiased. In the case of an unbiased initialization, the candidate solutions are picked randomly from the solution space. This helps in maximizing the solution diversity in the initial population. On the other hand, the initial population can also be biased by generating it through the mutation of

[1]Therefore it belongs to a particular signal modulation.

a seed. The seed is obtained with the help of an a priori guess concerning the optimum solution.

A.2.2 Fitness Evaluation

The fitness evaluation is performed using a *fitness metric* to calculate a *fitness value*[2] once after the initialization phase and then for every new generation formed after the mutation process as shown in Fig. A.1. The fitness value is an indicator of the goodness of the generated solutions. The fitness metric for signal detection is defined as

$$\Lambda_p\left(\bar{\mathbf{x}}^p\right) = \left\|\mathbf{y} - \mathbf{H}\bar{\mathbf{x}}^p\right\|^2, \qquad p = 1, 2 \ldots P_T, \tag{A.1}$$

which can also be written as (3.11)

$$\Lambda_p\left(\bar{\mathbf{x}}^p\right) = \sum_{i=1}^{N_R} \left| y_i - \sum_{j=1}^{N_T} h_{ij}\bar{x}_j^p \right|, \qquad p = 1, 2 \ldots . P_T. \tag{A.2}$$

The fitter candidate solutions correspond to smaller values of the fitness, $\Lambda_p\left(\bar{\mathbf{x}}^p\right)$.

A.2.3 Selection

After the completion of the fitness evaluation step, a set of P_T fitness values $\left\{\Lambda_p\left(\bar{\mathbf{x}}^p\right)\right\}_{p=1}^{P_T}$ is generated. These fitness values are first sorted in ascending order of magnitude and then a mating pool of size S_P is created by *selecting* the solutions corresponding to first S_P fitness values[3]. A key design challenge here is specifying a value of S_P that not only delivers sufficient diversity in the solutions to be created for the next generation, but can also expedite the convergence process.

A.2.4 Crossover

Once the best known individuals are chosen through selection, they serve as parents for producing offsprings. In *crossover*, a pair of parents is randomly selected and their

[2] This is done for each of the candidate solutions.

[3] Note that these fitness values correspond to the best S_P solutions in the current generation.

alleles are interchanged at randomly chosen locations to form two new individuals which belong to the new generation.

A.2.5 Mutation

The new individuals of the current generation just produced are subjected to *mutation* immediately after crossover. In mutation, the individuals are allowed to invert their alleles with a certain probability which is termed mutation probability, p_m. It is important to mention here that the mutation operation is critical to the success of GA's as it guarantees adequate diversity in the newly formed population which helps to avoid the algorithm getting trapped in a local minimum region.

A.2.6 Elitism

The crossover and mutation operations improve the average fitness of populations, but they do not guarantee preservation of good solutions having better fitness values from one generation to the next. To avoid this, some of the best known solutions in the current generation are taken to the next generation under the process of *elitism*.

A.2.7 Termination

The GA evaluates the *termination criteria* once the elitism process is over. It can either continue by returning to the selection step or it can stop execution if the termination criteria is met. Termination can occur either when the total number of generations G_T are processed or convergence has been reached.

Appendix B

A TUTORIAL ON ANT-COLONY OPTIMIZATION

B.1 INTRODUCTION

The existence of ants on earth dates back to around 100 million years ago with an approximated population of 10 thousand trillion [117, 118]. Ants are social insects living in the form of colonies with sizes ranging from 30 individuals to millions of individuals per colony. Apparently simple insects, but their tendency to behave intelligently when in groups have captured the attention of researchers and scientists to investigate and better understand their working mechanisms. Some of the complex collective behaviors exhibited by the ant colonies include their foraging ability, division of labor, cooperative support and self assembly [117]. It is interesting to note that most of their functional matters are self-organized without any central control. The ants use an indirect form of communication, called *stigmergy*, through which they communicate with each other by modifying their surroundings (e.g. a foraging ant deposits a chemical on the path it follows). Upon sensing the environment, each individual performs a simple probabilistic action (which is biased by the chemical deposited on the ground in the case of a foraging ant). When observed collectively over a certain time duration these actions appear highly structured with some interesting outcomes (e.g. formation of the shortest path between the nest and the food source).

This idea of self-organization and distributed control for efficiently solving real-world problems in ants lead researchers to develop artificial ant colony based algorithms to solve mathematical optimization problems. For this purpose, artificial agents are developed having capabilities of performing some basic random operations along with

artificial stigmergy to facilitate communication among them.

This chapter presents fundamental concepts and terminologies associated with ant colony optimization (ACO). The reader should be aware of the fact that the material presented in the remainder of this chapter make frequent use of [86] which serves as a more comprehensive treatment of the ACO metaheuristic. Section B.2 discusses the foraging ability of natural ants with the help of an experimental setup. The transformation from real to artificial ants for optimization purposes is also introduced. In Section B.3, the notion of combinatorial optimization is introduced and its importance is explained. An ACO metaheuristic is explained and the associated framework is briefly discussed in Section B.4. Ant colony based mathematical optimization algorithms are specified in Section B.5. Finally, Section B.6 presents the chapter summary.

B.2 REAL ANTS AND THE SHORTEST PATH PROBLEM

The visual ability of many of the real ant species is so weak that they can be classified as either completely blind or semiblind [86]. Earlier investigations into the ants' behavior have revealed the fact that the communication between them is accomplished by secreting certain chemicals, called *pheromones*, into the environment. Ants have a built in ability to sense these pheromone and thus can extract the contained message which can be [119] (i) an alert to other ants about a potential danger, (ii) an assembly call, (iii) identifying ants, (iv) marking the colony's boundary, (v) marking the path followed by ants from nest to the food source and (vi) calling for help when in need. Among all these, the foraging behavior using which ants are able to efficiently find a shortest path between the nest and the food source is of special importance from the optimization perspective.

To study the foraging behavior of ants, controlled experiments were designed and conducted in [120, 121]. Two different paths were made available to the ants[1] to move between the ant nest and the food source. The length of one of the two available paths was changed each time the experiment was performed. Initially, the ants releasing

[1]To be specific, the behavior of Argentine ant species *Iridomyrmex humilis* was investigated in these experiments.

from the nest randomly chose one of the two alternate paths to move in the direction of food. However, over a period of time, an interesting behavior (where all the ants employed a common pheromone trail) was observed which lead in the formation of a shortest path by the ants between the nest and food source. The experiments not only successfully demonstrated the intelligent behavior of ants but also helped in developing a theoretical model (based on transition probabilities) governing the ants' movement and path construction strategy. With the help of computer simulations, the developed probabilistic model was shown to be similar to the mechanism followed by the real ants.

This provided motivation to develop ant colony based algorithms to solve various engineering optimization problems.

B.3 COMBINATORIAL OPTIMIZATION

Combinatorial optimization (CO) field is concerned with finding an optimal object (from a finite set of all the available objects) by evaluating the fitness of each of the object using a known objective function. Most of the CO problems arising in the field of engineering and sciences are nondeterministic polynomial-time hard (NP-hard), i.e. the optimal solution cannot be found within polynomially bounded computational time [88]. It is important to note that the MUD problem we are dealing in this thesis requires an exhaustive search in a $|\chi|^{N_T}$ space to find an optimal solution and thus cannot be completed in polynomial time i.e. the combinatorial problem of finding the most likely transmitted signal vector is NP-hard in the number of transmitted signals [122].

To solve such NP-hard optimization problems in real time, problem-specific methods are often used which are able to find reasonably better suboptimal solutions in less time. The algorithms based on this methodology are often classified as *heuristics*. In contrast, a relatively newer category of optimization algorithms known as *metaheuristics* have emerged over the past couple of decades as a simple yet effective alternative. A metaheuristic is a combination of concepts and rules which can solve a wide range of optimization problems by exploiting the underlying heuristics of the problem. Metaheuristic optimization techniques have better likelihood of finding high-quality solutions

to hard problems in a time-efficient manner [86].

B.4 THE ACO METAHEURISTIC

Earlier investigations and theoretical findings on behavioral aspects of social insects lead to the development of some powerful metaheuristic optimization techniques. Among many of those, ant colony optimization (ACO) have emerged as a popular CO optimization technique. ACO is based on the collective social behavior which follows the survival of whole colony of ants rather than survival of an individual.

In ACO, the notion of artificial ants along with the artificial stigmergic model is designed to tackle a given optimization problem. An artificial ant or simply an ant is modeled as a computational entity or agent which probabilistically creates a solution to the problem. To construct a solution, the metaheuristic uses incremental strategy where solution components[2] are added one after the other to a partially constructed solution. Before applying the ACO metaheuristic, it is important to understand the available heuristics and represent the problem in a consistent form.

With every minimization or maximization problem $\mathcal{P}$, there is an associated triplet $(\mathcal{S}, f, \Omega)$, where

- $\mathcal{S}$ is the set of candidate solutions.
- f is the objective function that assigns a cost or fitness value to each of the candidate solution in $\mathcal{S}$.
- Ω is a set of constraints.

This information is used to specify an appropriate expression for the heuristic information, pheromone update rule and the transition probability.

A generic framework of ACO metaheuristic is shown in Fig. B.1. A set of solution components and pheromone model are derived from the given CO problem. The ACO engine repeatedly performs the solution construction process to build solutions and evaluate their fitness values. The pheromone update task modifies the pheromone

[2]Note that each of the solution component is chosen from the entire set of available options.

depending on the fitness values of the constructed solution. This iterative process stops when the termination criteria has reached.

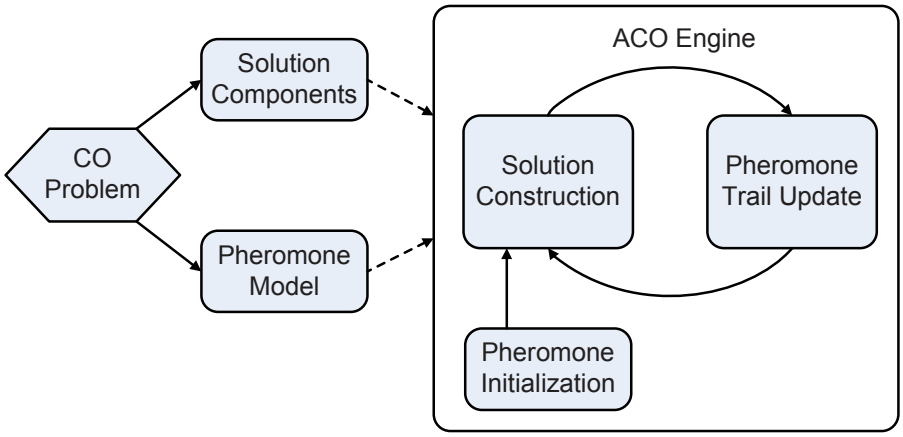

Figure B.1 Algorithmic framework of the ACO metaheuristic [13].

B.5 ANT BASED ALGORITHMS

The general procedure shown in Fig. B.1 specifies the ACO metaheuristic without providing the detailed internal working mechanism of the ACO engine which includes the specification of the probabilistic action choice rule and the pheromone update mechanism. In general, these details vary among the different ACO algorithms. To be specific, some of the main ACO algorithms which are used in MUD problem[3] are:

- Ant System (AS).
- MAX-MIN Ant System (MMAS).
- Ant Colony System (ACS).

The success of AS algorithm [95, 123], which is the first ant-based optimization algorithm provided motivation to the researchers to further explore and investigate the area. As a result, new and refined MMAS and ACS algorithms [96, 124–126] emerged based on the AS algorithm with certain modifications. For a detailed and comprehensive coverage of these algorithms, the reader is referred to [86].

[3]Note that these algorithms were originally proposed for the traveling salesman problem (TSP).

B.6 SUMMARY

The chapter began with a demonstration of the foraging behavior of natural ants with the help of a set of experiments. The probabilistic model derived from these experiments provided motivations for developing the notion of artificial ants and ant based optimization algorithms. The notion of NP-hard CO problem is explained and the need of devising low complexity suboptimum algorithms is emphasized. As a potential candidate, ACO metaheuritic approach is introduced. In the end, some of the ACO based algorithms are specified which include AS, MMAS and ACS.

REFERENCES

[1] J. Winters. Optimum combining in digital mobile radio with cochannel interference. *IEEE J. Sel. Areas Commun.*, 2(4):528–539, Jul. 1984.

[2] E.A. Fain and M.K. Varanasi. Diversity order gain for narrow-band multiuser communications with pre-combining group detection. *IEEE Trans. Commun.*, 48(4):533–536, Apr. 2000.

[3] M.O. Damen, K. Abed-Meraim, and J.-C. Belfiore. Generalised sphere decoder for asymmetrical space-time communication architecture. *Electron. Lett.*, 36(2):166–167, Jan. 2000.

[4] M. O. Damen, H. El Gamal, and G. Caire. On maximum-likelihood detection and the search for the closest lattice point. *IEEE Trans. Inf. Theory*, 49(10):2389–2402, 2003.

[5] P. Dayal and M.K. Varanasi. A fast generalized sphere decoder for optimum decoding of under-determined MIMO systems. In *Proc. of 41st Annual Allerton Conf. on Commun., Control, and Computing*, Oct. 2003.

[6] G. W. K. Colman and T. Willink. Genetic algorithm assisted array processing in overloaded systems. In *Proc. of IEEE 60th Veh. Tech. Conf., (VTC)*, volume 1, pages 261 – 265, 2004.

[7] B.W. Zarikoff. Investigation of an iterative groupwise soft input/output multiuser detection algorithm. Master's thesis, Simon Fraser University, June 2004.

[8] T. Cui and C. Tellambura. An efficient generalized sphere decoder for rank-deficient MIMO systems. *IEEE Commun. Lett.*, 9:423–425, May 2005.

[9] B.W. Zarikoff, J.K. Cavers, and S. Bavarian. An Iterative Groupwise Multiuser Detector for Overloaded MIMO Applications. *IEEE Trans. Wireless Commun.*, 6(2), Feb. 2007.

[10] G. Colman and T.J. Willink. Overloaded array processing using genetic algorithms with soft-biased initialization. *IEEE Trans. Veh. Tech.*, 57(4):2123–2131, Jul. 2008.

[11] M. Krause, D.P. Taylor, and P.A. Martin. List-based group-wise symbol detection for multiple signal communications. *IEEE Trans. Wireless Commun.*, 10(5):1636–1644, May 2011.

[12] John G. Proakis and Masoud Salehi. *Digital Communications.* McGraw-Hill, 2008.

[13] Christian Blum and Daniel Merkle. *Swarm Intelligence: Introduction and Applications.* Springer-Verlag, 2008.

[14] Z. Ding B. P. Lathi. *Modern Digital and Analog Communications Systems.* Oxford University Press, 2010.

[15] C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, Oct. 1948.

[16] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multiple antennas. *Wireless Personal Communications*, pages 311–335, 1998.

[17] B. Vucetic and J. Yuan. *Space-time Coding.* John Wiley & Sons Ltd, 2003.

[18] Arogyaswami Paulraj, Rohit Nabar, and Dhananjay Gore. *Introduction to Space-Time Wireless Communications.* Cambridge University Press, 2003.

[19] J. Winters. On the capacity of radio communication systems with diversity in a rayleigh fading environment. *IEEE J. Sel. Areas Commun.*, 5(5):871–878, Jun 1987.

[20] J. Winters, J. Salz, and R.D. Gitlin. The impact of antenna diversity on the capacity of wireless communication systems. *IEEE Trans. Commun.*, 42(234):1740–1751, Feb/Mar/Apr 1994.

[21] I.E. Telatar. Capacity of multiantenna gaussian channels. *European Transactions Telecommun.*, pages 585–595, 1999.

[22] J. Winters. Optimum combining for indoor radio systems with multiple users. *IEEE Trans. Commun.*, 35(11):1222–1230, Nov. 1987.

[23] S.J. Grant and J.K. Cavers. Performance enhancement through joint detection of cochannel signals using diversity arrays. *IEEE Trans. Commun.*, 46(8):1038–1049, Aug. 1998.

[24] S. Verdu. *Multiuser Detection*. Cambridge University Press, 1998.

[25] G.D. Golden, G.J. Foschini, R.A. Valenzuela, and P.W. Wolniansky. Detection Algorithm and Initial Laboratory Results using V-BLAST Space-Time Communications Architecture. *IEE Electronic Lett.*, 35:14–16, Jan. 1999.

[26] Simon O. Haykin. *Adaptive Filter Theory*. Prentice Hall, 4 edition, 2002.

[27] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger. Closest point search in lattices. *IEEE Trans. Inf. Theory*, 48:2201–2214, Aug. 2002.

[28] R. Bohnke, D. Wubben, V. Kuhn, and K.-D. Kammeyer. Reduced complexity mmse detection for blast architectures. In *Proc of. IEEE Global Telecommun. Conf. (GLOBECOM)*, volume 4, pages 2258–2262, Dec. 2003.

[29] D. Pham, K.R. Pattipati, P.K. Willett, and J. Luo. A generalized probabilistic data association detector for multiple antenna systems. *IEEE Commun. Lett.*, 8(4):205–207, Apr. 2004.

[30] B. Hassibi and H. Vikalo. On the sphere-decoding algorithm i. expected complexity. *IEEE Trans. Signal Process.*, 53(8):2806–2818, Aug. 2005.

[31] J.-K. Lain and Jyun-Yu Chen. Near-MLD MIMO Detection Based on a Modified Ant Colony Optimization. *IEEE Commun. Lett.*, 14(8):722–724, Aug. 2010.

[32] Cheng-Yu Hung and Wei-Ho Chung. An improved mmse-based mimo detection using low-complexity constellation search. In *Proc of. IEEE Global Telecommun. Conf. (GLOBECOM) Workshops*, pages 746 –750, Dec. 2010.

[33] C.L. Miller, D.P. Taylor, and P.T. Gough. Estimation of Co-Channel Signals with Linear Complexity. *IEEE Trans. Commun.*, 49(11), Nov. 2001.

[34] H. Vikalo and B. Hassibi. Maximum-Likelihood Sequence Detection of Multiple Antenna Systems over Dispersive Channels via Sphere Decoding. *EURASIP J. Applied Signal Process*, pages 525–531, May 2002.

[35] X. Zhu and R.D. Murch. Layered Space-Time Equalization For Wireless MIMO Systems. *IEEE Trans. Wireless Commun.*, 2(6), Nov. 2003.

[36] J. Jalden and B. Ottersten. On the Complexity of Sphere Decoding in Digital Communications. *IEEE Trans. Signal Process.*, 53(4), Apr. 2005.

[37] Y. Lee and W.-R. Wu. Adaptive Channel Aided Decision Feedback Equalisation for SISO and MIMO Systems. *IEE Commun. Proceedings*, 153(5), Oct. 2006.

[38] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. JHU Press, 1996.

[39] Harold Artes, D. Seethaler, and F. Hlawatsch. Efficient detection algorithms for mimo channels: A geometrical approach to approximate ml detection. *IEEE Trans. Signal Process.*, 51(11):2808–2820, 2003.

[40] J. Hicks, S. Bayram, W.H. Tranter, R.J. Boyle, and J. H. Reed. Overloaded array processing with spatially reduced search joint detection. *IEEE J. Sel. Areas Commun.*, 19(8):1584–1593, Aug. 2001.

[41] H. Bolcskei, D. Gesbert, and C. Papadias. *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge University Press, 2005.

[42] U. Fincke and M. Pohst. Improved Methods for Calculating Vectors of Short Length in a Lattice, including a Complexity Analysis. *Mathematics of Computation*, 44(11):463–471, Apr. 1985.

[43] K. Yen and L. Hanzo. Genetic-algorithm-assisted multiuser detection in asynchronous cdma communications. *IEEE Trans. Veh. Tech.*, 53(5):1413–1422, Sept. 2004.

[44] K. Yen and L. Hanzo. Hybrid genetic algorithm based detection schemes for synchronous cdma systems. In *Proc. of IEEE 51st Veh. Tech. Conf. (VTC), 2000-Spring*, volume 2, pages 1400 –1404, 2000.

[45] M.J. Juntti, T. Schlosser, and J.O. Lilleberg. Genetic algorithms for multiuser detection in synchronous cdma. In *Proc. of IEEE Intl. Symp. on Inf. Theory (ISIT)*, page 492, Jun-Jul 1997.

[46] C. Ergun and K. Hacioglu. Multiuser detection using a genetic algorithm in cdma communications systems. *IEEE Trans. Commun.*, 48(8):1374–1383, Aug. 2000.

[47] K.K. Soo, Y.M. Siu, W.S. Chan, L. Yang, and R.S. Chen. Particle-swarm-optimization-based multiuser detector for cdma communications. *IEEE Trans. Veh. Tech.*, 56(5):3006–3013, Sept. 2007.

[48] H.H. El-Mora, Sheikh A.U., and A. Zerguine. Application of particle swarm optimization algorithm to multiuser detection in cdma. In *Proc. of IEEE 16th Intl. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, volume 4, pages 2522 –2526 Vol. 4, Sept. 2005.

[49] Xu Binbin. Particle-swarm-optimization based turbo multiuser detection for stbc mc-cdma systems. In *Proc. of IEEE 19th Intl. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, pages 1–5, Sept. 2008.

[50] S.L. Hijazi and B. Natarajan. Novel Low-Complexity DS-CDMA Multiuser Detector based on Ant Colony Optimization. Sept. 2007.

[51] J.-K. Lain and J.-J. Lai. Ant Colony Optimisation-Based Multiuser Detection for Direct-Sequence CDMA Systems with Diversity Reception. *IET Commun.*, 1(4):556–561, 2007.

[52] R.G. Xu, L.-L. Yang Maunder, and L. Hanzo. Near-Optimum Multiuser Detectors Using Soft-Output Ant-Colony-Optimization for the DS-CDMA Uplink. *IEEE Signal Process. Lett.*, 16(2), Feb. 2009.

[53] International Telecommunication Union. Measuring the information society 2012. 2012.

[54] B. Sklar. *Digital Communications: Fundamentals and Applications.* Prentice Hall, 2 edition, 2001.

[55] William C. Jakes. *Microwave Mobile Communications.* Wiley-IEEE Press, 1994.

[56] T.S. Rappaport. *Wireless Communications: Priciples and Practice.* Prentice Hall, 2002.

[57] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.

[58] A. Leon-Garcia. *Probability and Random Processes for Electrical Engineers.* Addison-Wesley, 1994.

[59] M.K. Simon and M.-S. Alouini. *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis.* John Wiley & Sons, 2000.

[60] D. G. Brennan. Linear diversity combining techniques. *Proc. IRE*, 47(6):1075–1102, 1959.

[61] A. Goldsmith. *Wireless Communications.* Cambridge University Press, 2005.

[62] Andreas F. Molisch. *Wireless Communications.* IEEE Press and John Wiley & Sons Ltd, 2006.

[63] Vijay K. Garg and Joseph E. Wilkes. *Principles and Applications of GSM.* Prentice Hall, Upper Saddle River, NJ, USA, 1998.

[64] Vijay K. Garg. *IS-95 CDMA and CDMA2000: Cellular/PCS Systems Implementation.* Prentice Hall, Upper Saddle River, NJ, USA, 1999.

[65] D.P. Taylor. *Signal Space Analysis, Lecture Notes Commun. Engineering II.* University of Canterbury, Christchurch, New Zealand, 2012.

[66] E. Biglieri, R. Calderbank, A. Constantinides, A. Goldsmith, A. Paulraj, and H.V. Poor. *MIMO Wireless Communications.* Cambridge University Press, 2007.

[67] A. van Zelst. *MIMO OFDM for Wireless LANs.* PhD thesis, Department of Electrical Engineering, Eindhoven University of Technology, April 2004.

[68] C.P. Schnorr and M. Euchner. Lattice basis reduction: Improved practical algorithms and solving subset sum problems. In *Math. Programming*, pages 181–191, 1993.

[69] E. Viterbo and J. Boutros. A universal lattice code decoder for fading channels. *IEEE Trans. Inf. Theory*, 45:1639–1642, Jul. 2000.

[70] H. Vikalo. *Sphere decoding algorithms for digital communications.* PhD thesis, Stanford University, 2003.

[71] P.W. Wolniansky, G.J. Foschini, G.D. Golden, and R.A. Valenzuela. V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel. In *Proc. of URSI Intl. Symp. on Signals, Systems, and Electronics*, pages 295 –300, Sept. - Oct. 1998.

[72] S.W. Kim and P.W. Kim. Log-likelihood-ratio-based detection ordering in v-blast. *IEEE Trans. Commun.*, 54(1), Jan. 2006.

[73] Y.J. Chun and S.W. Kim. Log-likelihood-ratio ordered successive interference cancellation in multi-user, multi-mode mimo systems. *IEEE Commun. Lett.*, 12(11):837–839, Nov. 2008.

[74] Sang-Rim Lee and Inkyu Lee. Optimal detection ordering for v-blast. In *IEEE Intl. Conf. on Commun. (ICC)*, pages 5288 –5293, Jun. 2007.

[75] M.K. Varanasi. Group detection for synchronous cdma communication over frequency-selective rayleigh fading channels. pages 849–858, Sep. 1993.

[76] M.K. Varanasi. Group detection for synchronous gaussian code-division multiple-access channels. *IEEE Trans. Inf. Theory*, 41(4):1083–1096, Jul. 1995.

[77] M.K. Varanasi. Parallel group detection for synchronous cdma communication over frequency-selective rayleigh fading channels. *IEEE Trans. Inf. Theory*, 42(1):116–128, Jan 1996.

[78] John H. Holland. *Adaptation in Natural and Artificial Systems, Ann Arbor, MI.* The University of Michigan Press, 1975.

[79] K. De Jong. A 10 year perspective. In *Proc. of Intl. Conf. on Genetic Algorithms and their Applications*, pages 169–177, 1985.

[80] K. De Jong. Learning with the genetic algorithm: An overview. *Machine Learning*, 3:121–137, Oct. 1988.

[81] David E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, 1989.

[82] Melanie Mitchel. *An Introduction to Genetic Algorithms.* The MIT Press, 1996.

[83] David A. Coley. *An Introduction to Genetic Algorithms for Scientists and Engineers.* World Scientific Publishing, 1999.

[84] K. Zyczkowski and H. Sommers. Truncations of random unitary matrices. *J. Phys. M, Math. Gen.*, 33(10):20452057, Mar. 2000.

[85] M.L. Mehta. *Random Matrices and Statistical Theory.* New York: Academic, 1967.

[86] M. Dorigo and T. Sttzle. *Ant Colony Optimization.* MIT Press, 2004.

[87] K.T. Tasneem, P.A. Martin, and D.P. Taylor. Iterative soft detection of cochannel signals using ant colony optimization. In *IEEE 23rd Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC).*, 2012.

[88] M.R. Gary and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness.* Freeman, San Francisco, 1979.

[89] M. Dorigo, V. Maniezzo, and A. Colorni. Ant system: optimization by a colony of cooperating agents. *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, 26(1):29–41, Feb. 1996.

[90] Marco Dorigo and Thomas Sttzle. The ant colony optimization metaheuristic: Algorithms, applications, and advances. In *Handbook of Metaheuristics*, pages 251–285. Kluwer Academic Publishers, 2002.

[91] Todd K. Moon and Wynn C. Stirling. *Mathematical Methods and Algorithms for Signal Processing*. Prentice Hall, Upper Saddle River, 2005.

[92] P. Robertson, E. Villebrun, and P. Hoeher. A comparison of optimal and sub-optimal map decoding algorithms operating in the log domain. In *IEEE Intl. Conf. on Commun. (ICC) 'Gateway to Globalization'*, volume 2, pages 1009 – 1013 vol.2, June 1995.

[93] H.V. Poor and S. Verdu. Probability of error in mmse multiuser detection. *IEEE Trans. Inf. Theory*, 43(3):858–871, May.

[94] Jin-Sung Kim, Sung-Hyun Moon, and Inkyu Lee. A new reduced complexity ml detection scheme for mimo systems. *IEEE Trans. Commun.*, 58(4):1302–1310, April.

[95] M. Dorigo, V. Maniezzo, and A. Colorni. Positive feedback as a search strategy. Technical report, Dipartimento di Elettronica, Politecnico di Milano, Milan, 1991.

[96] T. Stutzle and H.H. Hoos. Max-min ant system. *Future generation computer systems*, 16(8):889–914, 2000.

[97] Jos Akhtman and Lajos Hanzo. Reduced-complexity maximum-likelihood detection in multiple-antenna-aided multicarrier systems. In *5th Intl. Workshop on Multi-Carrier Spread Spectrum, Oberpfaffenhofen, Germany, Springer.*, 2005.

[98] B. Tarokh and H.R. Sadjadpour. Construction of OFDM M-QAM sequences with low peak-to-average power ratio. *IEEE Trans. Commun.*, 51(1):25 – 28, Jan. 2003.

[99] Tao Cui and C. Tellambura. Approximate ml detection for mimo systems using multistage sphere decoding. In *Proc. of Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, volume 1, pages 1054–1056, 2004.

[100] Tao Cui and C. Tellambura. Approximate ml detection for mimo systems using multistage sphere decoding. *IEEE Signal Process. Lett.*, 12(3):222–225, 2005.

[101] S.L. Hijazi and B. Natarajan. Near-Optimal Multiuser Detection in Asynchronous MC-CDMA via the Ant Colony Approach. 2007.

[102] C. Xu, B. Hu, L.-L. Yang, and L. Hanzo. Ant-Colony-Based Multiuser Detection for Multifunctional-Antenna-Array-Assisted MC DS-CDMA Systems. *IEEE Trans. Veh. Tech.*, 57(1), Jan. 2008.

[103] A.J. Viterbi. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Inf. Theory*, 13(2):260–269, 1967.

[104] G.D. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Trans. Inf. Theory*, 18(3):363–378, 1972.

[105] B.M. Hochwald and S. ten Brink. Achieving near-capacity on a multiple-antenna channel. *IEEE Trans. Commun.*, 51(3):389–399, Mar. 2003.

[106] H. Vikalo, B. Hassibi, and T. Kailath. Iterative decoding for mimo channels via modified sphere decoding. *IEEE Trans. Wireless Commun.*, 3(6):2299–2311, Nov. 2004.

[107] P.A. Martin and D.P. Taylor. High-throughput error correcting space-time block codes. *IEEE Commun. Lett.*, 8(7):458–460, Jul. 2004.

[108] Yang Sun and Joseph R. Cavallaro. High throughput vlsi architecture for soft-output mimo detection based on a greedy graph algorithm. In *Proc. of 19th ACM Great Lakes Symp. on VLSI (GLSVLSI)*, pages 445–450, New York, NY, USA, 2009. ACM.

[109] Todd K. Moon. *Error Correction Coding: Mathematical Methods and Algorithms.* John Wliey & Sons Ltd., 2005.

[110] T-P Hong, C-H Chen, Y-C Li, and Y-L Wu. Genetic-fuzzy data mining with divide and conquer strategy. *IEEE Trans. Evol. Comp.*, 12(2):252–265, Apr. 2008.

[111] Y. Leung, G. Li, and Z-B Xu. A genetic algorithm for the multiple destination routing problems. *IEEE Trans. Evol. Comp.*, 2(4):150–161, Dec. 1998.

[112] C.W. Ahn and R.S. Ramakrishna. A genetic algorithm for shortest path routing problem and the sizing of populations. *IEEE Trans. Evol. Comp.*, 6(6):556–579, Dec. 2002.

[113] B. Lorenzo and S. Glisic. Optimal routing and traffic scheduling for multihop cellular networks using genetic algorithm. *IEEE Trans. Mobile Comp.*, Oct. 2012.

[114] V. Maniezzo. Genetic evolution of the topology and the weight distribution of neural networks. *IEEE Trans. Neural Networks*, 5:39–53, Jan. 1994.

[115] F. H F Leung, H.K. Lam, S.H. Ling, and P.K.-S. Tam. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Trans. Neural Networks*, 14(1):79–88, Jan. 2003.

[116] S. Abedi and R. Tafazolli. Genetically modified multiuser detection for code division multiple access systems. *IEEE J. Sel. Areas Commun.*, 20(2):463–473, Feb 2002.

[117] A.P. Engelbrecht. *Fundamentals of Computational Swarm Intelligence.* John Wliey & Sons Ltd., 2005.

[118] B. Holldobler and Edward O. Wilson. *Journey to the Ants: A Story of Scientific Exploration.* Belknap Press of Harvard University Press, 1994.

[119] R.K. Vander Meer, M.D. Breed, M.L. Winston, and K.E. Espelie. *Pheromone Communication in Social Insects: Ants, Wasps, Bees, and Termites.* Boulder, CO: Westview Press, 1997.

[120] S. Goss, S. Aron, J.L. Deneubourg, and J. Pasteels. Self-organized shortcuts in the argentine ant. *Naturwissenschaften*, 76(12):579–581, 1989.

[121] J. Deneubourg, S. Aron, S. Goss, and J. Pasteels. The self-organizing exploratory pattern of the argentine ant. *J. Insect Behavior*, 3:159–168, 1990.

[122] S. Verdu. Computational complexity of optimum multiuser detection. *Algorithmica*, 4:303–312, 1989.

[123] A. Colorni, M. Dorigo, and V. Maniezzo. Distributed optimization by ant colonies. In *Proc. of European Conf. on Artificial Life*, pages 134–142, Cambridge, MA, 1992. MIT Press.

[124] T. Stutzle and H.H. Hoos. Max-min ant system and local search for the traveling salesman problem. In *Proc. of IEEE Intl. Conf. on Evolutionary Comput.*, pages 309–314, Apr. 1997.

[125] M. Dorigo and L.M. Gambardella. Ant colonies for the traveling salesman problem. *BioSystems*, 43(2):73–81, 1997.

[126] M. Dorigo and L.M. Gambardella. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evolutionary Comput.*, 1(1):53–66, Apr.