

ANALYSIS AND DESIGN OF EFFICIENT WIRELESS NETWORKS

Hua Wang

COMMUNICATIONS & SIGNAL PROCESSING LABORATORY
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109-2122

December, 2004

Technical Report No. 355
Approved for public release; distribution unlimited.

ABSTRACT

ANALYSIS AND DESIGN OF EFFICIENT WIRELESS NETWORKS

by
Hua Wang

Chairman: Wayne E. Stark

We present a generic integrated design methodology that is suitable for many kinds of mobile systems. The integrated design methodology takes into account the coupling among the subsystems and simultaneously optimizes their operation under an energy constraint. We show that significant improvement in performance can be achieved by using the integrated design methodology compared with traditional design methodologies. We evaluate the tradeoff between energy consumption and performance for several network scenarios.

Routing is an efficient method for connectivity and low energy consumption of wireless networks. When each node is equipped with an omni-directional antenna, a point-to-multipoint connection is often available for routing. When the design goal is to minimize the maximum power consumed by the nodes in a network, we provide an algorithm with polynomial-time complexity that assigns power to each node for unicast, broadcast, and multicast sessions. When the design goal is to minimize the total power consumed by the nodes in a network, we provide an algorithm with

polynomial-time complexity that assigns power to each node for a unicast session and show that the computational complexity of routing algorithms for broadcast and multicast sessions is NP-hard.

We introduce transport efficiency to capture both bandwidth efficiency and energy efficiency of wireless networks. We show that for linear networks the optimal transport efficiency is inversely proportional to the end-to-end distance for one physical layer model and observe through numerical results that the same is true for many other physical layer models. We investigate the interference caused by space-time coding and an ordinary end-fire antenna array to neighboring networks. We show that the ordinary end-fire antenna array gives higher transport efficiency than space-time coding when the number of receiving antennae is small and space-time coding gives higher transport efficiency than the ordinary end-fire antenna array when the number of receiving antennae is large. We indicate that cooperative communication between linear networks can improve transport efficiency, but it gives marginal benefit if the cooperating networks are separated too far apart.

© Hua Wang 2004
All Rights Reserved

To my wife TingTing Hong,
my parents Zhengjun Wang and Guizhi Hua,
and my brother Feng Wang.

ACKNOWLEDGEMENTS

Most of all, I would like to thank my advisor Professor Wayne E. Stark for his encouragement, guidance, and support. His confidence and patience in me helped me pass difficult times of research. His extensive knowledge in the field of communications has always been a precious asset to me. I have benefited tremendously from his insight and advice. His way of thinking will surely shape the way I do research in my future career.

I would like to thank other committee members, Professor Demosthenis Teneketzis, Professor Stéphane Lafortune, and Professor Robert Smith for their careful reading of my dissertation. They have been great resources for my research.

I am very grateful to Mr. Michael Frieze and Mrs. Linda Frieze, whose financial support at the beginning of my doctoral program made this dissertation possible.

My thankfulness goes to all the teachers who taught me and led me into the wonderful world of human knowledge. I would also like to thank Ms. Linda Cox, Ms. Beth Olsen, Ms. Ann Pace, Ms. Susan Yale, and Ms. Beth Lawson, whose work made my life at the department of electrical engineering and computer science much easier.

My lab members, who are always great sources for discussion on everything, made my life at Ann Arbor a happy and joyful time. At the risk of leaving out many names, I would like to thank John Choi, Sangjin Hong, Tingfang Ji, Paul Liang, Troy Nolan, Andrew Worthen, Kar-Peoo Yar, Ping-Cheng Yeh, Ali Yilmaz, Do-Sik

Yoo, and Salam Zummo for their help at the Wireless Communications Lab.

To my parents, whose unconditional love and support have always been with me in my life, I owe a forever gratitude. I thank my brother with all my heart for taking care of our parents while I am away from home and for all the happy time we were together. To my beloved TingTing, whose loving and caring made me go through difficult moments in life, I would like to express my deep appreciation.

The support from Department of Defense Research & Engineering (DDR&E) Multidisciplinary University Research Initiative (MURI) on “Low-Energy Electronics Design for Mobile Platforms” managed by Army Research Office (ARO) under grant DAAH04-96-1-0377 and a grant from Office of Naval Research allowed me to work on many of the topics presented in this dissertation.

The results presented in Chapter II were obtained from the research project “Low-Energy Electronics Design for Mobile Platforms” that started at the University of Michigan in 1996. Many individuals at the University of Michigan have contributed to this project and to numerous discussions that led to the elaboration of the design methodology and associated simulation experiments presented in Chapter II. I wish to extend special thanks to the following individuals for their help at different stages of this work: George Barrett, Olivier Boivineau, Vuk Borich, Rami Debouk, Jack East, Riten Gupta, Al Hero, Sangjin Hong, Tara Javidi, Tingfang Ji, Linda Katehi, Seksan Kiatsupaibul, Il-Suek Koh, Paul Liang, JongTae Lim, Christopher Lott, Pinaki Mazumder, David Neuhoff, Troy Nolan, Kamal Sarabandi, John Shumpert, Robert Smith, Kimberly Wasserman, Andrew Worthen, and Do-Sik Yoo. In addition, I acknowledge the helpful comments by Jim Freebersyser, Jim Harvey, Jim Mink, and Bill Sander.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	viii
LIST OF TABLES	xiii
LIST OF APPENDICES	xiv
CHAPTER	
I. OVERVIEW OF RESEARCH IN WIRELESS COMMUNICATION NETWORKS	
1.1 Point-to-Point Communication	1
1.2 Broadcast and Multiple Access Communication	2
1.3 Network Coding	3
1.4 Challenges in Wireless Communication Networks	4
1.5 Contributions	5
1.6 Organization of Dissertation	7
II. LOW ENERGY WIRELESS COMMUNICATION NETWORK DESIGN	
2.1 Introduction	10
2.2 System Design Methodology	14
2.3 Design Example: Models for System Decomposition	18
2.3.1 Device Layer	18
2.3.2 Processing Layer	21
2.3.3 Network Layer	26
2.3.4 Performance Metric	38
2.4 Design Example: Global Optimization	40
2.5 Performance Results	44
2.5.1 Integrated Design for Mobility Model 1	44
2.5.2 Merits of Integrated Design	47

2.6	Conclusion and Future Research	49
III.	ROUTING IN BROADCAST MODE	51
3.1	Motivation for Routing in Ad Hoc Mobile Wireless Networks	51
3.2	Current Wireless Network Routing Protocols	53
3.3	Definitions	55
3.4	Problems of Routing in Broadcast Mode	57
3.5	Previous Research	60
3.5.1	Previous Results Related to Problem 3.4.1 through 3.4.3	60
3.5.2	Previous Results Related to Problem 3.4.4 through 3.4.6	61
3.6	Some Results of Routing in Broadcast Mode	61
3.6.1	Algorithm for Problem 3.4.1 through 3.4.3	61
3.6.2	Algorithm for Problem 3.4.4 through 3.4.6	75
3.7	Conclusion and Future Research	77
3.7.1	Conclusion	77
3.7.2	Future Research	78
IV.	OPTIMAL TRANSPORT EFFICIENCY OF LINEAR NET- WORKS	79
4.1	Introduction	79
4.2	Linear Network Topology	83
4.3	Performance Analysis	87
4.3.1	Threshold Model	91
4.3.2	Cutoff-Rate Model	98
4.3.3	Uncoded Model	99
4.3.4	Convolutional-Coded Model	99
4.4	Numerical Results	100
4.4.1	Transport Efficiency for Packet Size 224 Bits	101
4.4.2	Transport Efficiency for Packet Size 1000 Bits	107
4.4.3	Comparison of All Models	112
4.4.4	Observation	112
4.4.5	Discussion	113
4.5	Conclusion and Future Research	114
V.	TRANSPORT EFFICIENCY OF LINEAR NETWORKS WITH SPACE-TIME CODING AND ORDINARY END-FIRE AN- TENNA ARRAYS	115
5.1	Introduction	115
5.2	Space-Time Coding and Beam-Forming	117
5.3	Ordinary End-Fire Antenna Array	122
5.4	Interfering Communication for Linear Networks	128

5.4.1	Interfering Linear Networks	128
5.4.2	Transport Efficiency for Space-Time Coding	131
5.4.3	Transport Efficiency for the Ordinary End-Fire Antenna Array	135
5.5	Numerical Results	139
5.6	Conclusion	148
VI. DISTRIBUTED SPACE-TIME CODING AND COOPERATIVE COMMUNICATION		149
6.1	Introduction	149
6.2	Capacity for Distributed Space-Time Coding with Individual Constraint and Covariance Feedback	150
6.3	Calculation of Capacity	168
6.4	Cooperative Communication for Linear Networks	170
6.4.1	Cooperative Linear Networks	170
6.4.2	Transport Efficiency	170
6.5	Numerical Results and Discussion	172
VII. CONCLUSION AND FUTURE RESEARCH		175
7.1	Summary of Contributions and Conclusion	175
7.2	Future Research	177
APPENDICES		178
BIBLIOGRAPHY		200

LIST OF FIGURES

<u>Figure</u>		
1.1	Point-to-point communication scenario.	2
1.2	Point-to-point communication system modeling.	2
1.3	Point-to-multipoint communication scenario.	3
1.4	multipoint-to-point communication scenario.	3
2.1	Three nodes in a network.	12
2.2	Layered design/optimization.	15
2.3	Characteristics of the power amplifier.	19
2.4	Processing layer block diagram.	21
2.5	P_e as a function of E_{ct} and E_{cr} for convolutional code with the length of information sequence being 224 bits.	24
2.6	P_e as a function of E_{ct} and E_{cr} for convolutional code with the length of information sequence being 1018 bits.	25
2.7	Mobility model 1.	28
2.8	Mobility model 2.	29
2.9	Markov chain for mobility model 2, $\rho = 0.05$	29
2.10	Two-Path Propagation Model.	31
2.11	TDMA for single-hop transmission protocol.	33
2.12	Coupling of different layers.	40

2.13	Performance comparison of different algorithms with different drop areas.	45
2.14	Performance comparison of different design methodologies.	50
3.1	Categorization of routing protocols in ad hoc mobile wireless networks.	53
3.2	Example of inappropriateness of solution to Problem 3.4.1.	59
3.3	Demonstration of algorithm for Problem 3.4.3.	65
3.4	Demonstration of algorithm for Problem 3.4.3 (continued).	66
3.5	Demonstration of algorithm for Problem 3.4.3 (continued).	67
3.6	Demonstration of algorithm for Problem 3.4.3 (continued).	68
3.7	Demonstration of algorithm for Problem 3.4.3 (continued).	69
3.8	Demonstration of algorithm for Problem 3.4.3 (continued).	70
4.1	Linear wireless network, where node 1 wants to transmit a packet to node $k + 1$	83
4.2	Simplified power amplifier model for the class AB amplifier in Figure 2.3. A constant power $P_h = 35$ mW is turned into heat when the amplifier is in operation.	85
4.3	Packet error probability versus signal-to-noise ratio per information bit for the four different models, with each packet containing 224 bits or 1000 bits. The minimum signal-to-noise ratio for the channel capacity to be $\frac{1}{2}$, $\frac{3}{4}$, $\frac{7}{8}$, $\frac{15}{16}$, and 0.96 bit / (channel use) for the binary input antipodal signaling (BPSK) AWGN channel is shown. The packet error rate for the cutoff-rate model is shown for cutoff-rate $\frac{1}{2}$ and cutoff-rate 0.97.	90
4.4	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for threshold model.	101
4.5	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for cutoff-rate model. The length of information sequence is 224 bits.	102

4.6	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for uncoded model. The length of information sequence is 224 bits.	102
4.7	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for convolutional-coded model. The length of information sequence is 224 bits.	103
4.8	Optimal number of hops for threshold model.	103
4.9	Optimal number of hops for cutoff-rate model. The length of information sequence is 224 bits.	104
4.10	Optimal number of hops for uncoded model. The length of information sequence is 224 bits.	104
4.11	Optimal number of hops for convolutional-coded model. The length of information sequence is 224 bits.	105
4.12	Optimal packet error probability for cutoff-rate model. The length of information sequence is 224 bits.	105
4.13	Optimal packet error probability for uncoded model. The length of information sequence is 224 bits.	106
4.14	Optimal packet error probability for convolutional-coded model. The length of information sequence is 224 bits.	106
4.15	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for cutoff-rate model. The length of information sequence is 1000 bits.	107
4.16	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for uncoded model. The length of information sequence is 1000 bits.	108
4.17	Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for convolutional-coded model. The length of information sequence is 1000 bits.	108
4.18	Optimal number of hops for cutoff-rate model. The length of information sequence is 1000 bits.	109

4.19	Optimal number of hops for uncoded model. The length of information sequence is 1000 bits.	109
4.20	Optimal number of hops for convolutional-coded model. The length of information sequence is 1000 bits.	110
4.21	Optimal packet error probability for cutoff-rate model. The length of information sequence is 1000 bits.	110
4.22	Optimal packet error probability for uncoded model. The length of information sequence is 1000 bits.	111
4.23	Optimal packet error probability for convolutional-coded model. The length of information sequence is 1000 bits.	111
4.24	Optimal transport efficiency in log-log scale for the threshold model, cutoff-rate model, uncoded model, and convolutional-coded model.	112
4.25	Optimal number of hops for the threshold model, cutoff-rate model, uncoded model, and convolutional-coded model.	113
5.1	Far-zone geometry for a linear antenna array, where d is the distance between adjacent elements in the antenna array, θ is the angle between the line connecting the reference point of the antenna array (usually the origin) with the point of field observation and the line where radiating elements lie.	123
5.2	Amplitude pattern $\left \overline{(AF)} \right $ of the ordinary end-fire antenna for t radiating elements when $d = \frac{\lambda}{4}$	125
5.3	Two parallel linear wireless networks that are separated by distance d_b . The upper linear network is referred to as linear network A and the lower linear network is referred to as linear network B	129
5.4	Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 1$ receiving antenna is used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	140
5.5	Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 2$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	141

5.6	Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 3$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	142
5.7	Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 4$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	143
5.8	Optimal adjacent distance with respect to the distance between linear networks when $r = 1$ receiving antenna is used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	144
5.9	Optimal adjacent distance with respect to the distance between linear networks when $r = 2$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	145
5.10	Optimal adjacent distance with respect to the distance between linear networks when $r = 3$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	146
5.11	Optimal adjacent distance with respect to the distance between linear networks when $r = 4$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.	147
6.1	Optimal $g^{(ct)}$ versus the distance between two linear networks, where each node has $r = 1$ receiving antenna.	173
6.2	Optimal adjacent distance versus the distance between two linear networks, where each node has $r = 1$ receiving antenna.	173

LIST OF TABLES

Table

2.1	Variables used in wireless system modeling.	31
3.1	List of current wireless routing protocols.	54
3.2	Weight between each pair of nodes in Example 3.6.1.	64
4.1	Optimal Code Rate.	95

LIST OF APPENDICES

Appendix

A.	SIMULATED ANNEALING ALGORITHM	179
B.	PROPER COMPLEX RANDOM VARIABLES, VECTORS, AND PROCESSES	182
	B.1 Proper Complex Random Variables	182
	B.2 Proper Complex Random Vectors	184
	B.3 Proper Complex Random Processes	190
C.	CAPACITY OF MULTIPLE-ANTENNA COMPLEX GAUSSIAN CHANNELS	194
	C.1 Multiple-Antenna Complex Gaussian Channels	194
	C.2 Capacity of Deterministic Channels	196
	C.3 Capacity of Rayleigh Fading Channels	197

CHAPTER I

OVERVIEW OF RESEARCH IN WIRELESS COMMUNICATION NETWORKS

There has been tremendous progress in the wired and wireless networks during the past decade. With the ever popularity of IEEE 802.11x wireless LAN protocol, more and more wireless devices have been developed to accommodate the demand for wireless connectivity. Wireless networks have significantly different characteristics compared with wired networks. Wireless channels are susceptible to interference and fading, thus more prone to errors than wired channels. Since many wireless devices are powered by batteries, energy efficiency is a big concern for the wireless communications. We start with a survey of the development of communication theory and practices that motivate us the research topics in this dissertation. A perfect resource for many communication topics is [65].

1.1 Point-to-Point Communication

A simple scenario of communication is when there is only one transmitter and one receiver as shown in Figure 1.1. For such scenario, the communication system is usually modeled by the diagram shown in Figure 1.2. For the purposes of this dissertation, we are interested in the channel coding aspect of point-to-point



Figure 1.1: Point-to-point communication scenario.



Figure 1.2: Point-to-point communication system modeling.

communication. In 1948 Shannon published his seminal work on channel coding theory for the point-to-point communication [73] in an additive white Gaussian noise (AWGN) channel, where he characterized the point-to-point communication channel by a parameter, which he named the *capacity* of the channel. Shannon showed that reliable communication is only possible when the communication rate is less than the channel capacity. For more than fifty years, researchers have tried to design coded systems of reasonable complexity to achieve low error probability with signal-to-noise ratio close to the Shannon limit. A true breakthrough occurred when turbo codes were invented [7] in 1993. The later rediscovery of the Gallager's low-density parity check codes [27, 28] by MacKay [50] and the development of turbo-like coding and modulation techniques [51] allow point-to-point communication in AWGN channels to be very close to the Shannon limit albeit large delays. To some extent, the point-to-point communication problem has been successfully solved.

1.2 Broadcast and Multiple Access Communication

Other than the point-to-point communication scenario, we have broadcast scenario and multiple access scenario shown in Figure 1.3 and Figure 1.4 respectively. A typical example of broadcast scenario is when a transmitter is equipped with an omni-directional antenna and there are many receivers around it. Typical examples

of multiple-access scenario include Ethernet and cellular networks.

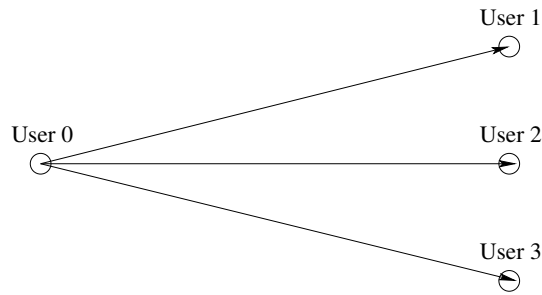


Figure 1.3: Point-to-multipoint communication scenario.

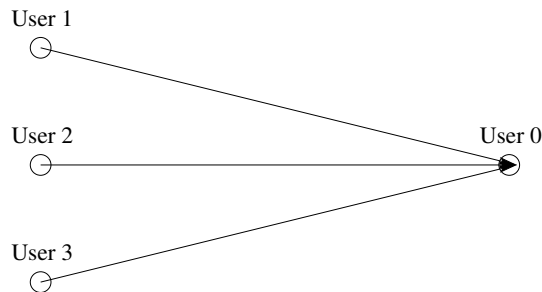


Figure 1.4: multipoint-to-point communication scenario.

There have been extensions of Shannon's theory from the point-to-point communication channels to the broadcast and multiple access communication channels. The most successful extensions have been [21, 23] for the broadcast channels and [75] for the multiple access channels. Cover gives a nice summary of the results for the capacity of broadcast channels in [22]. Unfortunately there has not been much interest in the design of practical coding and modulation schemes for these channels.

1.3 Network Coding

Network coding considers the situation where a source node wants to deliver its information bits to multiple destination nodes, possibly via multiple hops. The current research in network coding is mainly limited to wired networks. If a source

node is allowed to transmit only one bit via all its fixed connections, a fundamental question is how many times the source has to transmit in order to deliver all its information bits to the destinations. A method that can possibly reduce the number of transmissions by the source is called network coding [2, 47, 49], where each participating node in the network performs some transformation to all the bits it receives via its incoming links and transmits the transformed bits out of its outgoing links.

The network coding technique, although promising for some applications, such as multicasting in the Internet, requires cooperations of many nodes in a network in some synchronized fashion, which may be too complicated for actual implementation. Probably for this reason, there has not been much interest in the design of the practical coding schemes for network coding.

1.4 Challenges in Wireless Communication Networks

There are many research problems typical of wireless networks. Since many wireless devices operate on batteries, energy efficiency is a big challenge to system designers. Since wireless channels are more prone to errors, reliable communication is also a serious concern. The throughput of a wireless communication network may be limited by both energy efficiency and communication reliability.

In order to design efficient and reliable wireless networks, a profound general theory, analogous to the Shannon's theory for the point-to-point communication, is desired. However there has not been much progress along this direction [25]. There have been some research [33, 84, 85] on the capacity of different wireless networks, but its imminent practical application can not be easily seen.

The development of more efficient coding and modulation techniques, such as space-time coding, antenna array, and orthogonal frequency division multiplexing

(OFDM), provides wireless network designers with higher link capacity and higher energy efficiency. The effective integration of techniques at different layers of wireless communication systems will produce better wireless communication networks. This motivates us to investigate the integrated design methodology for wireless communication systems and analyze the relation of network throughput to parameters at other design layers in this dissertation.

1.5 Contributions

In this section we summarize our contributions in this thesis to the research in wireless communications.

1. We describe a generic integrated design methodology that is suitable for many kinds of mobile systems. The energy constraint for a communication system causes coupling between different design layers and makes the traditional separate layer design and optimization inappropriate. Drastically different from the traditional design, the integrated design methodology takes into account the coupling among the subsystems and simultaneously optimizes their operation under an energy constraint. Using our methodology, we are able to optimize a communication system across multiple design layers in a reasonable amount of time. We illuminate why integrated design methodology is better than the traditional design methodologies and show the improvement in performance that the integrated design methodology achieves over traditional design methodologies. The tradeoff between energy consumption and network performance is also demonstrated.
2. We consider routing in wireless networks. To achieve connectivity and energy efficiency, routing is an essential and efficient means for end-to-end communi-

cation in wireless networks. For wireless networks where each node is equipped with an omni-directional antenna, a point-to-multipoint connection can often be exploited for routing purposes. When the goal is to minimize the maximum power consumed by the nodes in a network, we provide a polynomial-time complexity algorithm to assign power to each node for unicast, broadcast, and multicast sessions. When the design goal is to minimize the total power consumed by all nodes in a network, we provide a polynomial-time complexity algorithm to assign power to each node for a unicast session and show that the computational complexity of routing algorithms for broadcast and multicast sessions is NP-hard. Our routing algorithms do not require the underlying graph after the power assignment be strongly connected.

3. We study the effects of power amplifier characteristics and receiver processing energy on the performance of a linear network. We introduce transport efficiency that accounts for both bandwidth efficiency and energy efficiency. We explicitly include the amplifier characteristics and receiver processing energy in the analysis of transport efficiency. We show that the optimal transport efficiency between two end nodes of a linear network at large distances is inversely proportional to the end-to-end distance for the threshold system. We observe that the same is true for any other type of communication systems under consideration from our numerical results.
4. We investigate the interference caused by space-time coding and the ordinary end-fire antenna array to neighboring networks. Our analytical and numerical results suggest that the ordinary end-fire antenna array gives higher transport efficiency than space-time coding when the number of receiving antennae is

small and gives transport efficiency close to that for space-time coding when two linear networks are close to each other. On the other hand, space-time coding gives higher transport efficiency than the ordinary end-fire antenna array when the number of receiving antennae is large and when two linear networks are far apart.

5. We show that cooperative communication with space-time coding between linear networks can improve transport efficiency, but it only gives marginal benefit if the cooperating networks are separated too far apart. This suggests that cooperative communication among networks should be used carefully since it usually incurs significant signal processing at the receiver.

1.6 Organization of Dissertation

This dissertation focuses on the analysis and design of wireless communication networks. In Chapter II we discuss about an integrated design methodology for low energy wireless communication networks, which is an important issue for mobile units operating on batteries. In Chapter III we describe energy-efficient routing algorithms for a wireless network where each node in the network is equipped with an omni-directional antenna, as it is the most common case for today's communication devices. In Chapter IV we investigate the transport efficiency of a linear network with different transmitter and receiver models and show how amplifier characteristics and receiver processing energy affect the transport efficiency. In Chapter V we derive interference models between linear networks when either space-time coding or an end-fire antenna array is used by each node and use transport efficiency to compare the network performance for these two techniques. In Chapter VI we address the problem of cooperative communication with space-time coding between different

linear networks. In Chapter VII we conclude the dissertation with a summary of contributions and indicate possible future research.

CHAPTER II

LOW ENERGY WIRELESS COMMUNICATION NETWORK DESIGN

Energy-efficient wireless communication network design is an important and challenging problem. Its difficulty lies in the fact that the overall performance depends, in a coupled way, on the following subsystems: antenna, power amplifier, modulation, error control coding, and network protocols. In addition, given an energy constraint, improved operation of one of the aforementioned subsystems may not yield a better overall performance. Thus, to optimize performance one must account for the coupling among the above subsystems and simultaneously optimize their operation under an energy constraint. In this chapter we present a generic *integrated design* methodology that is suitable for many kinds of mobile systems and achieves global optimization under an energy constraint. By pointing out some important connections among different layers in the design procedure, we explain why our *integrated design* methodology is better than traditional design methodologies. We present numerical results of the application of our design methodology to a situational awareness scenario in a mobile wireless network with different mobility models. These results illustrate: (i) the improvement in performance that our integrated design methodology achieves over traditional design methodologies; and (ii) the tradeoff between

energy consumption and performance.

2.1 Introduction

Energy-efficient wireless communication network design is an important and challenging problem. It is important because mobile units operate on batteries with limited energy supply. It is challenging because there are many different issues that must be dealt with when designing a low energy wireless communication system (such as amplifier design, coding, modulation design, resource allocation and routing strategies), and these issues are coupled with one another. Furthermore, the design and operation of each component of a wireless communication system present tradeoffs between performance and energy consumption. *The key observation is that constraining the energy of the nodes in a wireless network imposes a coupling among the design components that cannot be ignored in performing system optimization.* Therefore, the challenge is to exploit the coupling among the various components of a wireless communication system and understand the tradeoff between performance and energy consumption in each individual component/subsystem, in order to come up with an overall integrated system design that has optimal performance with respect to some performance metric and achieves low energy (power). Traditional design methodologies that optimize each layer separately may not be appropriate in terms of overall system optimality. The purpose of this chapter is to present a methodology for the design, simulation and optimization of wireless communication networks that achieves maximum performance under an energy constraint. The presentation of our methodology also gives some insight as to why traditional design methodologies may not achieve overall system optimality. The integrated design methodology is applied to several scenarios of mobile ad hoc networks. The results

show that significant gains are possible with an integrated design approach compared to traditional designs.

Before we proceed, we illustrate through simple examples, the coupling among the different components of a wireless communication system, and highlight the tradeoff between performance and energy consumption at individual components of the system. To illustrate the coupling among different components of a wireless communication system we first need to describe some key features of the amplifier's operation.

Consider the design and operation of an amplifier. The amplifier boosts the power of the intended transmitted signal so that the antenna can radiate sufficient power for reliable communication. However, typical power amplifiers have maximum efficiency in converting DC power into RF power when the amplifier is driven into saturation. In this region of operation, the amplifier voltage transfer function is nonlinear. Because of this nonlinearity, the amplifier generates unwanted signals (so called *intermodulation products*) in the band of the desired signal and in adjacent bands. When the amplifier drive level is reduced significantly (large backoff), the amplifier voltage transfer characteristic becomes approximately linear. In this case it does not generate *intermodulation products*. However, with large backoff the amplifier is not able to efficiently convert DC power into RF power. Thus, there is considerable wasting of power at low drive levels, whereas at high drive levels the amplifier generates more interfering signals.

We can now illustrate the coupling among individual components arising in the design of a wireless system. Consider packet routing in a wireless network that contains no base stations (i.e., an ad hoc network). For simplicity, consider a network with nodes A, B, and C, as shown in Figure 2.1. If node A wants to transmit a

message to node C, it has two options: Transmit with power sufficient to reach node C in a single transmission, or transmit first from A to B with smaller power, and then from B to C. Since the received signal power typically decays with distance as d^α , for α between 2 and 4, there is significantly smaller power loss due to propagation in the second option because $d_{AC}^\alpha > d_{AB}^\alpha + d_{BC}^\alpha$. However, even though node A transmits with smaller output power, it does not necessarily proportionally decrease the amount of power actually consumed because of the amplifier's effect discussed above. Furthermore, besides the energy required for packet transmission, there are energy requirements for packet reception and information decoding. The probability of packet error that is achieved depends on the energy allocated to the receiver. Thus the optimal network protocol (direct transmission from A to C or routing from A to B to C) depends on the amplifier characteristics as well as the energy needed to demodulate and decode a packet. Consequently, there is a coupling among amplifier design, coding and modulation design, decoding design, and routing protocols.

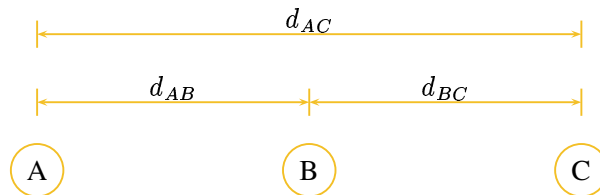


Figure 2.1: Three nodes in a network.

To highlight the tradeoff between energy consumption and system wide performance, consider the *situational awareness problem* in a mobile wireless network. In this problem, the objective of each node is to be aware of the position of every other node during a given time period. If energy consumption is ignored, and the overall performance metric is the average (over all nodes and over time) position estimation error, this error is minimized when all nodes continuously communicate their posi-

tions with one another. Such a strategy requires significant energy. If, on the other hand, the objective is to minimize the average position estimation error under an energy constraint, the nodes will have to jointly decide when to communicate and whom to communicate to during the given period, since a continuous communication strategy would use all available energy too quickly and could lead to large average position estimation error subsequent to the energy depletion of the battery.

Traditional design methodologies for wireless communication systems that attempt to optimize each layer separately may achieve global system optimality only by coincidence. However, through an understanding of the interactions and coupling among the functions at the different layers, it is possible to design a wireless communication system in a manner that truly integrates the functions of all layers. Therefore, we propose a methodology that decomposes the system into coupled layers and exploits the interactions among them to come up with an energy efficient design. The goal of the decomposition, besides a better understanding of the design procedure for global system optimality, is to obtain a computationally tractable approach to quantifying system performance with respect to different optimization criteria. Tackling such a problem can be a formidable task. We are not aware of any previous design and optimization attempts that encompass all the layers. Most previous research [15, 67, 69, 72, 74, 87] on low-energy ad hoc mobile wireless networks focuses on the optimization at the component/subsystem level. We hope that our work will provide some guidelines for further research in global system optimization.

The remainder of this chapter is organized as follows. In Section 2.2 we present a methodology for system design that incorporates the effect of the different layers on system performance. This methodology is fairly general and can be applied to many different applications besides the situational awareness scenario we consider later in

this chapter. In Section 2.3 we give detailed descriptions of the component models for the amplifier, propagation, coding, modulation, and network protocols for the system under investigation. In Section 2.4 we explain how global optimization works together with each system layer. In Section 2.5 optimization results for the situational awareness application are given. We conclude the discussion in Section 2.6.

2.2 System Design Methodology

In this section, we first describe the system decomposition and optimization, both of which form the constituent parts of our design methodology. We then comment on the decomposition and optimization. We consider a wireless network consisting of mobile nodes which need to communicate with one another in order to take some action or to share information, such as their respective positions. The overall goal is to characterize and optimize some performance metric under an energy constraint.

As we have pointed out in Section 2.1, in order to develop a systematic and computationally tractable design methodology, we divide the problem into interacting design layers as shown in Figure 2.2.

The system decomposition consists of three layers, namely, the device layer, the processing layer, and the network layer. Each layer interacts with layers above or below it in a well defined manner (described below). The device layer and the processing layer each perform local optimization in a manner that will be explained in more detail later in this section. In addition, there is a global optimization of an application-dependent performance measure that encompasses all layers.

At the device layer, we consider physical components at each node, including the antenna, the amplifier characteristics, and the circuit. At the processing layer, we consider the signal processing operations, including the modulation, coding, demod-

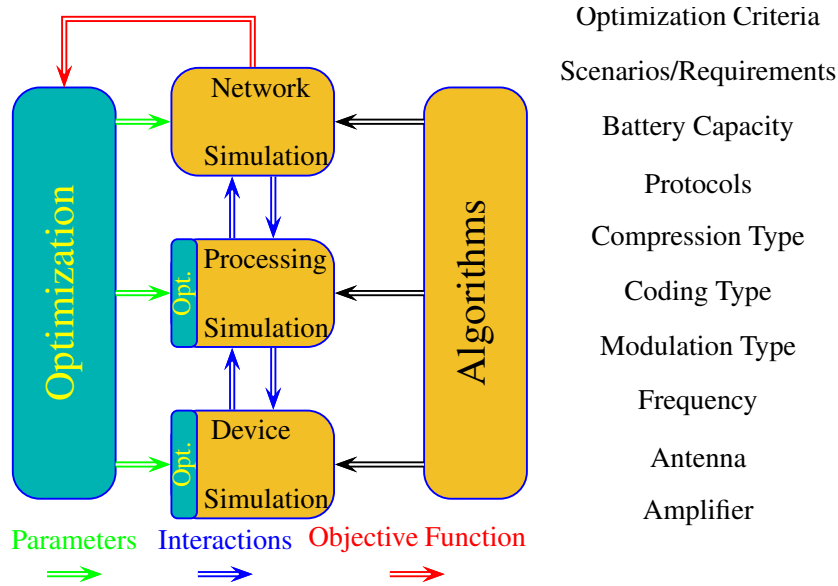


Figure 2.2: Layered design/optimization.

ulation, and decoding algorithms. At the network layer, we consider the collective operation of all mobile units, including the routing protocols, information distribution issues, communication environment, mobility modeling, and overall performance measure.

The system-wide objective is to optimize performance metric, that reflects, the collective operation of the mobile units, under an energy constraint. This is achieved by simultaneously optimizing over a number of parameters that characterize the objective. Some of these parameters also describe the coupling among the different layers, thus, they necessitate the development of an integrated design methodology. The integrated design methodology we propose is described by the following steps:

- Step 1. Identify the direct interactions (key coupling parameters) among layers; indirect interactions will “trickle through” the model.

For example: (i) the packet error probability, provided by the processing layer, is a key coupling parameter that directly affects network layer

performance; (ii) certain receiver parameters, such as, the numbers of bits of quantization for the equalizer input data, the equalizer coefficients and the decoder, affect network layer decisions indirectly through the packet error probability. For instance, hard decision decoding (one bit quantization) requires roughly 2 dB larger transmitted power to maintain the same level of packet error probability but decreases the amount of receiver energy necessary to process a packet.

- Step 2. At each layer consider a local performance measure that captures the contribution of that layer to the system-wide (global) performance criterion. Such a performance measure is a function of three types of parameters: (i) those that affect directly only the local performance criterion of the individual layer; (ii) those that are controllable and affect directly the performance of multiple layers; (iii) those that are uncontrollable and affect directly the performance of multiple layers. Fix the parameters of type (ii) and type (iii) and optimize the local performance criterion with respect to the parameters of type (i).

For example: At the processing layer, the packet error probability is a possible local performance criterion. It is a function of the three types of parameters described above. The type (i) parameters include the receiver parameters mentioned in Step 1. The type (ii) parameters include the energy constraints for transmitting and receiving a packet. The type (iii) parameters include the distances between each pair of nodes in the wireless network. We fix the parameters of type (ii) and type (iii) and optimize the packet error probability with respect to the type (i) parameters.

- Step 3. Using the results of Step 2, construct a model of each individual layer that

is a function of only the parameters of type (ii) and type (iii). Optimize the global performance criterion with respect to the parameters of type (ii) under an energy constraint.

For instance: Consider the situational awareness problem where the system-wide (global) performance criterion is the position estimation error of the network nodes averaged over type (iii) parameters from different layers (such as the distance between each pair of nodes). Optimize, under a constraint on the energy available to each node, the position estimation error with respect to type (ii) parameters (such as the energy for transmitting and receiving a packet) from the network and device layers.

The following comments are in accordance with each of the steps described in the *integrated design* methodology:

- In Step 2 the local optimization should be consistent with the global optimization. The effect of the local optimization is to “filter” out the parameters of type (i).
- In Step 2 and Step 3, the local and global optimization may have to be simulation-based because complete analytical expressions for the performance criteria of these steps may not be available.
- In Step 3 the model constructed for each individual layer may be based on table lookups. The global performance criterion is only a function of the parameters of type (ii), and the corresponding global optimization captures the complex interactions among layers.

We illustrate our modeling philosophy and *integrated design* methodology within the context of a situational awareness example presented in the next section.

2.3 Design Example: Models for System Decomposition

We apply our *integrated design* methodology to a particular network design problem, namely, the situational awareness problem in wireless mobile networks. In the situational awareness problem a number of mobile nodes desire to keep track of the location of all the other nodes over some time duration. The nodes operate with batteries and thus have a (finite) energy constraint. The transmission of information by a node requires a certain amount of energy as does the processing of any received signal. The goal of the design is to minimize the mean absolute error of the position estimates. There are a plethora of parameters that could be considered for optimization. We focus on a small set of parameters to illustrate the design and simulation methodology. In addition, we describe the system decomposition and justify our choice of the coupling parameters among different layers. We present the system optimization in Section 2.4. We proceed to describe each layer in a bottom-up manner.

2.3.1 Device Layer

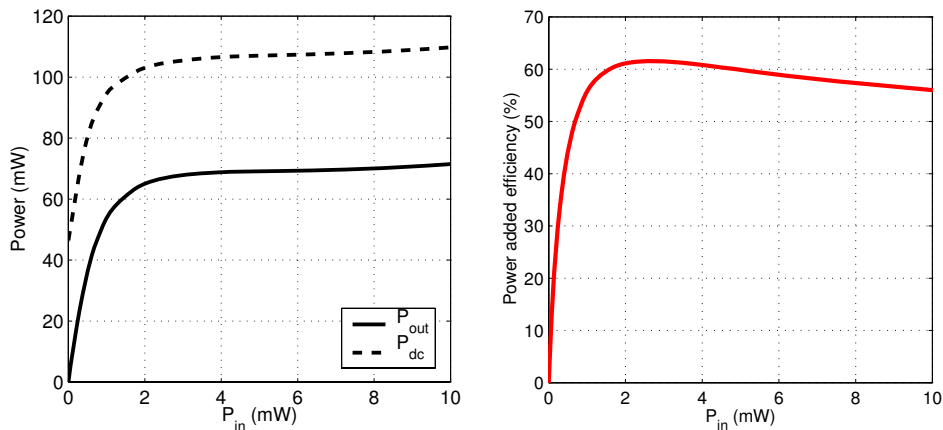
In this subsection, we present the model for the device layer and the coupling parameters between the device layer and the processing layer considered. We justify why the coupling parameters we choose are appropriate for the wireless communication systems under investigation. While not all parts of a transmitter and a receiver have been considered in the model of this chapter, we have chosen a few parameters that have an important coupling and illustrate the tradeoff between performance and energy consumption.

At the device layer, we assume each node has an omni-directional dipole antenna and a small amplifier. The amplifier is usually the component that consumes the

most energy in a wireless device. In our simulation we have a relatively small power amplifier model, with characteristics shown in Figure 2.3. Let P_{in} denote the input power, P_{out} denote the output power, and P_{dc} denote the consumed DC power. The power added efficiency is defined as

$$\text{Power added efficiency} = \frac{P_{out} - P_{in}}{P_{dc}}. \quad (2.1)$$

The characteristics of our class AB power amplifier [12] are tabulated for use at the processing layer.



(a) Radiated and DC power.

(b) Power added efficiency.

Figure 2.3: Characteristics of the power amplifier.

When the input power to the amplifier is low, the input-output voltage relation of the amplifier is fairly linear; however, in this region the amplifier operates at a very low efficiency, i.e., the ratio of useful output energy to consumed energy is low. If we drive the amplifier harder with higher input power, the amplifier operates at higher efficiency, with large input-output non-linearity. This non-linearity generates in-band and out-of-band signals called intermodulation signals, which adversely affect the performance of the processing layer, which in turn affects the design of the

network layer protocols. Out-of-band signals of a user, say user A , affect the performance of other users using adjacent frequencies, and out-of-band signals of other users affect the performance of user A . Because our global objective is to achieve high-precision situational awareness (i.e., low estimation errors) for every node under an energy constraint, it is important to understand the role of the amplifier power added efficiency in the overall optimization problem. In the current literature on energy-efficient routing protocols, the transmitted energy (power) is usually chosen as the routing metric. Unfortunately, this does not correspond to the actual consumed energy with most amplifiers. To account for the energy consumed by the amplifiers, we consider the following parameters associated with an amplifier's operation: the consumed DC power P_{dc} , the output power P_{out} , and the AM-to-AM voltage characteristics. These parameters also describe the coupling between the device layer and the processing layer, and depend on (are functions of) the energy consumed for a packet transmission. We note here that the intermodulation interference also depends on the modulation scheme chosen. Constant envelope modulation schemes have no intermodulation interference but have poor spectral efficiency. Non-constant modulation schemes can have significant intermodulation interference but have better spectral efficiency. As discussed below, we consider a non-constant modulation scheme. For such a modulation scheme, we characterize the relation between the average amplifier output power and the energy constraint E_{ct} for transmitting a packet by

$$P_{out} = g_1(E_{ct}). \quad (2.2)$$

This relation is tabulated for use by higher layers.

2.3.2 Processing Layer

At the processing layer one of the fundamental questions involves allocation of energy to transmitter versus receiver. The receiver uses energy to process and recover the data. More energy is used at the receiver by, for example, increasing the number of quantization levels allowed in the demodulation and decoding. A key question is to determine the optimal energy allocation between transmission and reception. In this subsection, we present the model for the processing layer, and the coupling parameters between the processing layer and the network layer. We further justify why the coupling parameters are appropriate for the wireless communication systems under investigation. A typical diagram of the processing layer is given in Figure 2.4. Note that we only consider the power consumed by the power amplifier, demodulator and channel decoder. The power consumed by other parts of the block diagram is typically much less than the power consumed by these three subsystems. In addition, the performance of these three subsystems is more sensitive to energy than other parts of the system.

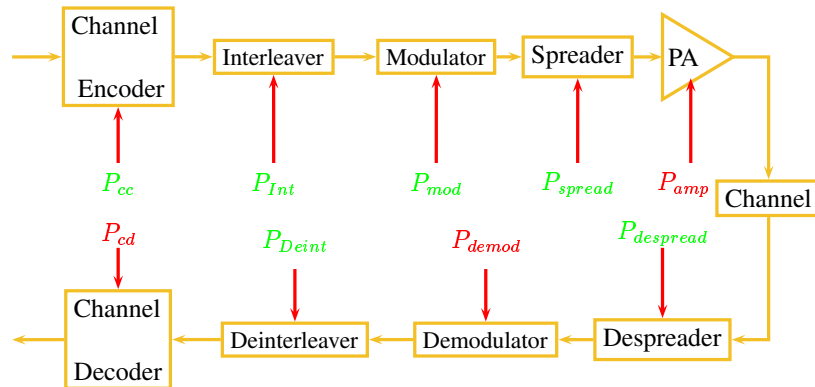


Figure 2.4: Processing layer block diagram.

We now describe the components of the processing layer; define the performance measures at the processing layer; identify the parameters that affect only the per-

formance of the processing layer, as well as those that affect the performance of additional layers; and present the results of optimizing the performance measures with respect to the parameters that affect only the performance of the processing layer.

We employ convolutional codes as error control codes. We consider BPSK with raised cosine filtering along with frequency-hopped spread spectrum for modulation and spreading. We use the Additive White Gaussian Noise (AWGN) channel as the channel model. We consider a tapped delay line model for quantization/matched filtering and coherent demodulation. We employ Viterbi decoding and iterative decoding. We fix several parameters such as the packet length, the preamble size for synchronization and equalization, the number of bits per hop, the sample rate at the input to the equalizer/matched filter, and the number of taps in the equalizer/matched filter. Some of the key characteristics of the above architecture of the processing layer are: (i) the combination of coding and frequency hopping can provide significant immunity to channel fading. (ii) The inclusion of error control coding mitigates fading, interference and thermal noise. Error control coding can significantly reduce the energy required for transmission for a given packet error rate at the expense of reduced spectral efficiency (bits/sec/Hz) and increased processing energy.

For the situational awareness problem considered in this chapter, the length of information sequence for convolutional coding is 224 or 1018 bits. These information bits are encoded using a standard rate 1/2 constraint length 7 convolutional code to form 460 (2048) channel bits. The channel bits are modulated by BPSK and the modulated signal is shaped by a square root raised cosine filter with rolloff parameter 0.3, which is implemented using 19 taps (without quantization). The symbol

duration is $20 \mu\text{s}$ and the waveform of each symbol is transmitted using a frequency-hopped spread spectrum with 23 (64) bits per hop, i.e., 20 (32) hops per packet. Before transmission the waveform is passed through a nonlinearity which models the amplifier voltage-to-voltage characteristic. This procedure generates inband interference depending on the operating point of the amplifier. At the receiver, the received signal with additive white Gaussian noise is sampled at four times the transmission rate (50 kb/s) before being passed through an equalizer/matched filter consisting of a tap delay line model with 19 taps and N_E bits of quantization for the coefficients and input samples. The decoder is a Viterbi algorithm with N_D bits of quantization for branch metrics.

The performance measures at the processing layer are defined to be the probability of packet error P_e and the bit error rate BER . Both P_e and BER affect the performance of the network layer. In general, for any choice of coding and demodulation schemes, P_e and BER depend on the energy constraint E_{ct} for the transmitter to send a packet, the energy constraint E_{cr} for the receiver to process a packet, the received signal-to-noise ratio SNR , the number N_E bits of quantization for equalizer data input and equalizer coefficients, and the number N_D bits of quantization for decoding. The parameter N_E and N_D affect only the performance of the processing layer. Consequently: (i) for given E_{ct} , E_{cr} , and SNR , we locally optimize P_e and BER with respect to N_E and N_D ; (ii) we generate parameterized versions of P_e with respect to E_{ct} , E_{cr} , and SNR , and build a performance table for these parameterized versions of P_e . The network layer (and global optimization) utilizes this table for calculating its global performance by generating the E_{ct} , E_{cr} , and SNR .

Figures 2.5 and 2.6 show the locally optimized packet error rate as a function of the transmitter energy constraint and the receiver processing energy constraint for

convolutional codes with the length of the information sequence being 224 bits and 1018 bits respectively. The surfaces are, from top down, $SNR = 1$ dB, 2 dB, 3 dB, 4 dB.

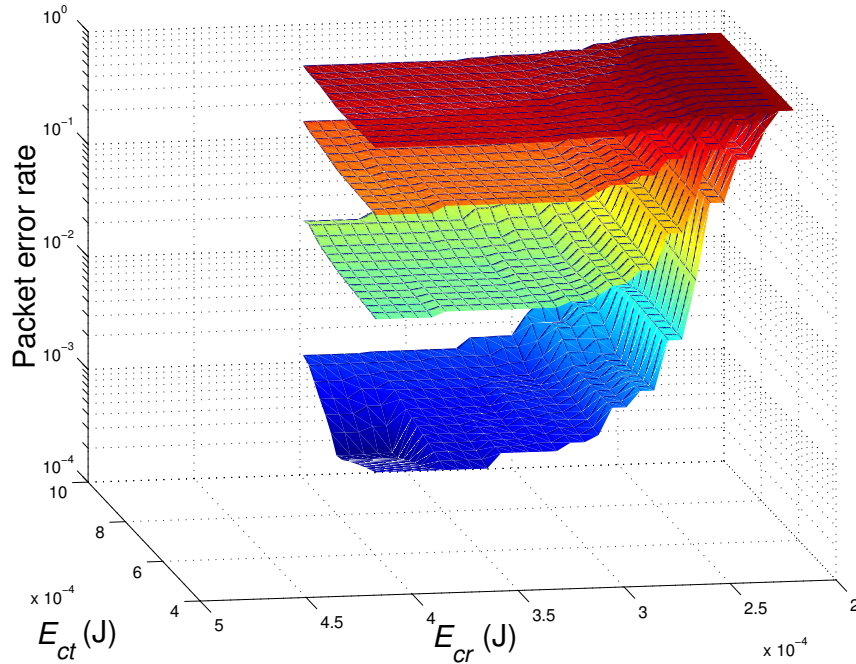


Figure 2.5: P_e as a function of E_{ct} and E_{cr} for convolutional code with the length of information sequence being 224 bits.

For each level of E_{ct} and E_{cr} , we estimate the power (energy) consumption by modeling individual components comprising the transmitter and receiver. We derive the actual energy consumption for transmitting a packet from the amplifier model at the device layer through table lookups. For the power (energy) required by a receiver to process an incoming packet, at the architectural and circuit levels, the main contribution to power consumption in complementary metal oxide semiconductor (CMOS) circuits is attributed to the charging and discharging of parasitic capacitors that occur during logical transitions [66]. The average switching energy

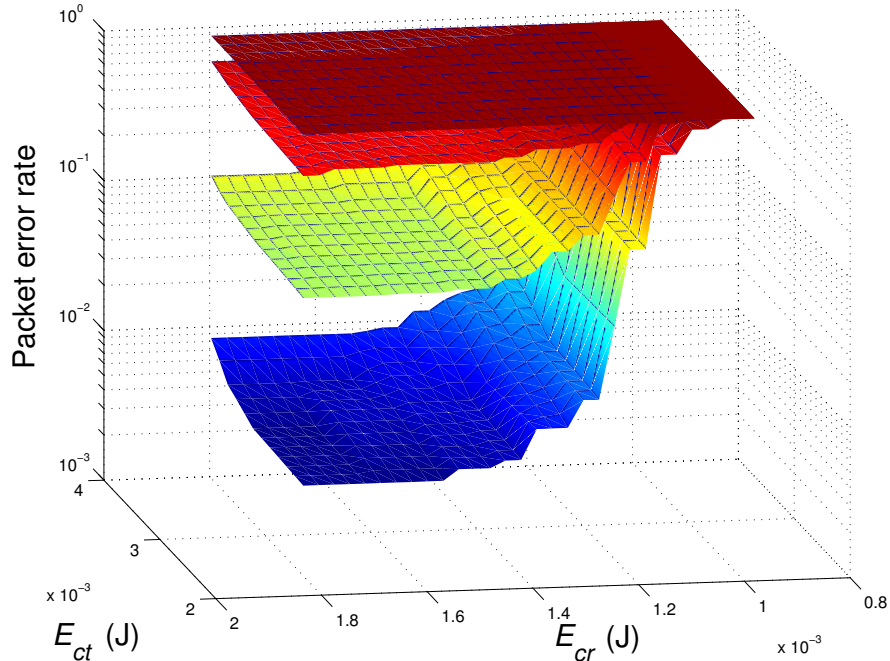


Figure 2.6: P_e as a function of E_{ct} and E_{cr} for convolutional code with the length of information sequence being 1018 bits.

E_{op} of a CMOS gate (or the power-delay product) is given by

$$E_{op} = C_{ave} V_s^2, \quad (2.3)$$

where C_{ave} is the average capacitance being switched per clock cycle and V_s is the supply voltage. The value of C_{ave} is the product of physical capacitance and switching activity. We assume that the switching activity is 50%. The capacitance of the gates (or components) is obtained based on 0.5- μm CMOS standard cell technology. Even though the quadratic dependence of energy on voltage makes it clear that operating at the lowest possible voltage is desirable for minimizing the energy consumption, the supply voltage of 3.3V is used in the power estimation process. The physical layout is obtained by synthesizing the algorithms (or functions) using Epoch CAD tool. At the architectural level, the algorithms and functions, such as pulse shaping filters and decoders, are optimized locally for low power consumption. For example,

FIR filters require multiplications of unknown data with known coefficients. This is typically implemented with a shift and add type of circuit. The number of 1's in the coefficient determines the amount of energy needed to perform the operation [35, 37]. Scaling all the coefficients to minimize the number of 1's will lead to lower power with identical performance [35, 37]. Similarly, power optimization is carried out for the convolutional decoders [36, 89].

For each bit transmitted, the total value of switching capacitance is estimated in order to compute energy consumed per bit. Using this value, the energy per packet is evaluated based on the number of the operations needed by each algorithm for demodulating and decoding a packet. We then build a table for the actual energy consumption E_{ta} for transmitting a packet and the actual energy consumption E_{ra} for receiving a packet based on the locally optimized N_E and N_D given the E_{ct} and E_{cr} constraint. The network layer utilizes this table by providing E_{ct} and E_{cr} to the processing layer. The actual energy can differ from the constrained energy because only discrete quantization levels are allowed.

In summary, at the processing layer a table is generated from simulation which contains the performance as a function of the energy E_{ct} allowed at the transmitter, the energy E_{cr} allowed at the receiver, and the signal-to-noise ratio SNR at the receiver. This table is then used at the network layer in determining the overall performance of the system.

2.3.3 Network Layer

In this subsection, we describe the model for the network layer and the parameters of the network protocols that affect global performance. We explain why these parameters are appropriate for the wireless communication systems under investiga-

tion.

We consider a network of nine nodes. The nodes move according to a single specific mobility model. In the situational awareness problem, each node attempts to keep track of the positions of all the other nodes. This is accomplished by communication and estimation. All nodes share their respective position information according to a specific communication protocol. In the rest of this section we present the mobility models, the propagation models, the communication protocols, and the estimation schemes used by the nodes.

Mobility Models

We describe two mobility models which we use in the various optimization problems that we consider. In both mobility models, each node in the network moves to a new location at the end of every T_m seconds. In the examples considered in this chapter $T_m = 1$.

In mobility model 1, we consider a region of size 6 km \times 6 km and a group of nine nodes initially deployed in an area as shown in Figure 2.7. All nodes travel toward the same destination which is located at $\underline{G} = (6000, 6000)$ m. Each node travels at average speed v , where $v = 1$ m/s (not drawn to scale in Figure 2.7). Let $\underline{w}_k^{(i)} = [w_{x,k}^{(i)}, w_{y,k}^{(i)}]^T$ be node i 's position at time kT_m . $\underline{w}_k^{(i)}$ is determined as follows:

1. Calculate a normalized direction vector $\underline{\theta}_{k-1}^{(i)} = [\theta_{x,k-1}^{(i)}, \theta_{y,k-1}^{(i)}]^T$ from node i 's position at time $(k-1)T_m$ toward the goal:

$$\underline{\theta}_{k-1}^{(i)} = \frac{\underline{G} - \underline{w}_{k-1}^{(i)}}{\|\underline{G} - \underline{w}_{k-1}^{(i)}\|}. \quad (2.4)$$

where $\|\underline{z}\| = \|[z_x, z_y]^T\| = \sqrt{(z_x)^2 + (z_y)^2}$ is the standard norm for $\underline{z} \in \mathbb{R}^2$.

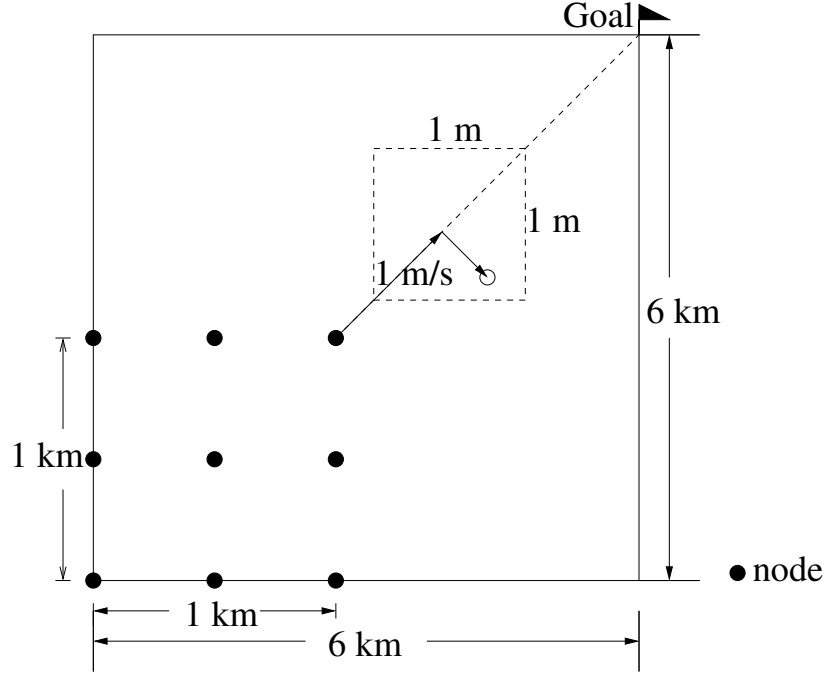


Figure 2.7: Mobility model 1.

2. Generate a random vector $\underline{\gamma}_k^{(i)} = [\gamma_{x,k}^{(i)}, \gamma_{y,k}^{(i)}]^T$, where $\gamma_{x,k}^{(i)}, \gamma_{y,k}^{(i)}$ are independent and uniformly distributed over $[-0.5, 0.5]$ m.
3. The position of node i at time kT_m is determined by

$$\begin{aligned} \underline{w}_k^{(i)} &= \underline{w}_{k-1}^{(i)} + vT_m \underline{\theta}_{k-1}^{(i)} + \underline{\gamma}_k^{(i)} \\ &= \underline{w}_{k-1}^{(i)} + vT_m \frac{\underline{G} - \underline{w}_{k-1}^{(i)}}{\|\underline{G} - \underline{w}_{k-1}^{(i)}\|} + \underline{\gamma}_k^{(i)}. \end{aligned} \quad (2.5)$$

If the position of any one of the nodes is within vT_m of \underline{G} , all nodes stop moving.

In mobility model 2, all nodes are initially deployed in a region of size 1332 m \times 1332 m and move within this region as shown in Figure 2.8. The mobility of each node is characterized by a two-state discrete-time Markov chain as shown in Figure 2.9, where the two states are labeled *Stay* and *Move*. The position of node i at time kT_m is determined as follows:

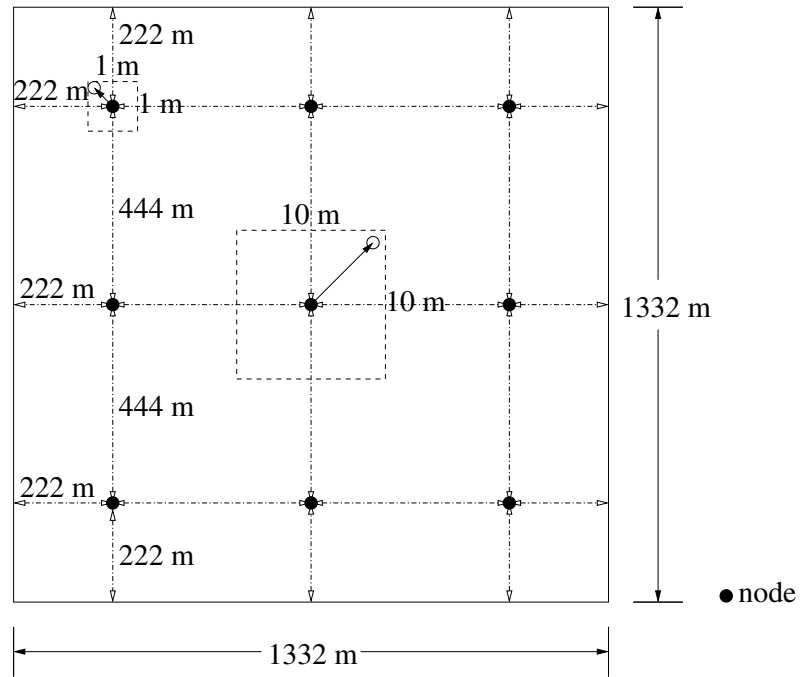
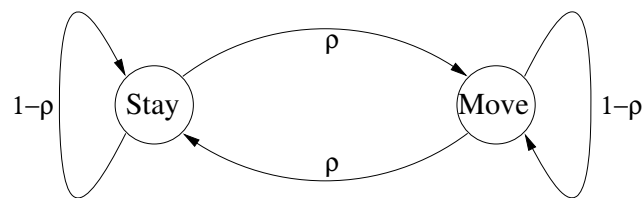


Figure 2.8: Mobility model 2.

Figure 2.9: Markov chain for mobility model 2, $\rho = 0.05$.

1. The Markov chain at each node changes state with probability ρ and remains at the same state with probability $1 - \rho$.
2. Depending on the current state, generate a random vector $\underline{\gamma}_k^{(i)} = [\gamma_{x,k}^{(i)}, \gamma_{y,k}^{(i)}]^T$. The random variables $\gamma_{x,k}^{(i)}, \gamma_{y,k}^{(i)}$ are independent and uniformly distributed over $[-s, s]$ m, where $s = 0.5$ when the Markov chain is the state *Stay* and $s = 5$ when the Markov chain is the state *Move*.
3. The position of node i at time kT_m is

$$\underline{w}_k^{(i)} = f \left(\underline{w}_{k-1}^{(i)} + \underline{\gamma}_k^{(i)} \right), \quad (2.6)$$

where

$$f(\underline{z}) = \begin{bmatrix} g(z_x) \\ g(z_y) \end{bmatrix} \quad (2.7)$$

and

$$g(x) = \begin{cases} 0 & \text{if } x < 0, \\ x & \text{if } 0 \leq x \leq 1332, \\ 1332 & \text{if } x > 1332. \end{cases} \quad (2.8)$$

Propagation Model

The transmitted signal from each node experiences propagation loss and fading. In the results that follow, we assume a two-path propagation model shown in Figure 2.10. Table 2.1 shows the variables and their typical values needed to explain the model.

The two-path propagation model consists of a direct path and a path reflected off the ground with 180 degree phase change at the reflection point from the transmitter to the receiver. The cumulative effect of both paths determines an attenuation A

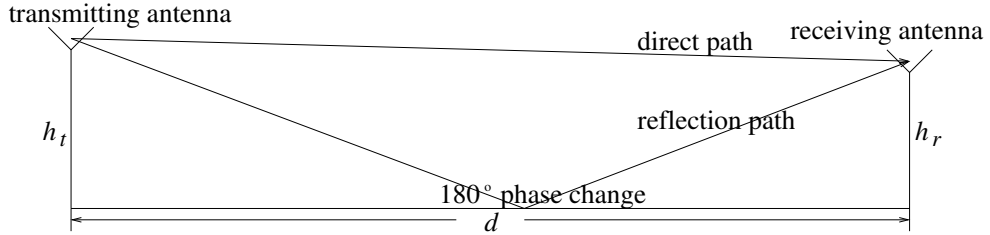


Figure 2.10: Two-Path Propagation Model.

Table 2.1: Variables used in wireless system modeling.

Variable	Meaning	Value	Unit
λ_c	carrier wavelength	10	m
h_t	height of transmitting antenna	1	m
h_r	height of receiving antenna	1	m
d	propagation distance	(0, 9000)	m
P_t	radiated power	$[10^{-7}, 1]$	W
P_r	received power	(2.9)	W

between received power and radiated power, which is given by [68]

$$A = \frac{P_r}{P_t} = 4 \left(\frac{\lambda_c}{4\pi d} \right)^2 \sin^2 \left(\frac{2\pi h_t h_r}{\lambda_c d} \right). \quad (2.9)$$

It should be noted that (2.9) is valid only when the receiver is in the far zone of the transmitter. In order for the receiver to be in the far zone of the transmitter, the distance d between the transmitter and the receiver has to satisfy the following three conditions:

1.

$$d \geq \frac{2D^2}{\lambda_c}, \quad (2.10)$$

where D is the largest physical linear dimension of the antenna.

2.

$$d \gg D \quad (2.11)$$

3.

$$d \gg \lambda_c \quad (2.12)$$

For instance, suppose $D = 0.5$ m, then it is required that $d \gg 10$ m for the receiver to be in the far zone of the transmitter. When $d \gg \max\{h_t, h_r\}$, we can approximate A in (2.9) by

$$A \approx \frac{h_t^2 h_r^2}{d^4}. \quad (2.13)$$

The two-path propagation model characterizes the large-scale propagation loss of many fading channels reasonably well and this is the reason why we adopt it in our network layer simulation.

Communication Protocols

We describe the medium access control (MAC) strategy as well as the routing protocols. We consider two communication protocols, namely, a single-hop transmission protocol (which may be considered as a single-hop routing protocol) and a multi-hop routing protocol. For both communication protocols, the omni-directionality of the antenna at each node makes the potential connections among nodes point-to-multiple-points, i.e., if a node sends out a packet, the electro-magnetic wave will propagate in all directions and may be received by many other nodes. Therefore in the design of wireless communication protocols, communication occurs in a broadcast medium, which is very different from traditional wired networks, where the connections are mostly point-to-point. Recall that in the situational awareness problem the objective is for each node to keep track of the positions of all other nodes as accurately as possible, so the omni-directionality property of the antenna that allows multiple receptions can be exploited in the design.

Single-hop Transmission Protocol

In the single-hop transmission protocol, each node transmits its position information packets every T seconds, where T is a design parameter. The medium access control is Time Division Multiple Access (TDMA), where each node is assigned a transmission slot of duration T/I , where $I = 9$ for our case, as shown in Figure 2.11. The slot duration is much larger than a packet duration. In a given slot, each packet

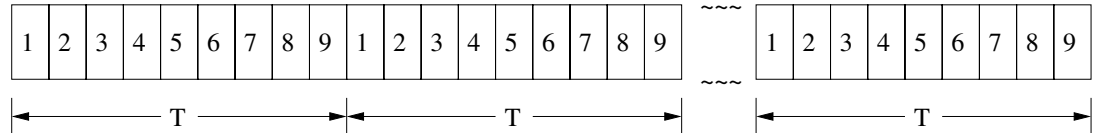


Figure 2.11: TDMA for single-hop transmission protocol.

transmission is followed, with probability q , by a retransmission, and so forth, until the slot ends. The retransmission probability q is considered as a design parameter. Since each node operates on a battery with limited capacity, we constrain the energy used for each packet transmission or retransmission to be upper bounded by E_{ct} , which we consider as a design parameter. The position information packets may be received by many other nodes, each of which consumes a certain amount of energy to process the packets. Again, due to the limited capacity of the battery, we constrain the energy used to process the incoming packets to be upper bounded by E_{cr} , which we consider as a design parameter. When a node receives a packet, it does not send back any acknowledgment, nor does it forward the packet it receives to other nodes. As a consequence, every packet in the single-hop transmission protocol travels only one hop. We therefore choose T, q, E_{ct}, E_{cr} as the design parameters at the network layer that affect global performance.

Multi-hop Routing Protocol

In the multi-hop routing protocol, there are mainly two issues we need to consider. These are (i) setting up a routing path; (ii) transmitting position information packets.

We discuss each issue separately.

(i) For routing, we adapt the Open Shortest Path First (OSPF) protocol [54, 53], which is a link state routing protocol, to our situational awareness scenario and take advantage of the omni-directionality of the antenna. The objective of our routing algorithm is to determine the first-hop nodes a node should reach, and the power this node should use to reach these first-hop nodes when it transmits position information packets originated by itself or when it forwards position information packets received from other nodes. Our routing algorithm works as follows:

- When nodes are initially deployed in an area, they do not have the knowledge of the topology of the network. In order to obtain and maintain their knowledge of the network topology, each node sends out a *hello* packet every T_r seconds, where T_r is a design parameter. The power used for transmitting the *hello* packet is fixed at 5 mW, which, under the propagation model described in 2.3.3, and for the transmission rate of 1 Mbps and the typical noise power spectral density of -174 dBm/Hz, gives a signal-to-noise ratio of approximately 0 dB at a receiver located 0.5 km from the transmitter. The *hello* packet contains the identity and position of the advertising node. When a node correctly receives a *hello* packet, it regards the sender as its neighbor. The receiver does not forward the *hello* packet, so the *hello* packet only travels a single hop.
- Based on our propagation model, the link metric for the routing algorithm is defined to be d^4 , where d is the distance between a pair of transmitter and receiver nodes. Each node computes the link metrics to its neighbors and stores them in its link database. Every node sends out packets containing the link metrics in its database every T_r seconds. The power used for the transmission

of packets containing link metrics is the same as that for the hello packets.

- When a node receives the packets containing the link metrics, it stores them in its link database. Gradually each node accumulates enough knowledge of the link metrics connecting each pair of nodes, thus obtaining the topology of the network. We need to emphasize that T_r , which is a design parameter, is very small compared to the speed of the nodes so that each node completes the updating of its link database before the positions of the nodes change significantly.
- Based on the link database, each node calculates the shortest path spanning tree from it to all other nodes using Dijkstra's algorithm [54, 53] and stores the first-hop nodes to which it should forward the position information packets in the routing table.
- Let d_{max} be the largest distance among the distances between a node and any of its first-hop nodes. Since the antenna is omni-directional and the source node knows its farthest first-hop node, it can take advantage of this by choosing the radiated power for position information packets to be Cd_{max}^4 , where C is a design parameter. Each node transmits only once with this power instead of transmitting separately to each of its first-hop nodes.
- Since nodes are constantly moving, the neighbors of a node may be out of reach and no longer be considered as neighbors. If node A does not receive any *hello* packet of its neighbor node B for $8T_r$ seconds, node A removes node B 's link information from its link database.
- A node recalculates its routes to all other nodes whenever it receives informa-

tion about a new link or when a newly received link metric differs from previous one, stored in the node's link database, by more than 20%.

(ii) At the application layer, each node transmits its position information packets¹ every T_p seconds, using radiated power Cd_{max}^4 , where T_p is a design parameter. The corresponding consumed energy E_{ct} for each transmission is obtained using the amplifier model.

In addition to the issues of routing path setup and position information transmission discussed above, several other issues must also be taken into consideration; these issues are briefly discussed below.

When a node receives a packet of any type, which can be a routing protocol packet or a position information packet, it consumes energy E_{cr} to process it and accepts the packet with a packet error probability, which depends on E_{ct} , E_{cr} , and the signal-to-noise ratio SNR . However, since we do not have a fixed-size packet (mostly due to routing protocol packets), E_{cr} is different for received packets of different sizes. We choose a nominal \bar{E}_{cr} for a single packet of length 224 bits (or 1018 bits) and use this \bar{E}_{cr} as a design parameter. The energy E_{cr} consumed for packets of other sizes is proportional to \bar{E}_{cr} .

Since the link database is distributed, there is a potential of infinite loops in the network, thus wasting all the energy that a node can possibly have if a packet is trapped in one of these loops. To prevent a packet from looping around in the network forever, we implement a time-to-live (*tll*) field in each packet as is done in IP. The value of *tll* is the number of hops remaining for the packet to travel in the network. When a node receives a packet from other nodes, it reduces the packet's

¹While the hello packets contain the position information of nodes, we do not use this information in the application. We do not allow this because the application may be different from a situational awareness problem where the position information does not provide any information to the application.

tll by one. If *tll* equals 0, the node does not forward the packet any more.

We conclude the discussion of the multi-hop routing protocol by addressing the medium access control strategy. As in the case of the single-hop transmission protocol, we use a TDMA protocol at the medium access control layer except that in the present case the slot duration allocated to each node is equal to the largest packet duration plus a nominal guard interval. In addition, the medium access control layer of a node may buffer the packets from the network and application layers, and it sends the packets at the transmission slot assigned to the node by the MAC protocol.

In summary, at the network layer we choose $T_r, T_p, C, \bar{E}_{cr}$ as the design parameters that affect global performance.

Estimation Schemes

We consider two estimation schemes corresponding to the two mobility models we described in 2.3.3. In both estimation schemes, each node updates its estimate of another node's position every T_e seconds. In the examples considered in this chapter $T_e = 2$.

We first describe estimation scheme 1, which is used for mobility model 1. Let $\underline{\hat{w}}_k^{(i,j)} = [\hat{w}_{x,k}^{(i,j)}, \hat{w}_{y,k}^{(i,j)}]^T$ be node i 's estimate of node j 's position at time kT_e , where $j \in \{1, \dots, I\} \setminus \{i\}$ and I is the total number of nodes in the network. Depending on whether node i correctly receives a packet from node j between time $(k-1)T_e$ and kT_e , the position estimate is updated differently. If node i correctly receives a packet from node j , then

$$\underline{\hat{w}}_k^{(i,j)} = \underline{a}_{k-1}^{(i,j)} + v(kT_e - t_{k-1}) \underline{\hat{\theta}}_a^{(i,j)}, \quad (2.14)$$

where $\underline{a}_{k-1}^{(i,j)}$ is the actual position of node j at time t_{k-1} and

$$\hat{\underline{\theta}}_a^{(i,j)} = \frac{\underline{G} - \underline{a}_{k-1}^{(i,j)}}{\left\| \underline{G} - \underline{a}_{k-1}^{(i,j)} \right\|}. \quad (2.15)$$

If node i does not correctly receive a packet from node j , then

$$\underline{\hat{w}}_k^{(i,j)} = \underline{\hat{w}}_{k-1}^{(i,j)} + vT_e \hat{\underline{\theta}}_e^{(i,j)}. \quad (2.16)$$

where

$$\hat{\underline{\theta}}_e^{(i,j)} = \frac{\underline{G} - \underline{\hat{w}}_{k-1}^{(i,j)}}{\left\| \underline{G} - \underline{\hat{w}}_{k-1}^{(i,j)} \right\|}. \quad (2.17)$$

These estimates are based on the following strategy. Node i knows the mobility model for all the other nodes. Since the mobility model is such that nodes move toward the goal in a straight line subject to noise, the new estimate is the extrapolation toward the goal, of the position contained in the packet that was last received correctly, by an amount proportional to the product of velocity and time.

We now describe estimation scheme 2, which is used for mobility model 2. The estimate $\underline{\hat{w}}_k^{(i,j)}$ is determined by:

$$\underline{\hat{w}}_k^{(i,j)} = \underline{w}_l^{(j)}, \quad (2.18)$$

where $\underline{w}_l^{(j)}$ is the position of node j at time l contained in the packet from node j last received by node i . Since the mean of the increment of node j 's position at any time is zero, if node i does not correctly receive any packet between time l and kT_e , a reasonable estimate of node j 's position at time kT_e is node j 's position at time l .

2.3.4 Performance Metric

For the purpose of optimization, we use mean absolute error as the performance metric. For both estimation schemes, the estimation error of node j 's position made

by node i at time kT_e is defined as

$$\underline{e}_k^{(i,j)} = \underline{w}_k^{(j)} - \underline{\hat{w}}_k^{(i,j)}, \quad (2.19)$$

where $\underline{w}_k^{(j)}$ is the actual position of node j at time kT_e . The performance metric $J^{(i)}$ at node i is defined to be

$$J^{(i)} = \mathbb{E} \left[\frac{1}{K(I-1)} \sum_{j=1, j \neq i}^I \sum_{k=1}^K \left\| \underline{e}_k^{(i,j)} \right\| \right], \quad (2.20)$$

where KT_e is the time horizon under consideration. In the above equation, the expectation is with respect to the mobility, the noise in the receiver, and the randomness in retransmission (in the case of single-hop transmission protocol only).

The overall network performance measure is given by the average of the position estimation error contributed by all the nodes in the network:

$$J = \frac{1}{I} \sum_{i=1}^I J^{(i)}. \quad (2.21)$$

The goal is to minimize J over the parameters that affect global performance subject to a constraint on the energy used by each node. Let $E^{(i)}$ denote the energy used by node i over the time horizon KT_e . The constraint on energy is

$$\max_{1 \leq i \leq I} E^{(i)} \leq E. \quad (2.22)$$

Based on the discussion in Section 2.2 and Sections 2.3.1 through 2.3.3, we have the following objectives: For the single-hop transmission protocol the goal is to determine the design parameters

$$[T^*, q^*, E_{ct}^*, E_{cr}^*]^T = \arg \min_{\substack{[T, q, E_{ct}, E_{cr}]^T \\ \max E^{(i)} \leq E}} J(T, q, E_{ct}, E_{cr}) \quad (2.23)$$

and the corresponding performance $J^* = J(T^*, q^*, E_{ct}^*, E_{cr}^*)$. For the multi-hop routing protocol the goal is to determine the design parameters

$$[T_r^*, T_p^*, C^*, \bar{E}_{cr}^*]^T = \arg \min_{\substack{[T_r, T_p, C, \bar{E}_{cr}]^T \\ \max E^{(i)} \leq E}} J(T_r, T_p, C, \bar{E}_{cr}) \quad (2.24)$$

and the corresponding performance $J^* = J(T_r^*, T_p^*, C^*, \bar{E}_{cr}^*)$.

2.4 Design Example: Global Optimization

In this section, we illustrate how our methodology, described in Section 2.2, applies to a wireless system in the situational awareness scenario described in Section 2.3. We consider the single-hop transmission protocol and mobility model 1. Similar procedures apply to the other settings. The parameters that describe the coupling among the layers and that must be shared by the different layers in the setting under consideration are shown in Figure 2.12.

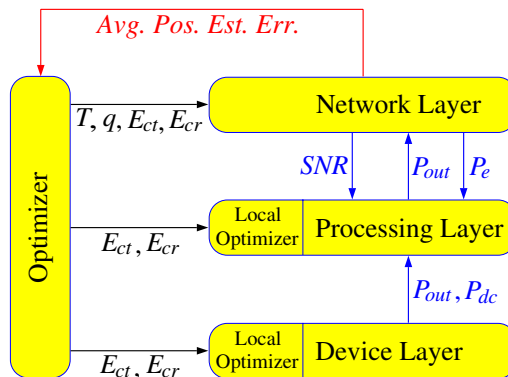


Figure 2.12: Coupling of different layers.

The integrated design methodology of Section 2.2 is applied in an iterative fashion. At each iteration the optimization is, in part, simulation-based because we do not have precise analytical expressions for the local and global optimization criteria we employ. The optimization program attempts to find the global minimum of the objective function J in (2.21). The global optimization and simulation modules perform the following steps in attempting to find the globally optimal solution (illustrated for scenario 1):

Step 1. The “optimizer” module determines the (new) parameters $[T, q, E_{ct}, E_{cr}]^T$, for which the network performance is to be evaluated.

Step 2. The “network simulator” module approximates the objective function in (2.21) for the given $[T, q, E_{ct}, E_{cr}]^T$ using Monte-Carlo simulation techniques. It returns the negative of the average position estimation error to the “optimizer” module (since the optimization used is based on the maximization of an objective function).

Step 3. Steps 1 and 2 are repeated until a terminating condition is reached.

We now describe each step in more detail. The “optimizer” module used in Step 1 is a type of *simulated annealing* algorithm. The method of *simulated annealing* is a technique that has attracted significant attention as it is suitable for optimization problems with a large number of parameters, especially ones where a desired global extremum is hidden among many local extrema. The *simulated annealing* algorithm used in our *integrated design* is called *Hide-and-Seek* and was developed by Romeijn and Smith [70]. The steps of the *Hide-and-Seek* algorithm are given in Appendix A.

We have implemented the “network simulator” module in *OPNET*, a widely used network development and analysis tool [59]. Our “network simulator” module has the following steps that involve the interactions among the three layers for the above global optimization of Step 2:

Step 2.1 Given the parameters E_{ct} and E_{cr} selected by the “optimizer” module, the device layer determines the amplifier output power P_{out} , the actual consumed energy E_{ta} for transmitting a packet, and the actual consumed energy E_{ra} for receiving a packet:

$$P_{out} = g_1(E_{ct}), \quad (2.25)$$

$$E_{ta} = g_2(E_{ct}), \quad (2.26)$$

$$E_{ra} = g_3(E_{ct}, E_{cr}). \quad (2.27)$$

The function g_1 captures the result of the local optimization of the amplifier model at the device layer, which is parameterized in terms of E_{ct} , as explained in Section 2.3.1. The resulting P_{out} is used at the network layer (see Step 2.2 below). We assume that the energy used by the transmitter is always all of E_{ct} and therefore $g_2(E_{ct}) = E_{ct}$. Due to the quantization at the demodulator and the decoder at the receiver, the actual consumed energy E_{ra} may be smaller than the energy constraint E_{ct} ; their relation is determined by the function g_3 that is well-defined and known.

Step 2.2 For each transmission scheduled at the network layer and for each receiver, the network layer determines the SNR at the output of the receiving antenna as follows. The radiated power P_t is the product of the amplifier output power P_{out} , the gain G_t of the transmitting antenna, and the efficiency η_t of the transmitting antenna, i.e.,

$$P_t = P_{out}G_t\eta_t. \quad (2.28)$$

In the examples considered in this chapter, $G_t = 1$ and $\eta_t = 0.2229$. The power received at the output of the receiving antenna is

$$P_r = A\eta_rG_rP_t, \quad (2.29)$$

where A is the channel attenuation given by (2.9), η_r is the efficiency of the receiving antenna, and G_r is the gain of the receiving antenna. In the examples considered in this chapter, $G_r = 1$ and $\eta_r = 0.2229$.

Substituting (2.9) and (2.28) into (2.29) yields

$$P_r = P_{out} 4 \left(\frac{\lambda_c}{4\pi d} \right)^2 G_t G_r \eta_t \eta_r \sin^2 \left(\frac{2\pi h_t h_r}{\lambda_c d} \right). \quad (2.30)$$

The SNR is

$$SNR = \frac{P_r \cdot T_s}{N_0}, \quad (2.31)$$

where T_s is the symbol duration and N_0 is the power spectral density of the thermal noise at the receiver. The remaining variables in the above equations are defined in Table 2.1.

Step 2.3 For each transmission scheduled at the network layer and for each receiver, the processing layer determines the packet error rate P_e using SNR obtained from the network layer (Step 2.2) and using E_{ct} and E_{cr} obtained from the “optimizer” module. P_e is obtained at the processing layer from E_{ct} , E_{cr} , and SNR by solving an optimization problem with respect to the parameters N_E and N_D as described in Section 2.3.2. For the sake of computational efficiency, this optimization at the processing layer is done off-line and its results are summarized using the function g_4 whose arguments are E_{ct} , E_{cr} , and SNR :

$$P_e = g_4(E_{ct}, E_{cr}, SNR). \quad (2.32)$$

Therefore, the “network simulator” module implements this step by table look-up. For a given P_e , the network layer flips a biased coin to determine if each packet is correctly received.

Step 2.4 Each node uses the estimation model described in Section 2.3.3 to update its estimates of the position of the other nodes. The “network simulator” module accumulates the position estimation errors of all nodes in the

network. When the simulation terminates, the network layer averages the accumulated value over all nodes and over time. Thirty “network simulation” runs are produced and their average is the approximation of the objective function in (2.21) that is returned to the “optimizer” module.

In Step 3, the termination condition that we chose for our experiments was to stop after 200 iterations of Steps 1 and 2.

It is critical to understand that the simulation and optimization effort has been carefully divided into a device layer simulation, a processing layer simulation, and a network layer simulation. More importantly, the interactions between layers have been identified and incorporated into the performance evaluation.

2.5 Performance Results

In this section, we present numerical results for two situational awareness scenarios that highlight our integrated design methodology and compare its performance to that of traditional design methodologies. In particular, within the context of mobility model 1, we use our design methodology to compare the performance of a single-hop transmission protocol with that of a multi-hop routing protocol. Furthermore, within the context of mobility model 2 and single-hop transmission, we compare our design methodology with traditional design methodologies and illustrate the improvement in performance achieved by our approach.

2.5.1 Integrated Design for Mobility Model 1

First we consider mobility model 1 (please refer to Section 2.3.3) where the nodes are initially located in either a $1 \text{ km} \times 1 \text{ km}$ or a $2 \text{ km} \times 2 \text{ km}$ area and move towards a single goal. We evaluate and optimize the objective function given in (2.21)

by Monte-Carlo simulation using the steps described in Section 2.4. The resulting performance is shown in Figure 2.13. As can be seen in Figure 2.13, the single-

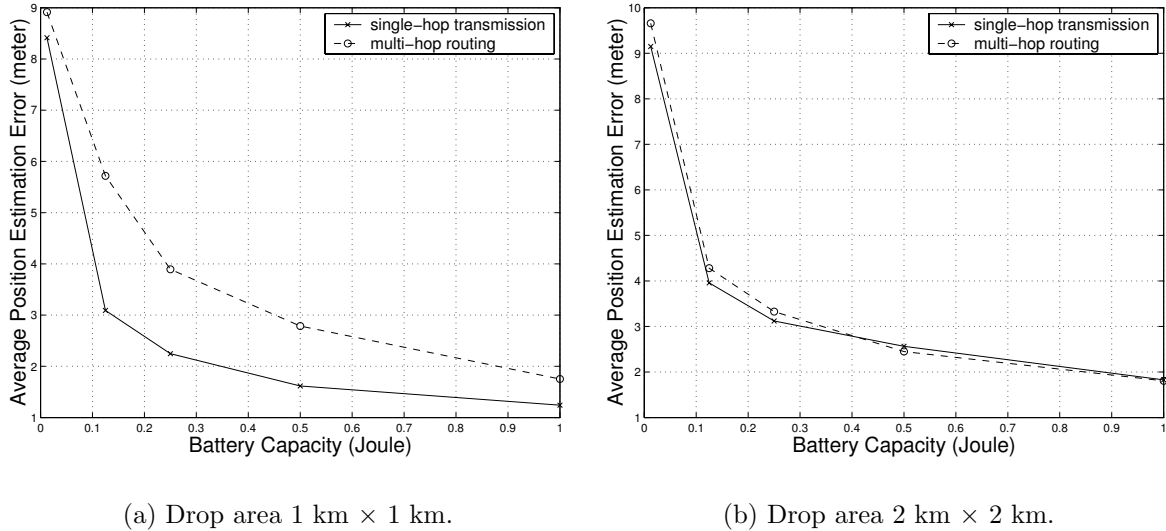


Figure 2.13: Performance comparison of different algorithms with different drop areas.

hop transmission protocol does better than the multi-hop routing protocol when the nodes are initially dropped in a $1 \text{ km} \times 1 \text{ km}$ area. When the nodes are dropped in a $2 \text{ km} \times 2 \text{ km}$ area, the multi-hop routing protocol does better than the single-hop transmission protocol for large battery capacities.

Many factors come into play to explain why single-hop transmission does better than multi-hop routing in the case of the smaller drop area. Among these, we mention: 1. propagation; 2. amplifier efficiency; 3. routing overhead. We briefly discuss each one of these factors:

1. Propagation:

Propagation loss becomes much less significant at small distance due to our propagation model, where the signal strength degrades proportionally to the

fourth power of the propagation distance. In particular, if the distance between a pair of nodes is increased from 1 km to 2 km, the signal attenuation will be increased by a factor of 16. Thus, as the distances between nodes becomes larger, the effect of propagation loss on performance grows dramatically.

2. Amplifier:

For small distances between nodes it is possible for the transmitter to reach all nodes in the network. Transmitting to nodes very close does not save much energy compared to transmitting to the farthest node since the amplifier will be operating in the region of low efficiency. Thus the energy consumed does not decrease proportionally to the decrease in desired output energy, as was discussed in the example in Section 2.1. So by routing messages through very close nodes the amount of energy saved does not increase proportionally to the decrease in power loss from transmission.

3. Routing Overhead:

In multi-hop routing a certain amount of energy is needed to update routing tables, which is not needed in single-hop transmission.

In summary, all these effects play a role in determining the overall system performance shown in Figure 2.13. While we have made these conclusions for a very specific set of parameters and models, we believe that when the nodes are close to each other single-hop transmission is generally more efficient than multi-hop routing. What changes is the threshold (in terms of distance) where one strategy is better than the other. Finally, it is of interest to understand where energy is being consumed in these systems. Within the context of the single-hop transmission protocol, about 30% of the totally consumed energy is used for transmitting packets in both

cases of drop areas ($1 \text{ km} \times 1 \text{ km}$ and $2 \text{ km} \times 2 \text{ km}$), and the other 70% is used for receiving packets. Within the context of the multi-hop routing protocol, about 25% of the totally consumed energy is used for transmitting both position information packets *and* routing protocol packets when the drop area is $1 \text{ km} \times 1 \text{ km}$; this percentage changes to about 35% when the drop area is $2 \text{ km} \times 2 \text{ km}$. When the multi-hop routing protocol is in use, depending on the different energy constraints, about 40-70% of the total consumed energy is spent on transmitting *and* receiving packets that are only used for updating routing tables. The remaining energy is used for transmitting and receiving position information packets.

2.5.2 Merits of Integrated Design

To determine the merits of the integrated design (ID) methodology, we compare its performance with two different traditional designs applied to the single-hop transmission protocol. In the first design, called AD-1, the optimization of the processing layer is done independently of the optimization at the network layer, but the network layer uses the results of the optimization at the processing layer. We can view this as a one way coupling between the processing layer and the network layer. In the second design, called AD-2, the two layers are designed totally independently.

Alternative Design 1

In alternative design 1 (AD-1), we partially decouple the optimization by imposing a constraint on the packet error probability for the transmission between the two most distant nodes. Let SNR_f be the signal-to-noise ratio between the two most distant nodes. In our scenarios, this is when the nodes are at the two diagonally opposite corners of the region under consideration. We first consider the following

optimization problem at the processing layer:

$$[\hat{E}_{ct}, \hat{E}_{cr}]^T = \arg \min_{\substack{[E_{ct}, E_{cr}]^T \\ P_e(E_{ct}, E_{cr}, SNR_f) \leq 10^{-2}}} E_{ct} + E_{cr}. \quad (2.33)$$

The goal in this first step is to minimize the total energy (transmitter and receiver) needed for the longest possible transmission distance in order to maintain a packet error probability of 0.01. The next step is to optimize the performance metric given by (2.21) over the network parameters using the results of the optimization at the processing layer design. Specifically, the goal is to determine

$$[\tilde{T}, \tilde{q}]^T = \arg \min_{\substack{[T, q]^T \\ \max E^{(i)} \leq E}} J(T, q, \hat{E}_{ct}, \hat{E}_{cr}). \quad (2.34)$$

In order to compare AD-1 with the integrated design approach, recall (2.23) for the integrated design, which we restate here for convenience:

$$[T^*, q^*, E_{ct}^*, E_{cr}^*]^T = \arg \min_{\substack{[T, q, E_{ct}, E_{cr}]^T \\ \max E^{(i)} \leq E}} J(T, q, E_{ct}, E_{cr}). \quad (2.35)$$

Comparing (2.35) with (2.34) and (2.33) reveals that

$$J(T^*, q^*, E_{ct}^*, E_{cr}^*) \leq J(\tilde{T}, \tilde{q}, \hat{E}_{ct}, \hat{E}_{cr}); \quad (2.36)$$

thus the performance of the integrated design is at least as good as AD-1.

Numerical results illustrating the comparison of AD-1 with ID are shown in Figure 2.14, within the context of mobility model 2. These results show a degradation in performance when AD-1 is used as compared with ID. In particular, there is a degradation of about 0.8 dB in energy when the average position estimation error is in the range of 60-100 m.

Alternative Design 2

We now consider alternative design 2 (AD-2). In this design we completely decouple the optimization at the network and processing layers. In AD-2, we proceed

as follows: (i) we optimize E_{ct} and E_{cr} as in AD-1; and (ii) we select the parameter values $T^o = 60$ s and $q^o = 0.01$ at the network layer, without doing any optimization. Design AD-2 is consistent with many traditional design methodologies where network layer parameters are selected based on engineering and application considerations, without any explicit optimization that would account for the processing and device layers. Therefore the performance of AD-2 is $J(T^o, q^o, E_{ct}^*, E_{cr}^*)$, and it is clear that

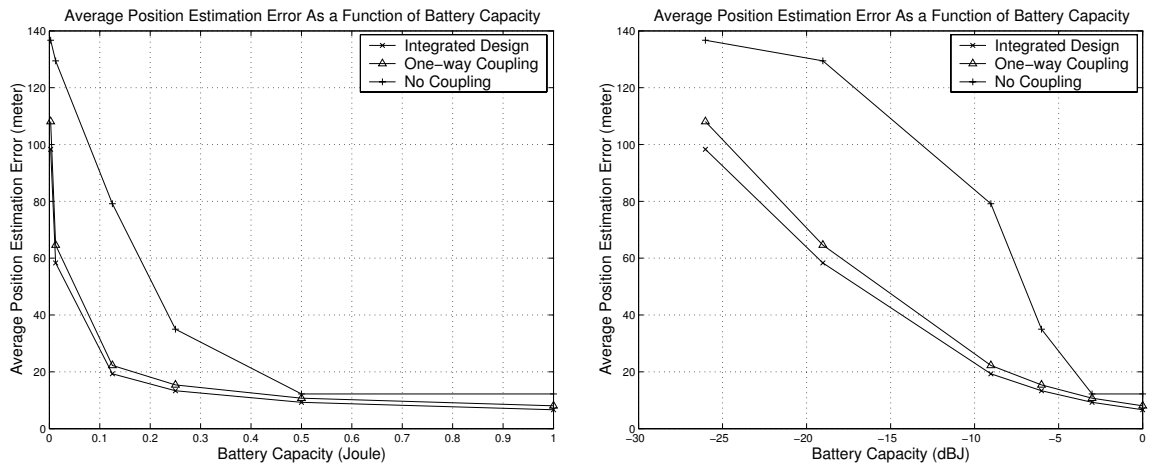
$$J(T^*, q^*, E_{ct}^*, E_{cr}^*) \leq J(\tilde{T}, \tilde{q}, \hat{E}_{ct}, \hat{E}_{cr}) \leq J(T^o, q^o, \hat{E}_{ct}, \hat{E}_{cr}). \quad (2.37)$$

Thus the performance of AD-2 is the worst among the three designs.

Numerical results illustrating the performance of AD-2 are also shown in Figure 2.14. The degradation in performance of AD-2 as compared with ID is quite significant. This is evident when we plot the energy in dB units (with a reference of 1 Joule). In particular, about 14 dB more energy is required to obtain the same average position estimation error when the average position estimation error is in the range of 60-100 m. However, this observation is for one scenario only and should not be generalized. Clearly, more simulation experiments, for a range of scenarios, mobility models and channel models, are needed to quantify more completely the benefits of our integrated design and optimization strategy.

2.6 Conclusion and Future Research

We have proposed an *integrated design* methodology and applied it to the optimization of the situational awareness problem in ad hoc mobile wireless networks. We have given evidence (presented in Section 2.5.2) why the integrated design methodology outperforms other design methodologies that do not account for or exploit coupling among layers. This evidence is supported by simulation experiments. In



(a) Normal scale.

(b) Log scale in battery capacity.

Figure 2.14: Performance comparison of different design methodologies.

future research, it would be of interest to classify other cases where an *integrated design* approach leads to large performance gains over traditional approaches.

CHAPTER III

ROUTING IN BROADCAST MODE

An ad hoc mobile wireless network (also called packet radio network, dynamic network, etc) is a collection of mobile platforms (usually called nodes) that are dynamically and arbitrarily located without any fixed infrastructure. The communication channel between each pair of mobile platforms in the network may be changing on a continual basis due to the mobility of each platform and the environment. Each mobile platform in such a network may function as a router or as an end transmitter or receiver. Examples of such networks include disaster relief operations where a temporary network is needed for rescuers, meetings or conventions where people want to share information quickly, data acquisition in a hostile environment where physically setting fixed connection between sensors is impossible, and military operations where small mobile units communicate their situational awareness data.

3.1 Motivation for Routing in Ad Hoc Mobile Wireless Networks

Because the degradation of signal strength can be very significant when an electromagnetic wave propagates through a wireless environment, a mobile platform may not have enough power to send its packets to another mobile platform in just a single hop such that the signal to noise ratio at the receiver is acceptable. For example, the

signal strength is usually degraded with the square of the propagation distance in free space, while if there is a ground reflecting plane, the signal strength degrades with the fourth power of the propagation distance. Because of the signal degradation, packets may need to be routed through the network in order for them to reach their final destinations with certain quality. Routing through multiple hops may save a mobile platform a significant amount of energy since the total energy consumed by each mobile platform involved in routing can be much less than the energy required by a mobile platform to send its packets to a destination in just a single hop, as demonstrated in Section 2.5. As can be seen in Figure 2.13, routing indeed gives us some benefit when the average distances between nodes are large. Routing may also allow us to configure wireless networks in a more flexible and robust fashion so that a failure of a single mobile platform does not affect the operation of a network significantly. In short, routing will play an important role in ad hoc mobile wireless networks as more and more wireless applications grow popular.

In general, finding the best routing protocol of a complex communication system can be too ambitious a task, especially when we want to design it analytically. Even with the help of simulation-based optimization, finding an optimal routing protocol over all parameters across different layers is still a formidable task. Therefore we make simplifications by considering the routing problems over one parameter, namely, the power assigned to each mobile platform to transmit a packet. We will also take advantage of the omni-directional antenna carried by each node.

The remainder of this chapter is organized as follows. In Section 3.2 we present a survey of wireless routing protocols. In Section 3.3 we give several definitions for wireless networks. In Section 3.4 we specify the problems of routing in broadcast mode. In Section 3.5 we list some previous research related to the routing problems

specified in 3.4. In Section 3.6 we present our results for the routing problems in broadcast mode specified in 3.4. We conclude the discussion and indicate future research in Section 3.7.

3.2 Current Wireless Network Routing Protocols

There have been many proposed wireless network routing protocols [52, 71], which can be divided into two broad categories: **table-driven** and **on-demand**, based on when and how routes are discovered and maintained. In **table-driven** routing protocols, consistent and up-to-date routing information to all other nodes are maintained at each node, whereas in **on-demand** routing protocols, routes are created only when they are needed. Figure 3.1 categorizes the current proposed routing protocols whose acronyms are explained in Table 3.1. For details of each routing protocol, please refer to its corresponding reference.

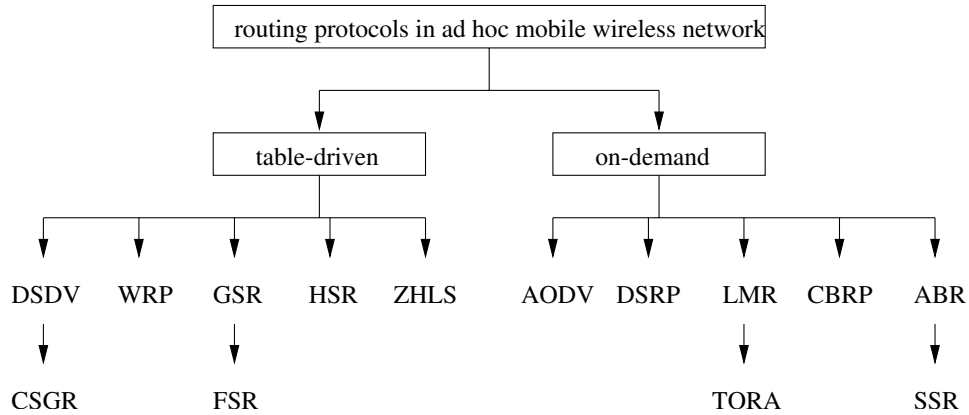


Figure 3.1: Categorization of routing protocols in ad hoc mobile wireless networks.

Ad-hoc network protocols are generally described using a layered structure. This structure may include the application layer, the network layer, the medium access control layer, and the physical layer. The routing protocols in Figure 3.1 function at the network layer, where each of them assumes that there exists a medium access

Table 3.1: List of current wireless routing protocols.

DSDV [61]	Dynamic Destination-Sequenced Distance-Vector Routing Protocol
CGSR [17]	Clusterhead Gateway Switch Routing Protocol
WRP [55]	The Wireless Routing Protocol
GSR [16]	Global State Routing
FSR [39]	Fisheye State Routing
HSR [39]	Hierarchical State Routing
ZHLS [34, 42]	Zone-based Hierarchical Link State Routing Protocol
AODV [62]	Ad Hoc On-demand Distance Vector Routing
DSRP [43, 44]	Dynamic Source Routing Protocol
LMR [20]	Lightweight Mobile Routing
TORA [60]	Temporally Ordered Routing Algorithm
CBRP [41]	Cluster Based Routing Protocols
ABR [82, 83]	Associativity Based Routing
SSR [24]	Signal Stability Routing

control (MAC) layer that reports to the network layer the state of the connections between each pair of mobile platforms. The routing protocols in an ad hoc mobile wireless network mostly deal with finding routing information and maintaining routes. Since *point-to-point* connections are always assumed in the ad-hoc networks for the routing algorithms in Figure 3.1, which resemble the connections in wired networks, these routing protocols can still use the basic algorithms that are used in wired networks, such as Bellman-Ford algorithms or Dijkstra's algorithm, to compute the shortest *path* (where the length of a path is usually defined as delay or number of hops) from a source to a destination. If nodes in an ad-hoc network have directional antennae, there is a very close relation between wired networks and wireless networks. However with the omni-directional antennae, connections can be point-to-multipoint and this makes wireless networks significantly different from wired networks. Consequently we must view routing problems in ad hoc mobile wireless networks very differently from the ones in wired networks.

3.3 Definitions

Conventionally an ad hoc mobile wireless network is modeled as a graph where two vertices have a directed (or an undirected) edge if and only if the corresponding mobile platforms can communicate from one to another (or between the two). However this representation of wireless networks hides many properties that are characteristic of wireless communications and is sometimes not very realistic. We will represent ad hoc mobile wireless networks in a different way in order to capture the key properties of wireless communications and describe wireless network routing in a regime of networks with point-to-multipoint connections (which we call **routing in broadcast mode**). But first, we need a few definitions.

Definition 3.3.1 A **node** is a mobile platform. The set \mathcal{N} of all nodes of interest is called a **node set**. The **cardinality** of a set \mathcal{N} is the number of elements in the set \mathcal{N} and is denoted by N .

Definition 3.3.2 A **location function** $\underline{\mu} : \mathcal{N} \rightarrow \mathbb{R}^3$ is a mapping such that $\underline{\mu}(n) = (x_n, y_n, z_n)^T \in \mathbb{R}^3 \forall n \in \mathcal{N}$, where $(x_n, y_n, z_n)^T$ is the location of node n .

Definition 3.3.3 An **adjustable parameter function** $\underline{\theta} : \mathcal{N} \rightarrow \mathbb{R}^M$ is a mapping such that $\underline{\theta}(n) = (\theta_n^{(1)}, \theta_n^{(2)}, \dots, \theta_n^{(M)})^T \in \mathbb{R}^M \forall n \in \mathcal{N}$, where $(\theta_n^{(1)}, \theta_n^{(2)}, \dots, \theta_n^{(M)})^T$ are adjustable parameters of node n .

Examples of adjustable parameters include interval between transmissions, coding and modulation techniques, transmitting power, etc. In this chapter we restrict our attention to the transmitting power and use $\eta : \mathcal{N} \rightarrow \mathbb{R}$ to represent a power assignment to each $n \in \mathcal{N}$, i.e., for each node $n \in \mathcal{N}$, $\eta(n)$ is the power that node n should use to transmit its packets.

Definition 3.3.4 An **environment function** $\psi : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ is a mapping such that $\psi((x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T) = r \in \mathbb{R}$, where r characterizes the properties of propagation from location $(x_1, y_1, z_1)^T$ to location $(x_2, y_2, z_2)^T$.

Examples of environment function include path loss function, shadowing, multipath fading, etc. In this chapter we restrict our attention to path loss, which is usually given by (2.30). In this case, the environment function will be

$$\psi((x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T) = \frac{P_{out}}{P_r} = \left[4 \left(\frac{\lambda_c}{4\pi d} \right)^2 G_t G_r \sin^2 \left(\frac{2\pi h_t h_r}{\lambda_c d} \right) \right]^{-1} \quad (3.1)$$

where $(x_1, y_1, z_1)^T$ is the location of transmitter, $(x_2, y_2, z_2)^T$ is the location of receiver, and

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \quad (3.2)$$

A **symmetric channel** satisfies

$$\psi((x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T) = \psi((x_2, y_2, z_2)^T, (x_1, y_1, z_1)^T) \quad (3.3)$$

Definition 3.3.5 An **ad hoc mobile wireless network** is a set $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \underline{\theta}, \underline{\psi})$, where \mathcal{N} , $\underline{\mu}$, $\underline{\theta}$, and $\underline{\psi}$ are defined in Definitions 3.3.1, 3.3.2, 3.3.3, and 3.3.4, respectively.

We consider source-initiated multicast sessions in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \underline{\theta}, \underline{\psi})$, where every node is permitted to initiate multicast sessions, namely, every node is allowed to transmit to any subset of nodes in \mathcal{N} . We assume that each node $n \in \mathcal{N}$ has an omni-directional antenna and we say that \mathcal{M} is in **broadcast mode** under this condition. When the data rate and noise power spectral density are fixed, the connectivity of \mathcal{M} depends on the power assignment function $\eta : \mathcal{N} \rightarrow \mathbb{R}$, where $\eta(j)$ is the power assigned to node j . In actuality, the

power assignment function corresponds to the situation where each node can choose its power level not exceeding the maximum power it can use. We assume that each node $n \in \mathcal{N}$ has a fixed **receiver sensitivity** S_n , i.e, if the received power P_r at node n satisfies $P_r \geq S_n$, we assume perfect reception at node n , otherwise, we assume a packet is lost at node n . We call this a **threshold receiver model**. Therefore **the least power function** $\beta : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$, which indicates the least power that a transmitter at location $(x_1, y_1, z_1)^T \in \mathbb{R}^3$ should use to ensure correct reception by a receiver at location $(x_2, y_2, z_2)^T \in \mathbb{R}^3$, is given by

$$\beta((x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T) = \psi((x_1, y_1, z_1)^T, (x_2, y_2, z_2)^T) S_2 \quad (3.4)$$

Thus it is possible for a node $n \in \mathcal{N}$ to reach a set of other nodes in \mathcal{N} with just one transmission, all we need is to assign the power level of node n the maximum of the power required for n to reach any one of the nodes in the set individually. In contrast, we don't have this possibility in wired networks where connections are composed of point-to-point links. Under the above assumptions on \mathcal{M} , we want to find solutions to the following six problems. Most importantly, we want to design **efficient algorithms** for finding such solutions, if possible, where **efficient algorithms** is defined to be the ones that have complexity polynomial in N .

3.4 Problems of Routing in Broadcast Mode

We formulate the problems of routing in broadcast mode into the following problems.

Problem 3.4.1 *Given two nodes n_0 and n_1 in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to n_1 (possibly in multiple hops) and $\max_{1 \leq n \leq N} \eta(n)$*

is minimized.

Problem 3.4.2 Given a node n_0 in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to every node in $\mathcal{N} \setminus \{n_0\}$ (possibly in multiple hops) and $\max_{1 \leq n \leq N} \eta(n)$ is minimized.

Problem 3.4.3 Given a node n_0 and a subset \mathcal{Q} of nodes in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to every node in \mathcal{Q} (possibly in multiple hops) and $\max_{1 \leq n \leq N} \eta(n)$ is minimized.

Problem 3.4.4 Given two nodes n_0 and n_1 in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to n_1 (possibly in multiple hops) and $\sum_{n=1}^N \eta(n)$ is minimized.

Problem 3.4.5 Given a node n_0 in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to every node in $\mathcal{N} \setminus \{n_0\}$ (possibly in multiple hops) and $\sum_{n=1}^N \eta(n)$ is minimized.

Problem 3.4.6 Given a node n_0 and a subset \mathcal{Q} of nodes in an ad hoc mobile wireless network $\mathcal{M} = (\mathcal{N}, \underline{\mu}, \eta, \underline{\psi})$, find a power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ such that the packet generated by node n_0 can be correctly sent to every node in \mathcal{Q} (possibly in multiple hops) and $\sum_{n=1}^N \eta(n)$ is minimized.

We see that Problem 3.4.1 and 3.4.2 are special cases of Problem 3.4.3 by letting $\mathcal{Q} = \{n_1\}$ and $\mathcal{Q} = \mathcal{N} \setminus \{n_0\}$, respectively. Similarly, Problem 3.4.4 and 3.4.5 are

special cases of Problem 3.4.6. Both types of optimization in the problems 3.4.1 to 3.4.6 find their applications in actual wireless networks. The argument for seeking solutions to Problem 3.4.1 through 3.4.3 is that battery life is usually a local resource so that the minimization of total consumed power for multicasting one packet has little practical value. However, we may also note that the solutions to Problem 3.4.1 through 3.4.3 may have their own drawback in that they may use more nodes involved in routing packets from a source to a destination, thus may cause more delay. An example of this situation is shown in Figure 3.2, where node 1 wants to send packets to node 3. This example shows that solutions to Problem 3.4.4 is more appropriate in terms of delay.

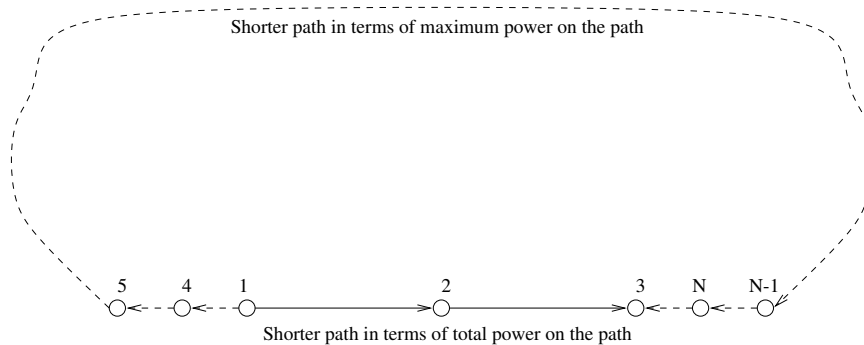


Figure 3.2: Example of inappropriateness of solution to Problem 3.4.1.

Because of the threshold receiver model, we note that the power assignment $\eta(n)$ for each node $n \in \mathcal{N}$ need only be one of N values, where these N values are: no power for transmission, least power required to reach the first nearest node, least power required to reach the second nearest node, etc, as given in (3.4). Thus there are a total of N^N functions of power assignment that we can enumerate to find whether a certain power assignment η is a solution to Problem 3.4.3 or Problem 3.4.6. Even though this is an algorithm for finding solutions to our problems, it requires a computational complexity exponential in N . Therefore our main objective is to

design efficient algorithms that find solutions to Problem 3.4.3 and Problem 3.4.6 respectively. We hope that these efficient algorithms, if exist, will play the same important role in ad hoc mobile wireless networks with *point-to-multipoint* connections, as their counterparts, such as Bellman-Ford algorithm, Dijkstra algorithm, and Floyd-Warshall algorithm, in wired networks with *point-to-point* connections.

3.5 Previous Research

There have been some related previous research papers on problems stated in Section 3.4, even though they do not quite address the same problems described in Section 3.4.

3.5.1 Previous Results Related to Problem 3.4.1 through 3.4.3

In [67] the authors present a polynomial-time algorithm for finding a solution to a special case of Problem 3.4.2, where it is assumed that each pair of nodes in \mathcal{N} see a symmetric channel, where the channel between node i and node j is symmetric if the degradation from node i to node j is the same as the degradation from node j to i . The algorithm in that paper finds a power assignment such that the maximum of the power used by all nodes in the ad hoc mobile wireless network is minimized under the condition that the underlying graph after the power assignment is **strongly connected**, namely, between any pair of nodes i and j , there is a directional path going from i to j and there is a directional path going from j to i . Note that in Problem 3.4.2, we don't have the assumption that each pair of nodes see a symmetric channel and we don't require the underlying graph to be strongly connected. Thus our problem is more general than the one solved in [67].

3.5.2 Previous Results Related to Problem 3.4.4 through 3.4.6

In [87] the authors state Problem 3.4.5 in a slightly different way and seek an approximate solution to Problem 3.4.5, since the an exact efficient algorithm can not be found by the authors. In [19, 45], however, the authors show that if an assumption of strong connectivity is made on the underlying graph after the power assignment in Problem 3.4.5, finding a exact solution to Problem 3.4.5 in 2-dimensional or 3-dimensional space will be NP-hard, but there is an efficient algorithm for finding a solution to Problem 3.4.5 in 1-dimensional space. We will link these results together in Section 3.6 to draw conclusions for Problem 3.4.4 through 3.4.6.

3.6 Some Results of Routing in Broadcast Mode

We consider a quasi-static network where the channel between each pair of nodes in an ad hoc mobile wireless network does not change or change very slowly with respect to route information updating and route computation. We also assume centralized route computation, where each node has the same view of the topology of the network. We want to design efficient algorithms for finding the solutions to Problem 3.4.1 through 3.4.6, if possible. We first deal with Problem 3.4.1 through 3.4.3. Since Problem 3.4.1 and 3.4.2 are special cases of Problem 3.4.3, we will just describe the algorithm for finding a solution to Problem 3.4.3.

3.6.1 Algorithm for Problem 3.4.1 through 3.4.3

We consider the underlying complete graph \mathcal{C} of an ad hoc mobile wireless network, where the weight of each edge is formed by the least power function between each pair of nodes. We will use a modified Dijkstra's algorithm in the description of the algorithm for Problem 3.4.3, even though any shortest path algorithm can be

modified in the same way and used in the algorithm for Problem 3.4.3. In the modified Dijkstra's algorithm, the weight of a path is redefined as the maximum of the weight of each edge on that path, instead of the usual definition as the cumulative weight of edges on that path [8, 11]. The basic idea of the algorithm given below is that we first build a minimum weight path tree from a source to all destinations in the complete graph and then assign power to each node considering only the nodes on the tree from the source to the set of desired destinations. As can be seen later, this will provide us with an efficient power assignment algorithm.

Algorithm 3.6.1 (Algorithm for Problem 3.4.3) *Let \mathcal{N} be the set of nodes in an ad hoc mobile wireless network \mathcal{M} . Without loss of generality, we assume that node 1 is the sender and $\mathcal{Q} := \{m_1, m_2, \dots, m_t\} \subseteq \mathcal{N} \setminus \{1\}$ is the set of destination nodes to which node 1 wants to send packets. Let $D : \mathcal{N} \rightarrow \mathbb{R}$ be a path weight function, i.e., the weight of a path from node 1 to node i is $D(i)$. Let \mathcal{S} be a set to record the nodes whose shortest path from node 1 have been found. Let $U : \mathcal{N} \rightarrow \mathcal{N}$ be an upstream-node function, i.e., the upstream-node of node i on the path from node 1 to \mathcal{Q} is $U(i)$. Let $F : \mathcal{N} \rightarrow 2^{\mathcal{N}}$ be a first-hop function, i.e., $F(i)$ is the set of first-hop nodes that node i should use in order to forward packets from node 1 to the nodes in \mathcal{Q} . Let $P : \mathcal{N} \rightarrow \mathbb{R}$ be a power assignment function, i.e., $P(i)$ is the power that node i should use in order to forward packets from node 1 to the nodes in \mathcal{Q} . The algorithm for Problem 3.4.3 runs over the following steps:*

1. *Initialization*

For each $i \in \mathcal{N}$, calculate the weight $d_{i,j}$ (which is the least power required for node i to successfully send packets to node j based on the least power function) to all nodes $j \in \mathcal{N} \setminus \{i\}$. Let $D(1) = 0$, $D(j) = d_{1,j}$ for $j \neq 1$, $U(j) = 1$ for $j \neq 1$, $F(j) := \emptyset$ for $j \in \mathcal{N}$, $P(j) = 0$ for $j \in \mathcal{N}$, $\mathcal{S} := \{1\}$, $\mathcal{A} := \{1\}$, $\mathcal{W} := \mathcal{Q}$.

2. *Building shortest path tree from node 1 to all nodes in $\mathcal{N} \setminus \{1\}$*

(a) *(Finding the next closest node) Find $i \notin \mathcal{S}$ such that $D(i) = \min_{j \notin \mathcal{S}} D(j)$.*

Set $\mathcal{S} := \mathcal{S} \cup \{i\}$. If $\mathcal{S} = \mathcal{N}$, stop and the shortest path tree is found.

(b) *(Updating weights) For all $j \notin \mathcal{S}$, let $D(j) = \min(D(j), \max(d_{i,j}, D(i)))$.*

If $D(j) > \max(d_{i,j}, D(i))$, let $U(j) = i$. Go to Step 2a.

3. *Building multicast tree from node 1 to \mathcal{Q}*

(a) *If $\mathcal{W} = \emptyset$, stop and the multicast tree \mathcal{T} is found.*

Else, choose $i \in \mathcal{W}$ and let $j = i$.

i. Set $F(U(i)) := F(U(i)) \cup \{i\}$.

ii. If $U(i) \neq 1$, set $i := U(i)$ and go to Step 3(a)i. Else, stop and a path from node 1 to one of nodes in \mathcal{Q} is found.

(b) *$\mathcal{W} := \mathcal{W} \setminus \{j\}$. Go to Step 3a.*

4. *Power Assignment*

(a) *If $\mathcal{A} = \emptyset$, stop and power assignment is done.*

Else, choose $i \in \mathcal{A}$ and let $\mathcal{A} := \mathcal{A} \setminus \{i\}$.

(b) *If $F(i) \neq \emptyset$, set $P(i) = \max_{j \in F(i)}(d_{i,j})$ and set $\mathcal{A} := \mathcal{A} \cup F(i)$. Go to Step 4a.*

Here is a concrete example to demonstrate in detail how the algorithm for Problem 3.4.3 works. It should be noted that the algorithm does not require symmetric channels and the underlying graph for the final power assignment of the ad hoc mobile wireless network is not necessarily strongly connected, as can be seen in the solution to the following example.

Example 3.6.1 *A group of nodes form an ad hoc mobile wireless network as shown in Figure 3.3 with node 1 as the sender and $\mathcal{Q} := \{3, 4, 7, 9, 10\}$ as the set of destination nodes. Table 3.2 shows the initialization steps, with $d_{i,j}$ the least power required for node i to successfully send packets to node j . Note that some of the channels between pairs of nodes are not symmetric, such as $d_{1,2}$ and $d_{2,1}$. Figure 3.3 through 3.8 show each step of the algorithm while running. When the algorithm terminates, the maximum power after the power assignment is 1024.*

Table 3.2: Weight between each pair of nodes in Example 3.6.1.

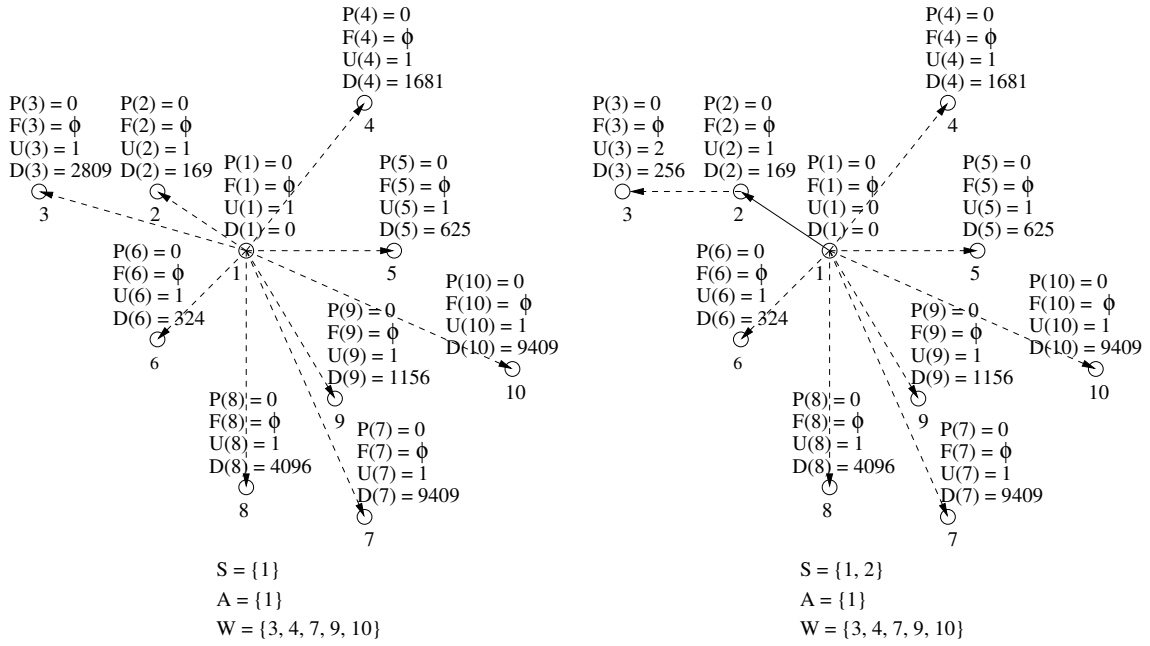
$d_{i,j} \setminus$	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$	$j = 10$
$i = 1$	0	169	2809	1681	625	324	9409	4096	1156	9409
$i = 2$	961	0	256	3364	4624	625	28900	11881	7225	32400
$i = 3$	2809	652	0	16900	21904	1681	58564	22201	22201	85264
$i = 4$	1681	3364	16900	0	676	12769	38416	34225	10201	11236
$i = 5$	625	4624	21904	676	0	5329	6724	7921	841	1024
$i = 6$	324	625	1681	12769	5329	0	7225	1156	1600	21025
$i = 7$	9409	28900	58564	38416	6724	7225	0	289	289	2500
$i = 8$	4096	11881	22201	34225	7921	1156	289	0	324	9409
$i = 9$	1156	7225	22201	10201	841	1600	289	324	0	1369
$i = 10$	9409	32400	85264	11236	1024	21025	2500	9409	1369	0

Theorem 3.6.1 (Correctness of Algorithm 3.6.1) *Algorithm 3.6.1 is an efficient algorithm that gives a solution to Problem 3.4.3 with complexity $O(N^2)$.*

Proof of Theorem 3.6.1: We claim that at the beginning of each Step 2a

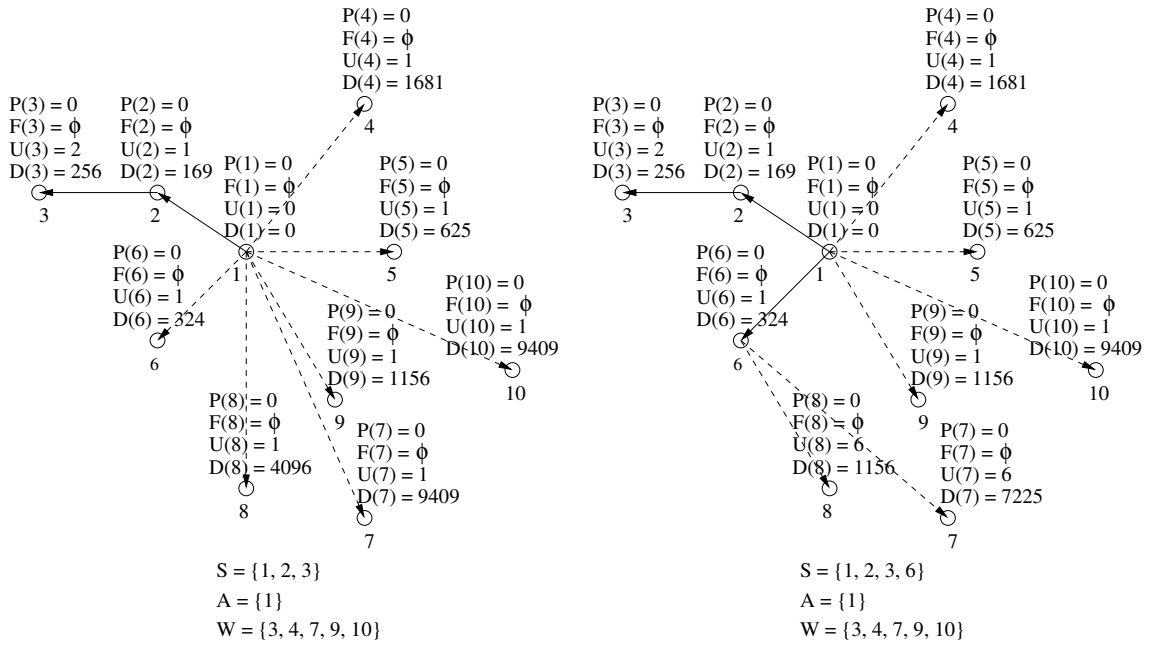
1. $D(l) \leq D(j)$ for all $l \in \mathcal{S}$ and $j \notin \mathcal{S}$.
2. $D(j)$ is, for each node j , the minimum weight from 1 to j using paths with all nodes, except possibly j , belonging to the set \mathcal{S} .

We show that Claim 1 holds by induction. Claim 1 holds at the initialization step, Step 1, of the algorithm because $\mathcal{S} = \{1\}$, $D(1) = 0$ and $D(j) = d_{1,j} \geq 0$ for



(a) 1: Initialization

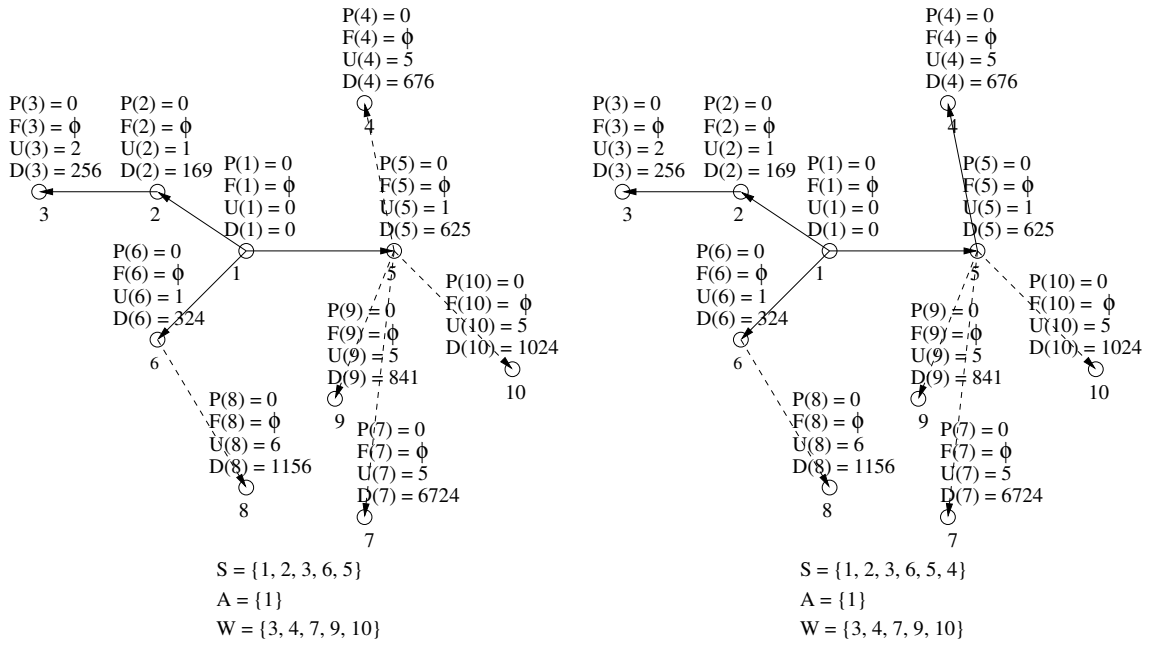
(b) 2: Step 2a, 2b



(c) 3: Step 2a, 2b

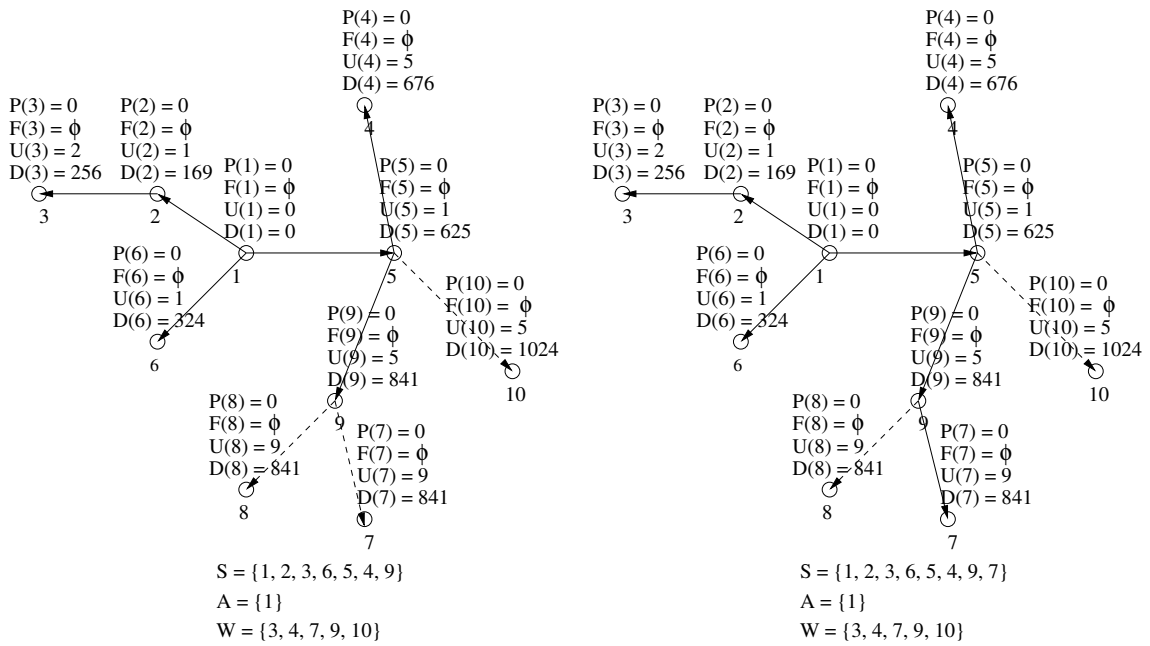
(d) 4: Step 2a, 2b

Figure 3.3: Demonstration of algorithm for Problem 3.4.3.



(a) 5: Step 2a, 2b

(b) 6: Step 2a, 2b



(c) 7: Step 2a, 2b

(d) 8: Step 2a, 2b

Figure 3.4: Demonstration of algorithm for Problem 3.4.3 (continued).

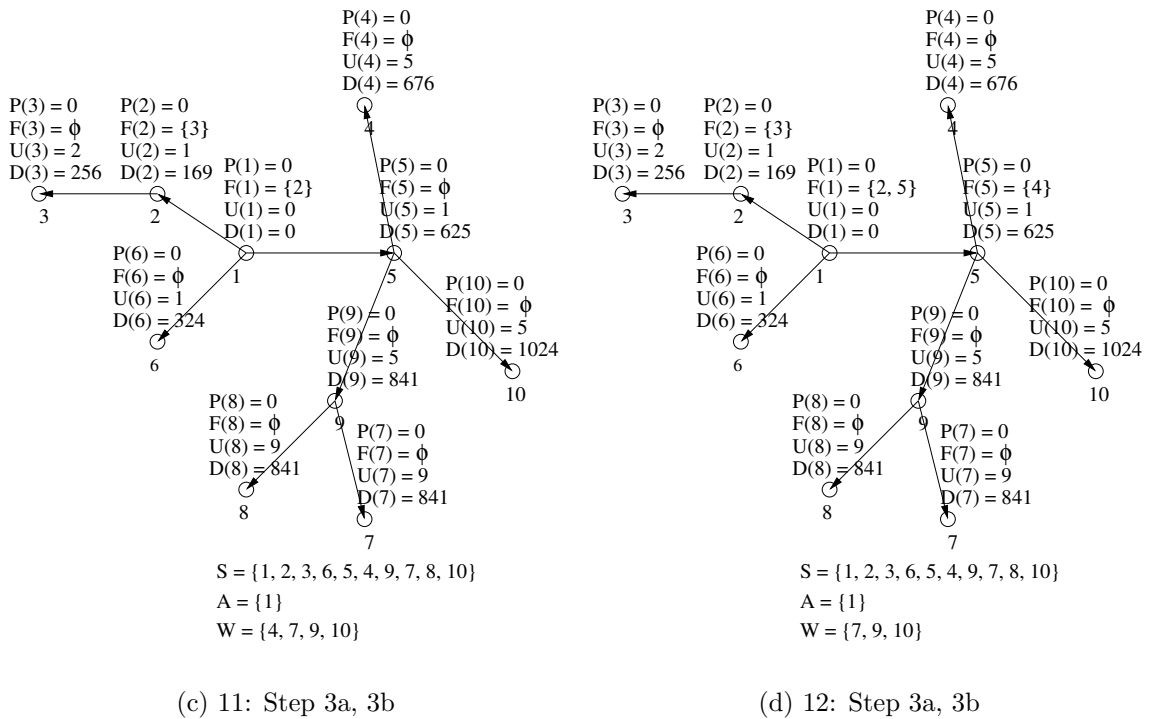
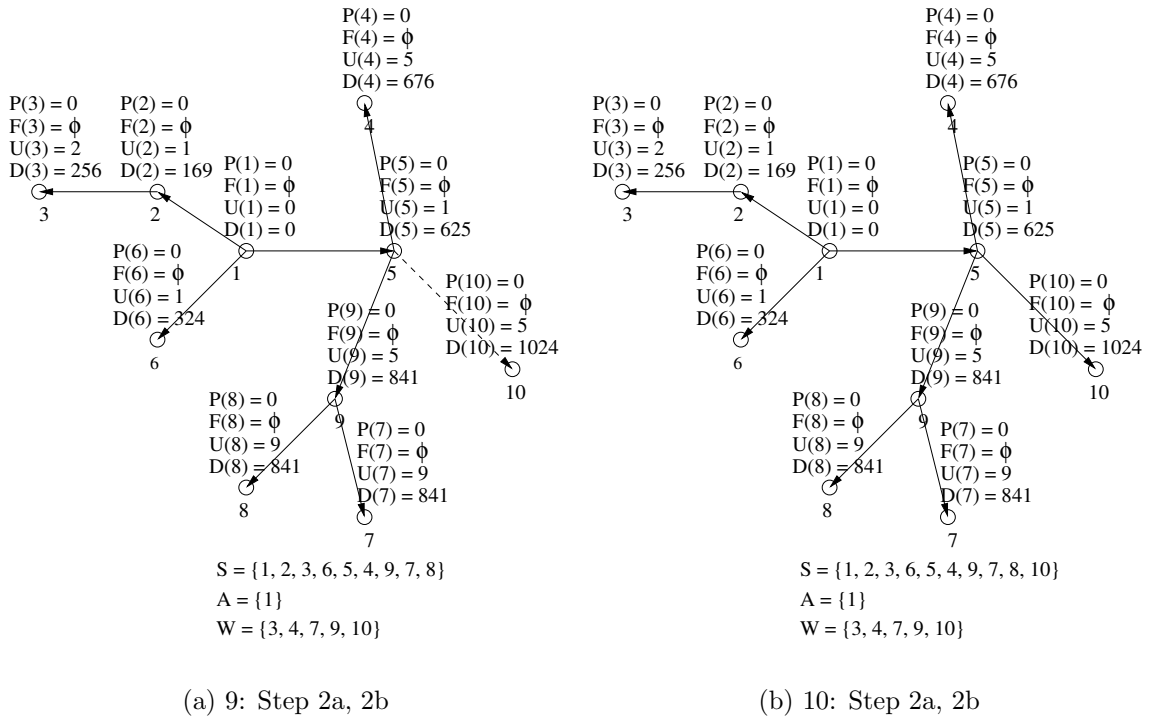
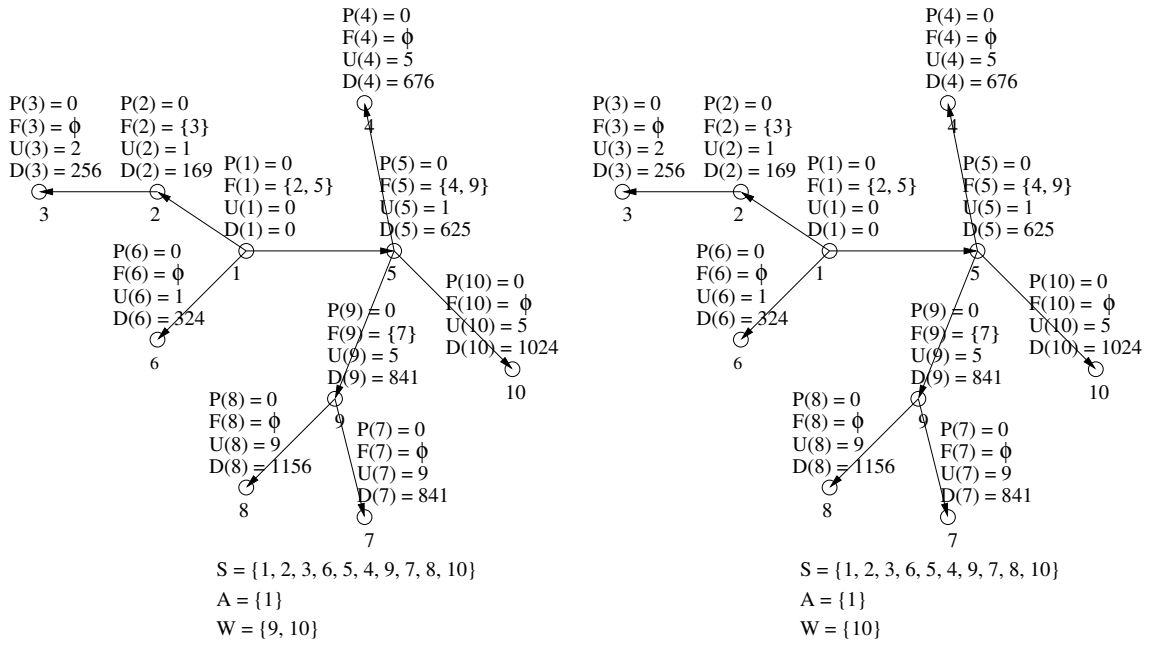
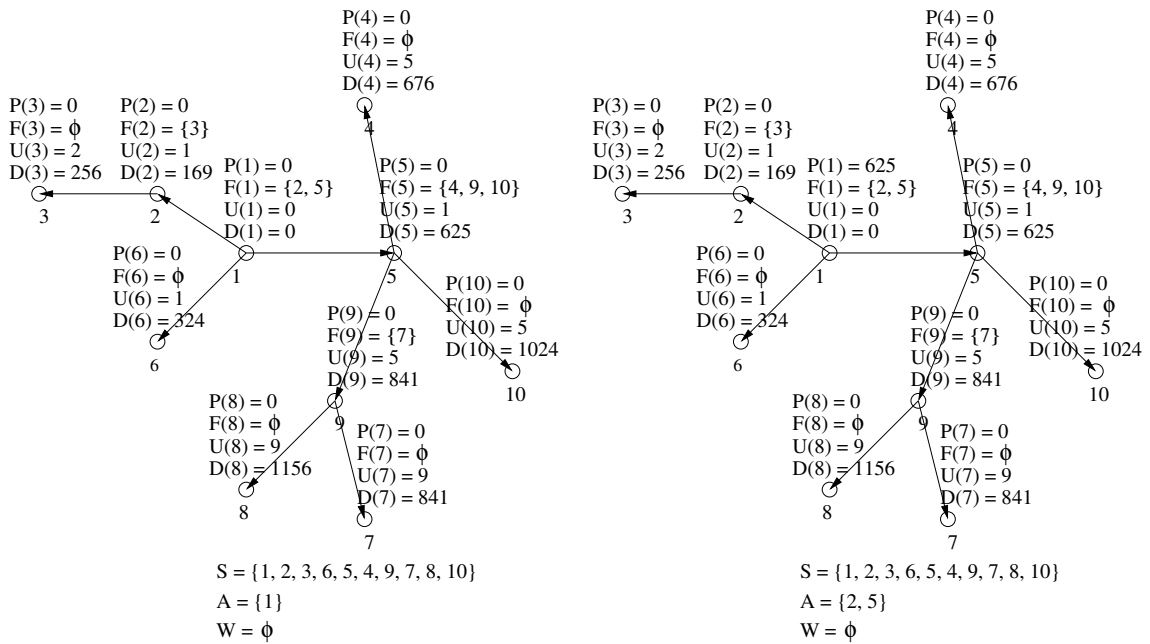


Figure 3.5: Demonstration of algorithm for Problem 3.4.3 (continued).



(a) 13: Step 3a, 3b

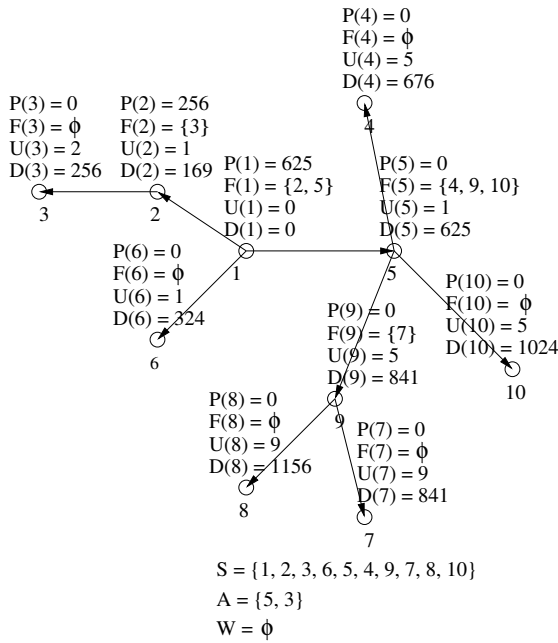
(b) 14: Step 3a, 3b



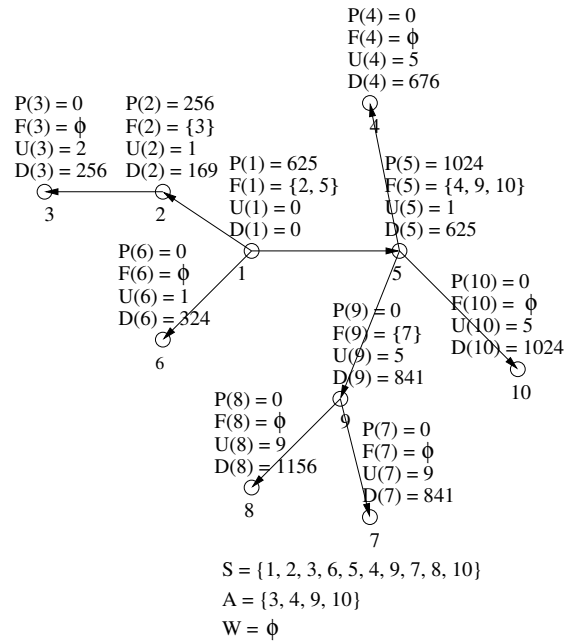
(c) 15: Step 3a, 3b

(d) 16: Step 4a, 4b

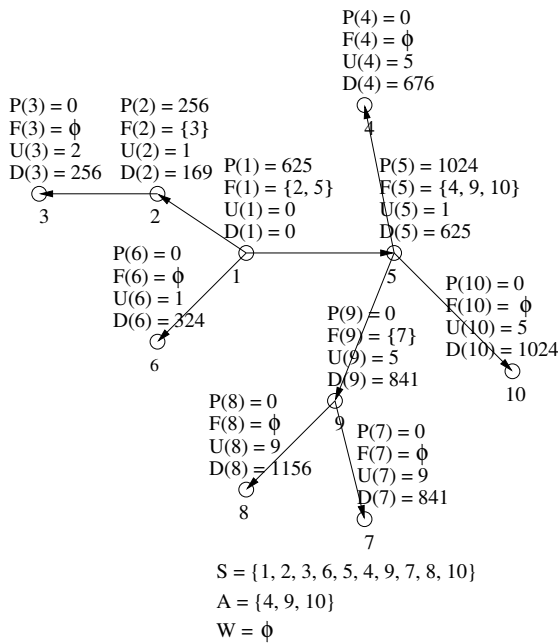
Figure 3.6: Demonstration of algorithm for Problem 3.4.3 (continued).



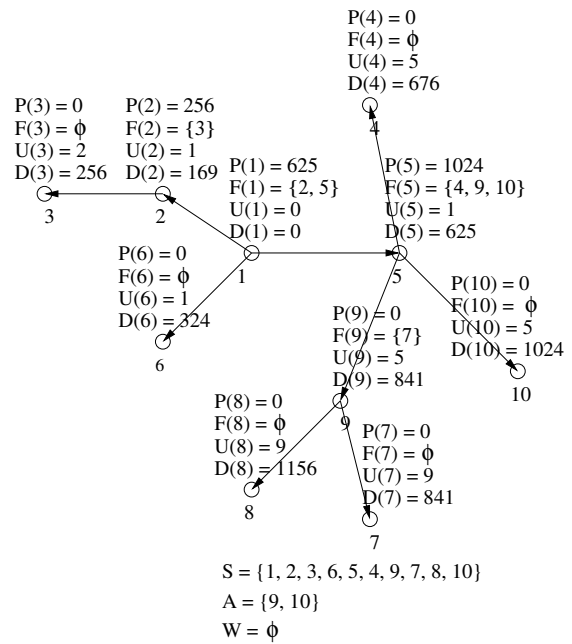
(a) 17: Step 4a, 4b



(b) 18: Step 4a, 4b

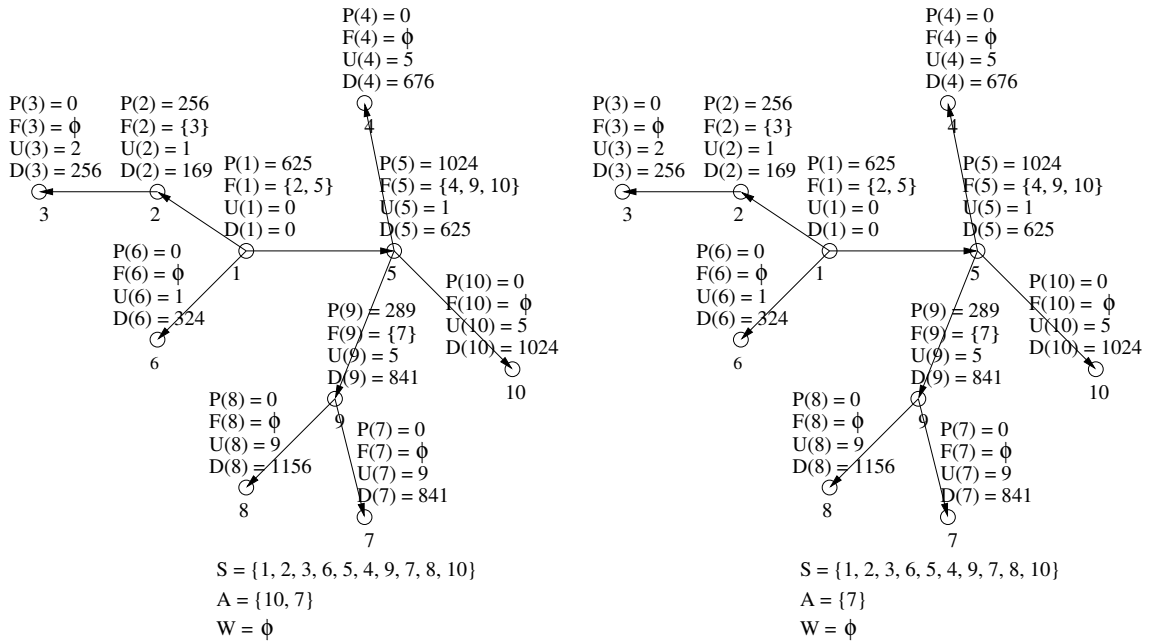


(c) 19: Step 4a, 4b



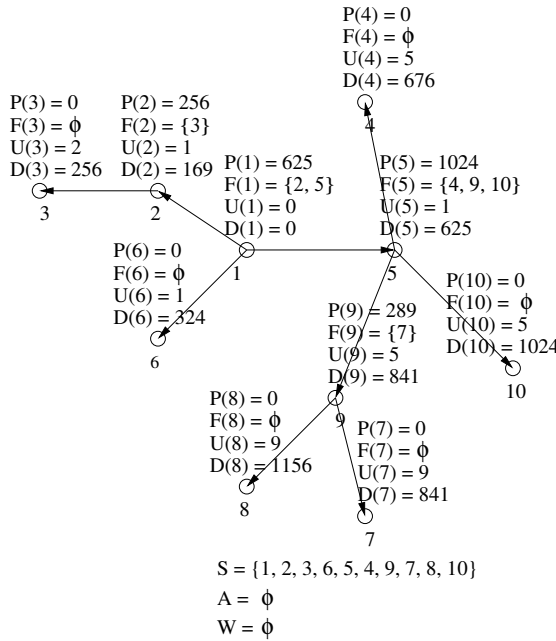
(d) 20: Step 4a, 4b

Figure 3.7: Demonstration of algorithm for Problem 3.4.3 (continued).



(a) 21: Step 4a, 4b

(b) 22: Step 4a, 4b



(c) 23: Step 4a, 4b

Figure 3.8: Demonstration of algorithm for Problem 3.4.3 (continued).

$j \neq 1$ when the algorithm is initialized. Now suppose Claim 1 holds at the beginning of Step 2a of some iteration, i.e., $D(l) \leq D(j)$ for all $l \in \mathcal{S}$ and $j \notin \mathcal{S}$, we show that at the end of Step 2b, i.e., at the beginning of Step 2a of next iteration, Claim 1 still holds. Since $D(i) = \min_{j \notin \mathcal{S}} D(j)$ in Step 2a, we have $D(i) \leq D(j)$ for all $j \notin \mathcal{S}$. By the induction hypothesis, we have $D(l) \leq D(i)$ for all $l \in \mathcal{S}$. After Step 2a is performed, we have $\mathcal{S} := \mathcal{S} \cup \{i\}$. Therefore for this newly updated \mathcal{S} , we have $D(l) \leq D(i)$ for all $l \in \mathcal{S}$ and $D(i) \leq D(j)$ for all $j \notin \mathcal{S}$. Consequently, after Step 2a, $D(l) \leq D(j)$ for all $l \in \mathcal{S}$ and $j \notin \mathcal{S}$. Since $D(i) \leq D(j)$ for all $j \notin \mathcal{S}$ and $D(i) \leq \max(d_{i,j}, D(i))$, we have $D(i) \leq \min(D(j), \max(d_{i,j}, D(i)))$ for all $j \notin \mathcal{S}$. Therefore after Step 2b is performed, which sets $D(j) = \min(D(j), \max(d_{i,j}, D(i)))$ for all $j \notin \mathcal{S}$, we have $D(i) \leq D(j)$ for all $j \notin \mathcal{S}$. Consequently, after Step 2b, $D(l) \leq D(j)$ for all $l \in \mathcal{S}$ and $j \notin \mathcal{S}$. Whence Claim 1 holds at the end of Step 2b and the induction proof is thus complete.

We show that Claim 2 holds by induction. Claim 2 holds at the initialization step, Step 1, of the algorithm because $\mathcal{S} = \{1\}$ and the minimum weight path from node 1 to node j using nodes in \mathcal{S} is $d_{1,j}$, which is the initial value for $D(j)$. Now suppose Claim 2 holds at the beginning of Step 2a of some iteration, i.e., $D(k)$ is, for each node k , the minimum weight from 1 to k using paths with all nodes, except possibly k , belonging to the set \mathcal{S} , we show that at the end of Step 2b, i.e., at the beginning of Step 2a of next iteration, Claim 2 still holds. Let i be the node such that $D(i) = \min_{j \notin \mathcal{S}} D(j)$ in Step 2a, then after Step 2a is performed, we have $\mathcal{S} := \mathcal{S} \cup \{i\}$ and for this newly updated \mathcal{S} , $D(i)$ is the minimum weight from 1 to i using paths with all nodes belonging to the set \mathcal{S} . Recall that in the proof for Claim 1, we have shown that after Step 2a is performed, we have $D(l) \leq D(i)$ for all $l \in \mathcal{S}$ and $D(i) \leq D(j)$ for all $j \notin \mathcal{S}$. Since $D(l) \leq D(i) \leq \max(d_{i,l}, D(i))$, it

follows that for the newly updated \mathcal{S} , $D(l)$ is still, for each node $l \in \mathcal{S}$, the minimum weight from 1 to l using paths with all nodes belonging to the set \mathcal{S} . After Step 2b is performed, we see that \mathcal{S} does not change and $D(l)$ remains the same for all $l \in \mathcal{S}$, so $D(l)$ is, for each node $l \in \mathcal{S}$, the minimum weight from 1 to l using paths with all nodes belonging to the set \mathcal{S} at the end of Step 2b. We now need to consider all $j \notin \mathcal{S}$ in Step 2b. Consider a path from 1 to j which has minimum weight among those paths with all nodes, except j , belonging to the set \mathcal{S} in Step 2b and let $D'(j)$ be the corresponding minimum weight. Such a path must consist of $d_{k,j}$ for some $k \in \mathcal{S}$, preceded by a minimum weight path from 1 to k with nodes in \mathcal{S} .

$$D'(j) = \min_{k \in \mathcal{S}} \{\max(d_{k,j}, D(k))\} = \min \left(\min_{k \in \mathcal{S} \setminus \{i\}} \{\max(d_{k,j}, D(k))\}, \max(d_{i,j}, D(i)) \right)$$

The induction hypothesis implies that $D(j)$ is, for each node $j \notin \mathcal{S} \setminus \{i\}$, the minimum weight from 1 to j using paths with all nodes, except j , belonging to the set $\mathcal{S} \setminus \{i\}$ (because we have updated \mathcal{S} in Step 2a), i.e., $D(j) = \min_{k \in \mathcal{S} \setminus \{i\}} \{\max(d_{k,j}, D(k))\}$, so we obtain $D'(j) = \min(D(j), \max(d_{i,j}, D(i)))$. Thus after Step 2b, which sets $D(j) = \min(D(j), \max(d_{i,j}, D(i)))$, we get that $D(j) = D'(j)$, which is the minimum weight from 1 to j using paths with all nodes, except j , belonging to \mathcal{S} . Whence Claim 2 holds at the end of Step 2b and the induction proof is thus complete.

We note that a new node is added to \mathcal{S} with each iteration, so the algorithm terminates after $N - 1$ iterations, with \mathcal{S} containing all nodes. By Claim 2 $D(j)$ is then equal to the minimum weight from 1 to j .

We claim that the power assignment $\eta : \mathcal{N} \rightarrow \mathbb{R}$ after Step 4b of Algorithm 3.6.1 is a solution to Problem 3.4.3. We prove this claim by contradiction. Suppose that there is another power assignment $\eta' : \mathcal{N} \rightarrow \mathbb{R}$ such that node 1 can successfully send packets to $\mathcal{Q} := \{m_1, m_2, \dots, m_t\}$ and such that $\max_{n \in \mathcal{N}} \eta'(n) < \max_{n \in \mathcal{N}} \eta(n)$.

For power assignment η , let i be the node such that $\eta(i) = \max_{n \in \mathcal{N}} \eta(n)$. Then there must exist a node j such that $d_{i,j} = \eta(i)$ and $d_{i,j}$ is on the minimum weight path in the complete graph \mathcal{C} from node 1 to one of the destination nodes in \mathcal{Q} according to Algorithm 3.6.1. Let $m \in \mathcal{Q}$ be such destination node and recall that $D(m) = \eta(i)$ is the weight of that path. For the power assignment η' , there is at least a sequence of nodes (a path) such that node 1 can successfully send packets to node m via these intermediate nodes. We should caution that there may be several paths formed from node 1 to node m for power assignment η' . For each such path, consider the sequence of nodes on the path. Let $\mathcal{B} = \{b_1, b_2, \dots, b_k\}$ be such sequence of nodes that defines this path, with $b_1 = 1$ and $b_k = m$. The power assignment η' guarantees that $\eta'(b_1)$ is enough for b_1 to successfully send packets to b_2 , $\eta'(b_2)$ is enough for b_2 to successfully send packets to b_3 , \dots , $\eta'(b_{k-1})$ is enough for b_{k-1} to successfully send packets to b_k . Recall that d_{b_1, b_2} is the least power required for b_1 to have a successful transmission to b_2 , d_{b_2, b_3} is the least power for b_2 to have a successful transmission to b_3 , \dots , d_{b_{k-1}, b_k} is the least power for b_{k-1} to have a successful transmission to b_k . Therefore the path defined by $\mathcal{B} = \{b_1, b_2, \dots, b_k\}$ with weights $d_{b_1, b_2}, d_{b_2, b_3}, \dots, d_{b_{k-1}, b_k}$ allows node 1 to successfully transmit packets to node m and $d_{b_1, b_2} \leq \eta'(b_1)$, $d_{b_2, b_3} \leq \eta'(b_2)$, \dots , $d_{b_{k-1}, b_k} \leq \eta'(b_{k-1})$. Since $\max_{n \in \mathcal{N}} \eta'(n) < \max_{n \in \mathcal{N}} \eta(n)$, we have $\max_{1 \leq l \leq k-1} (d_{b_l, b_{l+1}}) \leq \max_{n \in \mathcal{N}} \eta'(n) < \max_{n \in \mathcal{N}} \eta(n) = \eta(i) = D(m)$. The weight of the path determined by $\mathcal{B} = \{b_1, b_2, \dots, b_k\}$ in the complete graph \mathcal{C} is $\max_{1 \leq l \leq k-1} (d_{b_l, b_{l+1}})$. Therefore we have found another path from 1 to m in the complete graph \mathcal{C} that has a weight strictly less than $D(m)$, which is a contradiction to what we have proved that $D(m)$ is the minimum weight of the path from 1 to m in the complete graph \mathcal{C} . Thus there is no other power assignment $\eta' : \mathcal{N} \rightarrow \mathbb{R}$ such that $\max_{n \in \mathcal{N}} \eta'(n) < \max_{n \in \mathcal{N}} \eta(n)$. This completes the proof that Algorithm 3.6.1

indeed gives a solution to Problem 3.4.3.

We now consider the efficiency of Algorithm 3.6.1. Step 1 takes $O(N^2)$ number of operations. To estimate the computation complexity required by Step 2, we note that there are $N-1$ iterations and the number of operations per iteration is proportional to N . Therefore, in the worst case the computation is $O(N^2)$. Step 3 requires less than $O(tN)$ operations. Step 4 takes approximately $O(N)$ operations. We have therefore shown Algorithm 3.6.1 has complexity $O(N^2)$ and thus is an efficient algorithm by our definition. ■

In actuality, the procedure for a calculating node to find out its power for a (source, destination set) pair is as follows. The routing information of each node is flooded through the entire network and each node eventually gets a full view of the topology of the network. Based on the topological information, each node runs Algorithm 3.6.1 for this (source, destination set) pair to get a source-rooted tree from the source to the destination set. If the calculating node is on the tree, it calculates its own power for this (source, destination set) pair and store the tuple (source, destination set, first-hop nodes from the calculating node, power to use) in its routing table. So in Step 4b of the algorithm, a calculating node does not have to calculate and store other nodes' power assignment, even though the calculating node knows the power assignment of other nodes by Algorithm 3.6.1. If network is static or changes relatively slowly, we will expect that each node has the same view of the topology of the network and generates the same source-rooted tree for each (source, destination set) pair, even though the route computation is done distributively.

The (source, destination set, first-hop nodes) information is stored in the header of a packet. When a node receives a packet, it first checks whether itself is one of the

first-hop nodes for the received packet. If it is not, it simply discards the received packet. If it is, it then checks if itself is in the destination set. If it is, it keeps a copy of the received packet. It also checks if it has an entry in its routing table that matches (source, destination set) pair. If it has, it sends out the received packet using the power in the matching entry in its routing table. Otherwise, it simply discards the received packet.

3.6.2 Algorithm for Problem 3.4.4 through 3.4.6

We now solve the problems that minimize total power of the nodes in an ad hoc mobile wireless network, namely, Problem 3.4.4 through 3.4.6. We should notice that Problem 3.4.4 through 3.4.6 are very different from and much harder than the ones that minimize the maximum power, namely, Problem 3.4.1 through 3.4.3. To get an intuition why this is true, let us recall what Algorithm 3.6.1 gives us. After the step of power assignment in Algorithm 3.6.1 is finished for Problem 3.4.3, we know that one of the nodes has maximum power among all the nodes in the network. If we allow any other node to transmit packets at power between its assigned value by Algorithm 3.6.1 and this maximum value, we will have a different power assignment such that it is also a solution to Problem 3.4.3. Therefore in some sense, we don't really need to care about exactly how much power other nodes use except for the node that transmits packets at the maximum power. However, in the problems that minimize the total power, how much power each node uses directly affects the value of the total power. Therefore we may have to keep the exact power used by each node as the state information when designing a power assignment algorithm for an ad hoc mobile wireless network. Fortunately, we can still find an efficient algorithm for the power assignment when we have only one source and one destination. However, when

we have more than one destination nodes, the algorithm for the power assignment turns out to be NP-hard for 2-dimensional space or 3-dimensional space in general.

Algorithm 3.6.2 (Algorithm for Problem 3.4.4) *The algorithm for finding a solution to Problem 3.4.4 is almost the same as Algorithm 3.6.1. We simply replace $\max(d_{i,j}, D(i))$ in Algorithm 3.6.1 by $(d_{i,j} + D(i))$ to get the algorithm description for solving Problem 3.4.4. In this way, we actually define the weight of a path to be the sum of the weights of each edge on the path in the complete graph \mathcal{C} .*

Theorem 3.6.2 (Correctness of Algorithm 3.6.2) *Algorithm 3.6.2 is an efficient algorithm that gives a solution to Problem 3.4.4 with complexity $O(N^2)$.*

Proof of Theorem 3.6.2: The proof for the correctness and efficiency of Algorithm 3.6.2 is completely analogous to the proof for the correctness and efficiency of Algorithm 3.6.1, we simply replace $\max(d_{i,j}, D(i))$ by $(d_{i,j} + D(i))$ in the proof for Theorem 3.6.1 to get the proof for Theorem 3.6.2. Intuitively, if we find a minimum weight path from a source to a destination, where the weight of the path is defined to be the cumulative weight of each edge on the path, the total power used by nodes in the network to forward packets from the source to the destination will be minimized. ■

When a source node needs to send its packets to more than one destination nodes, the difficulty of solving such a problem grows drastically. In fact, we no longer have the same nice results as that for Problem 3.4.4. In a general situation, the algorithms for finding solutions to Problem 3.4.5 or 3.4.6 turn out to be NP-hard in a 2-dimensional space or a 3-dimensional space.

Theorem 3.6.3 (Problem 3.4.5 is NP-hard) *Problem 3.4.5 in a 2-dimensional space or a 3-dimensional space is NP-hard.*

Proof of Theorem 3.6.3: We note that if we restrict Problem 3.4.5 to a subproblem in which we require bidirectionality of communication between each pair of nodes in \mathcal{N} , Problem 3.4.5 in 2-dimensional space or 3-dimensional space becomes the problem in [19, 45], which has been proved to be NP-hard. Therefore Problem 3.4.5 in 2-dimensional space or 3-dimensional space is NP-hard by the restriction technique in [30].

■

Theorem 3.6.4 (Problem 3.4.6 is NP-hard) *Problem 3.4.6 in a 2-dimensional space or a 3-dimensional space is NP-hard.*

Proof of Theorem 3.6.4: We note that if we restrict $\mathcal{Q} = \mathcal{N} \setminus \{n_0\}$, Problem 3.4.6 becomes Problem 3.4.5, which in 2-dimensional space or 3-dimensional space has been proved to be NP-hard in Theorem 3.6.3. Therefore Problem 3.4.6 in 2-dimensional space or 3-dimensional space is NP-hard by the restriction technique in [30].

■

3.7 Conclusion and Future Research

3.7.1 Conclusion

We have shown that routing problems in ad hoc mobile wireless networks with *point-to-multipoint* connections are very different from routing problems in wired networks with *point-to-point* connections. We have found efficient algorithms to assign power to each node in order to minimize the *maximum power* used by nodes in an ad hoc mobile wireless network for unicast, broadcast, and multicast applications. The efficient algorithm that assigns power to each node in order to minimize the *total power* used by nodes in an ad hoc mobile wireless network exists only for unicast application. For broadcast and multicast applications, we have shown that

the problems to assign power to each node in order to minimize the *total power* used by nodes in an ad hoc mobile wireless network in a 2-dimensional space or a 3-dimensional space are NP-hard.

3.7.2 Future Research

For future research, we may need to seek efficient algorithms for finding *approximate solutions* to Problem 3.4.5 and 3.4.6. We may also need to study routing algorithms when we change the receiver model to include probability of reception and receiver processing energy.

More generally, we notice that each node repeats its received packet when it is involved in routing, so there are correlated packets in transit in a network when a packet is routed from a source to a set of destination nodes. Therefore there may be a possibility of diffusing the received packets at each node (dividing packets into pieces) and only sending out several selected pieces in order to efficiently use the transmission rate. If nodes in a network collaborate well, the receivers may recover the original packet, or part of the original packet of their own interest, from the received pieces. In this way, more applications can be accommodated simultaneously. There are techniques for the rate-efficient transmission of packets in *wired networks*, such as network coding, multiple description coding, and distributed source coding. We need to further investigate the rate-efficient transmission in *ad hoc mobile wireless networks*.

CHAPTER IV

OPTIMAL TRANSPORT EFFICIENCY OF LINEAR NETWORKS

In Chapter II we saw that there are many parameters affecting the performance of communication networks. In this chapter we want to study the role of some of the chosen parameters analytically and numerically. In particular, we want to study the effect of different transmitter and receiver systems on communication networks and include the amplifier characteristics and receiver processing energy from the physical layer in our analysis.

4.1 Introduction

For wireless systems that must operate on batteries, total energy consumption by nodes in a network is a critical design parameter. The energy consumed by a node includes energy consumed by the power amplifier, processing energy at the receiver, etc.. Since a lot of power consumed by a communication system resides in the power amplifier, it is important to consider the amplifier characteristics when we evaluate the performance of a network. The processing energy at the receiver is another important factor affecting the performance of a communication network. With the proliferation of more and more handheld devices that operate on batteries and

more and more sophisticated coding and modulation methods, the processing energy becomes a large portion of consumed energy. Apart from the energy concerns in the design of a wireless network, bandwidth is also a critical design parameter. Since in wireless networks, nodes often share common communication media, such as the same frequency band, how efficiently bandwidth is utilized affects the performance of a communication network.

In this chapter, we want to measure the performance of a network by taking into account of both the bandwidth efficiency and the energy efficiency and explicitly considering the power amplifier characteristics and receiver processing energy. From the bandwidth efficiency point of view, higher information bits per channel use is preferred since it will take less number of channel uses to transmit the same amount of information bits. However, with high rate code, the packet error probability is usually higher than that of a low rate code. Consequently, the effective successful transmission of information bits may be compromised. From the energy point of view, if there are too few relay nodes, each transmission has to travel a large distance, which causes large transmitting energy to be used. On the other hand, if there are too many relay nodes, too much energy will be turned into heat and too much processing energy will be used by the relay nodes. Therefore there are many parameters, such as code rate, input power to the power amplifier, distance between adjacent nodes, etc. that should be considered in designing efficient wireless networks. In summary, both bandwidth efficiency and energy efficiency are important measures of a communication network, which we should consider in deploying an actual network.

The energy and bandwidth tradeoff for the point-to-point single-hop communication has been studied to a large extent. The development of turbo codes and

LDPC [7, 50] is the culmination of the error-control coding for the point-to-point single-hop communication. The received energy is usually considered to study the energy and bandwidth tradeoff for the point-to-point single-hop communication, while the amplifier characteristics and the receiver processing energy have not been included in such tradeoff. However, in any practical system design, both the amplifier characteristics and the receiver processing energy are important.

In the literature, much of the study on network performance has ignored many factors from other design layers for tractability. Many simplified models for lower design layers have been abstracted for the analysis at the network layer. A most common receiver model seen in the literature is the threshold model, where received symbol is considered to be correct if the received signal-to-noise (or signal-to-interference plus noise) ratio is above certain threshold and the received symbol is considered to be wrong otherwise. While this is quite a rough model from the physical layer point of view, it usually makes the network analysis tractable, so many network researchers tend to adopt this model for their analysis.

In [18, 19, 45, 67] the authors studied low-energy wireless network from the view point of energy efficient routing paths and the complexity of finding such routing paths, where they all assumed a threshold model for the calculation of routing metrics. In [69] the authors studied a position-based network and designed a local optimization scheme that ensures strong connectivity of the entire network. For the stationary network, the authors showed that their scheme achieves global minimum energy. Even though the authors considered energy, they did not consider bandwidth efficiency explicitly. In [31, 32] the authors studied the performance of a network by a measure they defined as transport capacity. Roughly, it is defined as the product of the average bit rate and average communication distance between a source and a

destination. The transport capacity has a unit bit·m/s. The authors described two communication models, namely, the protocol model and the physical model, used in the network performance analysis. The transport capacity of a network, where n nodes are arbitrarily located in a disk of unit area, was given under these two models. However, the authors did not explicitly consider the bandwidth efficiency and energy efficiency. In [46] the authors studied the network throughput, defined as the expected number of successful source-destination sessions that a packet radio network can sustain and found the optimum transmission radii in order to maximize the network throughput. The analysis given in [46] is based on the parameters at the network layer only, no physical layer characteristics, such as power amplifier or receiver processing energy, was considered in the analysis.

We are not aware that the power amplifier characteristics and receiver processing energy are considered explicitly, or the bandwidth efficiency and energy efficiency are considered together, in any of the previous research. We will analyze the performance of a network by taking into account these factors. The remainder of this chapter is organized as follows. In Section 4.2 we describe a linear network where nodes are aligned on a straight line. Topologically this network can model a path in a dense wireless network. In Section 4.3 we introduce a performance measure, called transport efficiency, which takes into account of both bandwidth efficiency and energy efficiency in wireless networks. We consider amplifier characteristics and receiver processing energy explicitly when we compute transport efficiency. We show that the transport efficiency, when optimized over the input power to the power amplifier and the adjacent distance, is inversely proportional to the distance between a source and a destination for the threshold model. We observe that the same conclusion is true for the cutoff-rate model, uncoded model and convolutional-coded model. We

show numerical results in Section 4.4 and conclude the discussion in Section 4.5.

4.2 Linear Network Topology

For tractability, we will use a linear network topology shown in Figure 4.1 to investigate the effect of different communication models and the effect of amplifier nonlinearity, processing energy, etc. on the performance of the network. This network

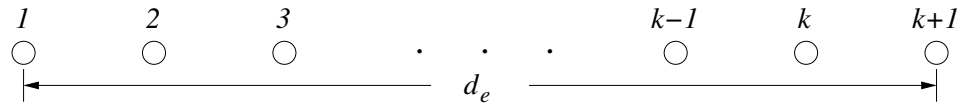


Figure 4.1: Linear wireless network, where node 1 wants to transmit a packet to node $k + 1$.

model, though simple, does represent quite many practical scenarios. For example, in a two or three dimensional network, the best route for the unicast traffic between a source and a destination is topologically a line, even though it may not be straight. We make the following assumptions.

- There are $k + 1$ nodes on a straight line, where node 1 wants to send unicast packets to node $k + 1$. Let d_e be the end-to-end distance between node 1 (source) and node $k + 1$ (destination).
- Each nodes uses a power amplifier whose characteristics is determined by two functions f_c and f_o . Let P_{in} be the input power to the power amplifier, P_{dc} be the consumed power, P_{out} be the output power of the power amplifier. Then the characteristics of the power amplifier is determined by:

$$P_{dc} = f_c(P_{in}) \quad (4.1)$$

and

$$P_{out} = f_o(P_{in}), \quad (4.2)$$

where $f_o(0) = 0$.

Normally f_c and f_o are strictly increasing functions of P_{in} . The difference $P_{dc} - P_{out}$ is the power converted into heat, which is also a function of P_{in} .

Define

$$P_h = P_{dc} - P_{out} = f_c(P_{in}) - f_o(P_{in}). \quad (4.3)$$

The set of allowed input power is

$$\mathcal{P} = \{P : 0 \leq P \leq P_{max}\}, \quad (4.4)$$

where P_{max} is the maximum allowed input power that the power amplifier can sustain.

For example, we may have a simplified amplifier, whose characteristics is shown in Figure 4.2. Usually P_h varies with respect to the input power, while in the simplified amplifier, this difference is a constant. For this simplified amplifier model, we have

$$f_o(P_{in}) = \begin{cases} s_1 P_{in}, & \text{if } 0 \leq P_{in} \leq 1.5 \text{ mW}; \\ c_1, & \text{if } 1.5 \text{ mW} < P_{in} \leq P_{max}; \end{cases} \quad (4.5)$$

where $s_1 = 60$ and $c_1 = 75$ mW.

$$f_c(P_{in}) = f_o(P_{in}) + 35 \text{ mW}. \quad (4.6)$$

Since there is no reason to let input power be larger than 1.5 mW, we will assume that $P_{max} = 1.5$ mW.

- Let the distance between node i and node $i + 1$ be $\theta_i d_e$, where $\theta_i > 0$ for $i = 1, \dots, k$ and $\sum_{i=1}^k \theta_i = 1$. In particular, when nodes are evenly distributed on a straight line, $\theta_i = \frac{1}{k}$ and the distance between adjacent nodes is

$$d_a = \frac{d_e}{k}. \quad (4.7)$$

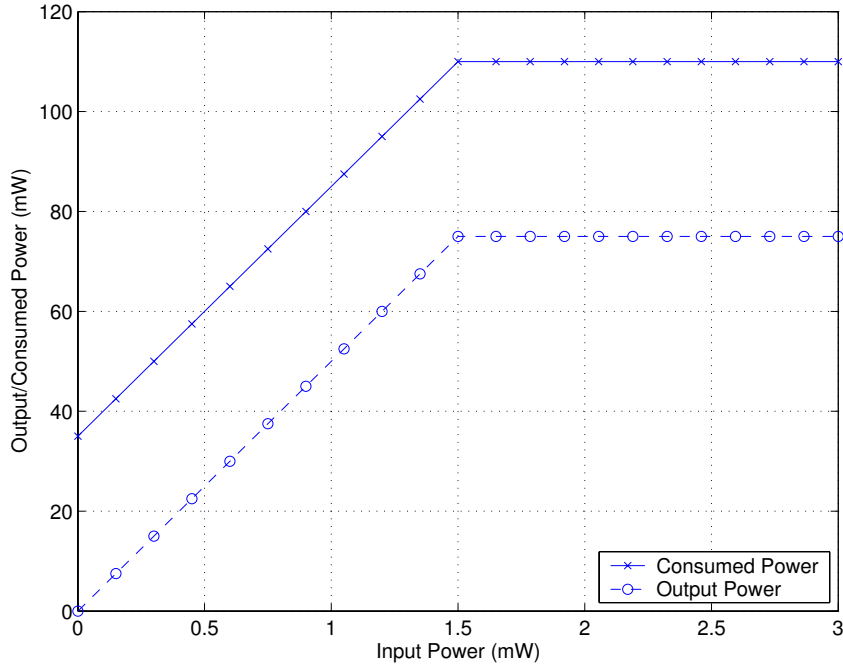


Figure 4.2: Simplified power amplifier model for the class AB amplifier in Figure 2.3. A constant power $P_h = 35$ mW is turned into heat when the amplifier is in operation.

In this case, we require that all nodes use the same input power.

- We assume that the average received signal power and transmitted signal power are related by a function which depends on the distance d between the transmitter and the receiver, namely,

$$P_r = \beta(d)P_t, \quad (4.8)$$

where $\beta(d)$ is the large-scale propagation loss depending on d . We assume that β is a continuous and strictly decreasing function of d . Without loss of generality, we require that $d \geq 1$ m in order for the receiver to be in the far zone of the transmitter.

For example, a commonly used propagation model is

$$\beta(d) = \frac{b}{d^q}. \quad (4.9)$$

where q is usually between 2 and 4 and b is a constant. For the propagation model given in (2.30), where the received power and output power of the power amplifier is related by

$$P_r = 4P_t \left(\frac{\lambda_c}{4\pi d} \right)^2 G_t G_r \eta_t \eta_r \sin^2 \left(\frac{2\pi h_t h_r}{\lambda_c d} \right) \approx P_t \eta_t \eta_r G_t G_r \frac{h_t^2 h_r^2}{d^4}, \quad (4.10)$$

we have $q = 4$ and

$$b = \eta_t \eta_r G_t G_r h_t^2 h_r^2. \quad (4.11)$$

- Each symbol is modulated using BPSK. Each transmitted symbol has a duration of T_s . Node 1 sends packets of length N to node $k + 1$. Therefore $Nf_c(P_{in})T_s$ is the consumed energy for a node to send a packet.
- There is no interference between the transmission of different users. This may be achieved via multiple access control protocols, such as TDMA, etc. Thus all transmissions are independent of each other.
- The packet success probability depends on received symbol signal-to-noise ratio

$$\gamma = \frac{E_s}{N_0}, \quad (4.12)$$

where

$$E_s = P_r T_s \quad (4.13)$$

is the received energy for a channel symbol.

- Each node, when receiving a packet, consumes E_p amount of energy to process a received channel symbol. This amount of energy is usually consumed by the demodulator and decoder as shown in Chapter II. In practice, different E_p for the same received signal-to-noise ratio may result in different packet error rate. It is also a fact that different coding and modulation techniques have different

E_p for the same packet error rate. In our analysis, we assume that E_p is a constant.

- Let each transmitter employ a coding technique with K information bits and N coded bits. The code rate $R = K/N$. We will assume that each packet has the same K , but the transmitter is allowed to choose R .
- Let $P_{1,s}(R, \gamma)$ be the probability of success for a packet when the symbol signal-to-noise ratio at the receiver is γ . The packet success probability for k -hops is given by $P_{k,s}(R, \gamma) = [P_{1,s}(R, \gamma)]^k$.
- We assume that a packet is considered lost if any one of the nodes except node 1 does not receive the packet correctly. So there is no automatic repeat request (ARQ) mechanism in the network.

4.3 Performance Analysis

We first assume that $k+1$ nodes are evenly placed on a straight line with adjacent nodes separated by d_a . Assume that node 1 has packets that are intended to be delivered to node $k+1$. Let $\mathcal{K}(d_e)$ be the set of allowed number of hops for a given end-to-end distance d_e in order for adjacent nodes to stay in the far zone of each other, i.e.,

$$\mathcal{K}(d_e) = \left\{ k \in \mathbb{N} : \frac{d_e}{k} > 1 \text{ m} \right\}. \quad (4.14)$$

Based on the propagation model in (4.8) and the amplifier model in (4.2), we have that the received power at a node is related to the input power at its upstream node by the following relation,

$$P_r = \beta(d_a) f_o(P_{in}). \quad (4.15)$$

We assume that all end-to-end transmissions are independent and identically distributed, which allows us to define a performance measure for a linear network from the performance of the transmission of one packet. A node needs energy $Nf_c(P_{in})T_s$ to send a packet to its downstream neighbor. A packet is received by a node at the expense of processing energy NE_p . Node 1 only consumes energy $Nf_c(P_{in})T_s$ without using any processing energy, while node $k + 1$ only consumes energy NE_p for processing a packet without using any transmitting energy. For a single packet that is sent from node 1 to node $k + 1$, the total amount of energy for transmitting and processing this packet is $kN[f_c(P_{in})T_s + E_p]$. We assume that if a packet is not successfully received, every information bit in the packet is lost.

The **energy efficiency** is defined by

$$\begin{aligned} E_{eff} &= \frac{KP_{k,s}\left(R, \frac{E_s}{N_0}\right)}{kN[f_c(P_{in})T_s + E_p]} \\ &= \frac{R}{k[f_c(P_{in})T_s + E_p]} \left[P_{1,s}\left(R, \frac{\beta(d_a)f_o(P_{in})T_s}{N_0}\right) \right]^k. \end{aligned} \quad (4.16)$$

and the **bandwidth efficiency** is defined by

$$B_{eff} = \frac{K}{N} = R \quad (\text{bits})/(\text{channel use}). \quad (4.17)$$

In general, it is possible to achieve higher energy efficiency by trading off bandwidth efficiency. However, bandwidth efficiency is not of negligible interest in wireless communications. Since all communication nodes share a common media, high bandwidth efficiency yields high throughput for nodes in a wireless network. One traditional approach is to optimize energy efficiency with the constraint that $R \geq R_0$. Define

$$B(R) = \begin{cases} 1, & R \geq R_0; \\ 0, & \text{otherwise.} \end{cases}$$

Then the constraint problem is equivalent to finding the largest $B(R) \times E_{eff}$.

By introducing a bandwidth-efficiency function $B(R)$, we can convert a constrained optimum energy efficiency problem into an unconstrained problem that considers not only energy efficiency but also bandwidth efficiency. In other words, we attempt to optimize $B(R) \times E_{eff}$ rather than E_{eff} . In general $B(R)$ is a monotonic increasing function of R . When $B(R) = R$, we define **transport efficiency** $\mu : \mathcal{K} \times \mathcal{P} \times [0, 1] \times (1, \infty) \rightarrow \mathbb{R}$ by

$$\begin{aligned} \mu(k, P_{in}, R, d_e) &= R \times E_{eff} \\ &= \frac{R^2}{k [f_c(P_{in})T_s + E_p]} \left[P_{1,s} \left(R, \frac{\beta(d_a)f_o(P_{in})T_s}{N_0} \right) \right]^k \end{aligned} \quad (4.18)$$

for the number of hops k , input power P_{in} , code rate R , and end-to-end distance d_e . The basic system design problem is to maximize the transport efficiency over the number of hops k , the input power P_{in} , and the code rate R , for a given end-to-end distance d_e .

We now evaluate the performance of following different transmitter and receiver models:

1. threshold model (capacity achieving transmitter and receiver);
2. cutoff-rate model;
3. uncoded model;
4. rate $\frac{1}{2}$ convolutional-coded model.

The packet error rate for these models versus the received signal-to-noise ratio is given in Figure 4.3. We see that there is a saving of 5.5 dB in received signal-to-noise ratio for the rate $\frac{1}{2}$ convolutional-coded model over the uncoded model when the packet has 224 information bits and packet error rate is 10^{-2} and there is a saving

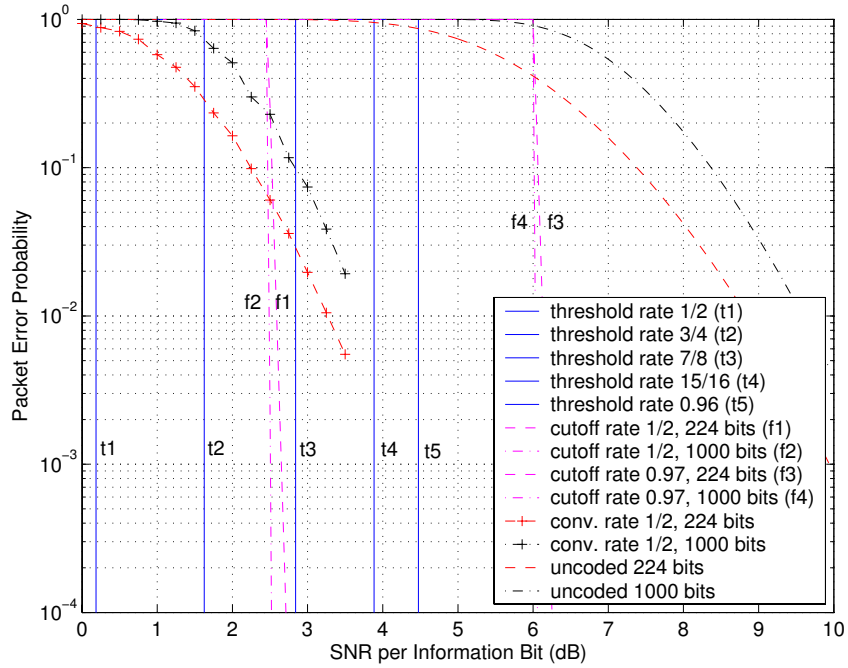


Figure 4.3: Packet error probability versus signal-to-noise ratio per information bit for the four different models, with each packet containing 224 bits or 1000 bits. The minimum signal-to-noise ratio for the channel capacity to be $\frac{1}{2}$, $\frac{3}{4}$, $\frac{7}{8}$, $\frac{15}{16}$, and 0.96 bit / (channel use) for the binary input antipodal signaling (BPSK) AWGN channel is shown. The packet error rate for the cutoff-rate model is shown for cutoff-rate $\frac{1}{2}$ and cutoff-rate 0.97.

of 8.5 dB in received signal-to-noise ratio for the threshold model when the channel capacity is $\frac{1}{2}$ over the uncoded model when the packet has 224 information bits and packet error rate is 10^{-2} .

4.3.1 Threshold Model

In the threshold model, the communication system employs capacity achieving transmitter and receiver. Because the output power of the power amplifier is bounded, i.e., there is a minimum and maximum constraint on the output power, we can not use the capacity for AWGN channel where there is only an average energy constraint on the transmitted channel symbols. Under the assumption of binary input antipodal signaling (BPSK) AWGN channel, we have that the capacity of the channel is given by [65, 77]

$$C(\gamma) = 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\sqrt{2\gamma})^2}{2}} \log_2 \left(1 + e^{-2x\sqrt{2\gamma}} \right) dx, \quad (4.19)$$

where γ is the received signal-to-noise ratio. Note that we need to have $R \leq C(\gamma)$ in order for the threshold receiver to be useful in practice. Define γ_t so that $C(\gamma_t) = R$, i.e.,

$$\gamma_t = C^{-1}(R). \quad (4.20)$$

For the threshold model, we have

$$P_{1,s} \left(R, \frac{E_s}{N_0} \right) = \begin{cases} 1, & \text{if } \frac{E_s}{N_0} \geq \gamma_t; \\ 0, & \text{otherwise.} \end{cases} \quad (4.21)$$

Therefore in this model, the system employs capacity achieving receiver. Let P_{in}^* satisfy

$$\frac{\beta(d_a) f_o(P_{in}^*) T_s}{N_0} = \gamma_t. \quad (4.22)$$

Note that choosing $P_{in} > P_{in}^*$ does not give any further benefit in $P_{1,s}$. The transport efficiency with $P_{in} = P_{in}^*$ is

$$\mu(k, P_{in}^*, C(\gamma_t), d_e) = \frac{[C(\gamma_t)]^2}{k \cdot [f_c(P_{in}^*)T_s + E_p]}. \quad (4.23)$$

We require that $P_{in}^* \in \mathcal{P}$. Since f_o is a continuous and strictly increasing function of P_{in} , we have that $P_{in}^* \in \mathcal{P}$ is equivalent to

$$0 \leq \frac{\gamma_t N_0}{\beta(d_a)T_s} \leq f_o(P_{max}). \quad (4.24)$$

Therefore

$$0 \leq d_a \leq \beta^{-1} \left(\frac{\gamma_t N_0}{f_o(P_{max})T_s} \right). \quad (4.25)$$

Normally we require that d_a is large enough so that the receiver is in the far zone of the transmitter. As indicated earlier, without loss of generality, we require that $d_a \geq 1$ m. For each $\gamma_t \in (0, \infty)$, define

$$\mathcal{D}_{\gamma_t} = \left\{ d \in \mathbb{R} : 1 \leq d \leq \beta^{-1} \left(\frac{\gamma_t N_0}{f_o(P_{max})T_s} \right) \right\}. \quad (4.26)$$

Define an indicator function $\mathbf{1}_{\mathcal{D}_{\gamma_t}} : (1, \infty) \rightarrow \mathbb{R}$ by

$$\mathbf{1}_{\mathcal{D}_{\gamma_t}}(d) = \begin{cases} 1, & \text{if } d \in \mathcal{D}_{\gamma_t}; \\ 0, & \text{otherwise.} \end{cases} \quad (4.27)$$

Proposition 4.3.1 *For a linear network with the threshold model where the source and destination nodes are separated by d_e , we have*

$$\sup_{k \in \mathcal{K}(d_e)} d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) \leq A(\gamma_t), \quad (4.28)$$

where $A(\gamma_t)$ does not depend on d_e . Furthermore,

$$\lim_{d_e \rightarrow \infty} \sup_{k \in \mathcal{K}(d_e)} d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) = A(\gamma_t). \quad (4.29)$$

Proof of Proposition 4.3.1:

We derive from (4.18) that

$$\mu(k, P_{in}^*, C(\gamma_t), d_e) = \left(\frac{[C(\gamma_t)]^2}{k (f_c(P_{in}^*) T_s + E_p)} \right) \mathbf{1}_{\mathcal{D}_{\gamma_t}} \left(\frac{d_e}{k} \right) \quad (4.30)$$

$$= \left(\frac{1}{d_e} \right) \left(\frac{[C(\gamma_t)]^2 d_a}{f_c(P_{in}^*) T_s + E_p} \right) \mathbf{1}_{\mathcal{D}_{\gamma_t}}(d_a). \quad (4.31)$$

Thus

$$d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) = \left(\frac{[C(\gamma_t)]^2 d_a}{f_c(P_{in}^*) T_s + E_p} \right) \mathbf{1}_{\mathcal{D}_{\gamma_t}}(d_a). \quad (4.32)$$

If we allow d_a to take any value in \mathcal{D}_{γ_t} rather than the discrete values d_e/k , $k \in \mathcal{K}(d_e)$ on the right-hand side of (4.32), it follows that the right-hand side of (4.32) is not a function of d_e because P_{in}^* is a function of d_a and γ_t , and \mathcal{D}_{γ_t} is a function of γ_t . Since $d_e \mu = 0$ outside \mathcal{D}_{γ_t} , we may maximize $d_e \mu$ on \mathcal{D}_{γ_t} . Define

$$A(\gamma_t) = \sup_{d_a \in \mathcal{D}_{\gamma_t}} \frac{[C(\gamma_t)]^2 d_a}{f_c(P_{in}^*) T_s + E_p}. \quad (4.33)$$

Then

$$\sup_{k \in \mathcal{K}(d_e)} d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) \leq A(\gamma_t). \quad (4.34)$$

Note $A(\gamma_t)$ does not depend on d_e . Let

$$d_a^* = \arg \sup_{d_a \in \mathcal{D}_{\gamma_t}} \frac{[C(\gamma_t)]^2 d_a}{f_c(P_{in}^*) T_s + E_p}. \quad (4.35)$$

Then d_a^* depends on γ_t , but does not depend on d_e . For such d_a^* , there exists an integer k^* , such that

$$\frac{d_e}{k^* + 1} \leq d_a^* \leq \frac{d_e}{k^*}. \quad (4.36)$$

If $d_e \rightarrow \infty$, we have that $k^* \rightarrow \infty$. Since

$$\left| \frac{d_e}{k^*} - d_a^* \right| \leq \left| \frac{d_e}{k^*} - \frac{d_e}{k^* + 1} \right| \leq \left| \frac{d_e}{k^*(k^* + 1)} \right| \leq d_a^* \frac{1}{k^*} \rightarrow 0 \quad \text{as } k^* \rightarrow \infty \quad (4.37)$$

and $\frac{[C(\gamma_t)]^2 d_a}{f_c(P_{in}^*)T_s + E_p}$ is a continuous function of d_a on \mathcal{D}_{γ_t} , we have that

$$\frac{[C(\gamma_t)]^2 d_e/k^*}{f_c(P_{in}^*)T_s + E_p} \rightarrow \frac{[C(\gamma_t)]^2 d_a^*}{f_c(P_{in}^*)T_s + E_p} \quad \text{as } d_e \rightarrow \infty. \quad (4.38)$$

Thus

$$d_e \mu(k^*, P_{in}^*, C(\gamma_t), d_e) \rightarrow d_e \mu(d_a^*, P_{in}^*, C(\gamma_t), d_e) \quad \text{as } d_e \rightarrow \infty. \quad (4.39)$$

Therefore

$$\lim_{d_e \rightarrow \infty} \sup_{k \in \mathcal{K}(d_e)} d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) = A(\gamma_t). \quad (4.40)$$

■

Proposition 4.3.1 tells us that the transport efficiency, when optimized over the number of hops, is inversely proportional to d_e when d_e is large. As an example, we will calculate the transport efficiency for the specific power amplifier model f_o and f_c given in (4.5) and (4.6), and the propagation model β given in (4.9), and find optimal P_{in}^* , d_a^* , and γ_t^* explicitly. From (4.22), we have

$$P_{in}^* = \frac{\gamma_t N_0 d_a^q}{s_1 b T_s}. \quad (4.41)$$

From (4.26), we have for each $\gamma_t \in (0, \infty)$,

$$\mathcal{D}_{\gamma_t} = \left\{ d \in \mathbb{R} : 1 \leq d \leq \left(\frac{b s_1 P_{max} T_s}{\gamma_t N_0} \right)^{\frac{1}{q}} \right\}. \quad (4.42)$$

From (4.32), we have

$$d_e \mu(k, P_{in}^*, C(\gamma_t), d_e) = \left(\frac{[C(\gamma_t)]^2 b d_a}{\gamma_t N_0 d_a^q + b(P_h T_s + E_p)} \right) \mathbf{1}_{\mathcal{D}_{\gamma_t}}(d_a). \quad (4.43)$$

Define $g : \mathcal{D}_{\gamma_t} \rightarrow R$ by

$$g(d_a) = \left(\frac{[C(\gamma_t)]^2 b d_a}{\gamma_t N_0 d_a^q + b(P_h T_s + E_p)} \right). \quad (4.44)$$

For each $d_a \in \mathcal{D}_{\gamma_t}$, we can take the derivative of g with respect to d_a ,

$$\frac{\partial g}{\partial d_a} = ([C(\gamma_t)]^2 b) \left(\frac{b(P_h T_s + E_p) - (q-1)\gamma_t N_0 d_a^q}{(\gamma_t N_0 d_a^q + b(P_h T_s + E_p))^2} \right), \quad (4.45)$$

and set it to zero to find optimal d_a^* ,

$$d_a^* = \left(\frac{b(P_h T_s + E_p)}{(q-1)\gamma_t N_0} \right)^{\frac{1}{q}}. \quad (4.46)$$

If $d_a^* \in \mathcal{D}_{\gamma_t}$, we have that d_a^* is feasible. Note that d_a^* in (4.46) does not depend on d_e . Substituting d_a^* into (4.43), we have

$$g(d_a^*) = [C(\gamma_t)]^2 \left(\frac{(q-1)^{(q-1)} b T_s}{q^q \gamma_t N_0 (P_h T_s + E_p)^{(q-1)}} \right)^{\frac{1}{q}}. \quad (4.47)$$

It follows from (4.33) that

$$A(\gamma_t) = [C(\gamma_t)]^2 \left(\frac{(q-1)^{(q-1)} b T_s}{q^q \gamma_t N_0 (P_h T_s + E_p)^{(q-1)}} \right)^{\frac{1}{q}}. \quad (4.48)$$

We can now optimize $A(\gamma_t)$ over the threshold γ_t to find the optimal γ_t^* and the optimal code rate $R^* = C(\gamma_t^*)$. Their values are given in Table 4.1 for different q when $T_s = 20 \mu\text{s}$, $N_0 = 3.9991 \times 10^{-21} \text{ W/Hz}$, $E_p = 250 \mu\text{J}$, $P_h = 35 \text{ mW}$, and $b = 4.97 \times 10^{-2} \text{ m}^4$. It is very interesting to note that if $P_h = 0$ and $E_p = 0$,

Table 4.1: Optimal Code Rate.

q	γ_t^* (channel bit)	$(E_b/N_0)^*$ per info. bit (dB)	R^* (bit/(channel use))
2	1.82	3.09	0.89
3	2.33	3.94	0.94
4	2.69	4.48	0.96

we would have $d_a^* = 0$, which means that we should place as many relay nodes as possible. Therefore in practice we need to consider the amplifier characteristics and receiver processing energy explicitly in the design of wireless networks.

Our next proposition shows that for a linear network with threshold model and for a given number of hops k , evenly placing relay nodes between source and destination

nodes yields higher transport efficiency than unevenly placing relay nodes under certain conditions.

Proposition 4.3.2 *If $k-1$ relay nodes are allowed to place on a straight line between source and destination nodes, where all nodes have the same threshold γ_t , and if $f_c\left(f_o^{-1}\left(\frac{a}{\beta(x)}\right)\right)$ is a convex cup function of x for any $a > 0$, evenly placing relay nodes between source and destination nodes yields higher transport efficiency than unevenly placing relay nodes.*

Proof of Proposition 4.3.2:

Suppose we place $k-1$ relay nodes in such a way that each hop has length $\theta_i d_e$, where $i = 1, 2, \dots, k$, $0 \leq \theta_i \leq 1$, $\sum_{i=1}^k \theta_i = 1$. For the threshold model where each node has the same threshold γ_t , the input power $P_{in}^{(i)*}$ at node i satisfies

$$\frac{\beta(\theta_i d_e) f_o\left(P_{in}^{(i)*}\right) T_s}{N_0} = \gamma_t. \quad (4.49)$$

Then

$$P_{in}^{(i)*} = f_o^{-1}\left(\frac{\gamma_t N_0}{\beta(\theta_i d_e) T_s}\right). \quad (4.50)$$

The packet success probability from node 1 to node $k+1$ is

$$P_{k,s} = \prod_{i=1}^k P_{1,s} \left(R, \frac{\beta(\theta_i d_e) f_o\left(P_{in}^{(i)*}\right) T_s}{N_0} \right). \quad (4.51)$$

The total consumed energy for a packet transmitted by node 1 to reach node $k+1$ is $\sum_{i=1}^k N \left(f_c\left(P_{in}^{(i)*}\right) T_s + E_p \right)$. The transport efficiency is

$$\begin{aligned} \mu_{uneven}(\theta_1, \dots, \theta_k, \gamma_t, d_e) &= \frac{[C(\gamma_t)]^2}{\sum_{i=1}^k \left(f_c\left(P_{in}^{(i)*}\right) T_s + E_p \right)} \prod_{i=1}^k \mathbf{1}_{\mathcal{D}_{\gamma_t}}(\theta_i d_e). \\ &= \frac{[C(\gamma_t)]^2}{\sum_{i=1}^k \left(f_c\left(f_o^{-1}\left(\frac{\gamma_t N_0}{\beta(\theta_i d_e) T_s}\right)\right) T_s + E_p \right)} \prod_{i=1}^k \mathbf{1}_{\mathcal{D}_{\gamma_t}}(\theta_i d_e). \end{aligned} \quad (4.52)$$

Define $f_t : (1, \infty) \times (0, \infty) \rightarrow \mathbb{R}$ by

$$f_t(x, y) = f_c \left(f_o^{-1} \left(\frac{yN_0}{\beta(x)T_s} \right) \right). \quad (4.53)$$

Then according to our assumption, we know that f_t is a convex cup function of x .

From (4.52), we have

$$\mu_{uneven}(\theta_1, \dots, \theta_k, \gamma_t, d_e) = \frac{[C(\gamma_t)]^2}{\sum_{i=1}^k (f_t(\theta_i d_e, \gamma_t) T_s + E_p)} \prod_{i=1}^k \mathbf{1}_{\mathcal{D}_{\gamma_t}}(\theta_i d_e). \quad (4.54)$$

Form (4.30), we have for $\theta_i = \frac{1}{k}$

$$\mu(k, P_{in}, \gamma_t, d_e) = \frac{[C(\gamma_t)]^2}{k (f_t(d_e, \gamma_t) T_s + E_p)} \mathbf{1}_{\mathcal{D}_{\gamma_t}}(d_e). \quad (4.55)$$

Comparing (4.54) with (4.55), we notice that we only need to find out the minimum of $\sum_{i=1}^k f_t(\theta_i d_e, \gamma_t)$ in order to see which strategy gives better transport efficiency.

We use Lagrange multiplier method to covert the following constraint optimization problem

$$\inf \left(\sum_{i=1}^k f_t(\theta_i d_e, \gamma_t) \right) \quad \text{subject to} \quad \sum_i \theta_i = 1 \quad (4.56)$$

to the unconditional optimization problem

$$\inf \left(\sum_{i=1}^k f_t(\theta_i d_e, \gamma_t) - \lambda \left(\sum_i \theta_i - 1 \right) \right), \quad (4.57)$$

where $\lambda \in \mathbb{R}$. Since f_t is a convex cup function of x , it follows that the optimal $\theta_i^* = \frac{1}{k}$ for $i = 1, \dots, k$. Therefore for each k , evenly placing relay nodes yields higher transport efficiency than unevenly placing relay nodes. ■

As an example, for the amplifier model given in (4.5) and (4.6), and the propagation model given in (4.9), we have

$$f_t(x, y) = \frac{yN_0}{bT_s} x^q + 35, \quad (4.58)$$

which is a convex cup function of x if $q > 1$. Therefore under these conditions, evenly placing relay nodes yields higher transport efficiency than unevenly placing relay nodes.

4.3.2 Cutoff-Rate Model

The cutoff rate is obtained via the calculation of the codeword error probability of an ensemble of codes. The codeword probability is given by

$$P_e = 2^{-N(R_0-R)}, \quad (4.59)$$

where N is the block length of a codeword, $R = K/N$ is the code rate, and R_0 is the **cutoff rate**. For the binary input antipodal signaling AWGN channel, the cutoff rate is given by

$$R_0 = 1 - \log_2(1 + e^{-E_s/N_0}), \quad (4.60)$$

where E_s is the received energy per channel symbol and is given by

$$E_s = P_r T_s = \beta(d_a) f_o(P_{in}) T_s. \quad (4.61)$$

We say that a communication system achieves cutoff rate if the codeword error rate between a transmitter and a receiver is given by (4.59).

The packet success probability is

$$\begin{aligned} P_{1,s} \left(R, \frac{E_s}{N_0} \right) &= 1 - P_e \\ &= 1 - 2^{-N(R_0-R)} \\ &= 1 - 2^{-K \left(\frac{R_0}{R} - 1 \right)}. \end{aligned} \quad (4.62)$$

Substituting (4.62) into (4.18), we have

$$\mu(k, P_{in}, R, d_e) = \left(\frac{R^2}{k(f_c(P_{in})T_s + E_p)} \right) \left(1 - 2^{-K \left(\frac{R_0}{R} - 1 \right)} \right)^k. \quad (4.63)$$

From (4.60) and (4.61), we know R_0 is a function of d_a and P_{in} . For each fixed k , we can find $\sup_{P_{in} \in \mathcal{P}} \sup_{R \in [0,1]} \mu$. We can also find the optimized transport efficiency $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \sup_{R \in [0,1]} \mu$. The optimization is not easily done analytically. Therefore we resort to numerical methods.

4.3.3 Uncoded Model

If each node does not perform any coding to the information blocks, we have the code rate $R = 1$. The packet success probability is

$$P_{1,s} \left(1, \frac{E_s}{N_0} \right) = \left(1 - Q \left(\sqrt{\frac{2\beta(d_a) f_o(P_{in}) T_s}{N_0}} \right) \right)^N, \quad (4.64)$$

where

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt. \quad (4.65)$$

The packet error probability $1 - P_{1,s} \left(1, \frac{E_s}{N_0} \right)$ for the uncoded model is shown in Figure 4.3. Substituting R and (4.64) into (4.18), we have

$$\mu(k, P_{in}, 1, d_e) = \left(\frac{1}{k(f_c(P_{in})T_s + E_p)} \right) \left[\left(1 - Q \left(\sqrt{\frac{2\beta(d_e/k) f_o(P_{in}) T_s}{N_0}} \right) \right)^N \right]^k. \quad (4.66)$$

For each fixed k , we can find $\sup_{P_{in} \in \mathcal{P}} \mu$. We can also find the optimized transport efficiency $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \mu$. The optimization is not easily done analytically. Therefore we resort to numerical methods.

4.3.4 Convolutional-Coded Model

For the convolutional-coded model, the transmitter uses a standard rate 1/2 constraint length 7 convolutional code and the receiver uses Viterbi decoding. The packet success probability between adjacent nodes is $P_{1,s} \left(R, \frac{E_s}{N_0} \right)$. The packet error probability $1 - P_{1,s} \left(1, \frac{E_s}{N_0} \right)$ for the convolutional-coded model is shown in Figure 4.3.

Due to the finite trellis length, we have two code rates based on the lengths of information sequences for this rate $\frac{1}{2}$ convolutional code. If the length of the information sequence is 224 bits, we have

$$R = \frac{224}{2(224 + 6)} = \frac{224}{460}. \quad (4.67)$$

If the length of the information sequence is 1000 bits, we have

$$R = \frac{1000}{2(1000 + 6)} = \frac{1000}{2012}. \quad (4.68)$$

Substituting R and the packet success probability into (4.18), we have

$$\mu(k, P_{in}, R, d_e) = \left(\frac{R^2}{k(f_c(P_{in})T_s + E_p)} \right) \left[P_{1,s} \left(R, \frac{\beta(d_e/k)f_o(P_{in})T_s}{N_0} \right) \right]^k. \quad (4.69)$$

For each fixed k , we can find $\sup_{P_{in} \in \mathcal{P}} \mu$. We can also find the optimized transport efficiency $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \mu$. The optimization is not easily done analytically. Therefore we resort to numerical methods.

4.4 Numerical Results

In this section, we give numerical results of the transport efficiency for the transmitter and receiver models described before. We assume the following values for our numerical evaluation.

transmission rate	$\frac{1}{T_s}$	=	50×10^3	b/s,
white noise	N_0	=	3.9991×10^{-21}	W/Hz,
packet length	K	=	224 or 1000	bits,
processing energy	E_p	=	0.000250	J,
power into heat	P_h	=	35	mW,
maximum transmit power	P_{max}	=	1.5	mW,

The propagation model is $\beta(d) = \frac{b}{d^4}$, where

$$b = 4.97 \times 10^{-2} \text{ m}^4. \quad (4.70)$$

4.4.1 Transport Efficiency for Packet Size 224 Bits

For the threshold model, we have the optimal $d_a^* = 1171$ m. We see that the optimal number of hops varies linearly with the end-to-end distance d_e when d_e is large. When the end-to-end distance is between 1 m and 1600 m, the optimal number of hops is 1. This means that single-hop transmission is better than multi-hop transmission when the source and destination are close to each other. For each k , we can optimize transport efficiency over P_{in} and R . The optimized transport efficiency for each fixed k and optimal k (where k is allowed to be a non-integer greater than or equal to 1) for different models is shown in Figures 4.4 to 4.7. The optimal number of hops for different models is given in Figures 4.8 to 4.11. The optimal packet error probability is shown in Figures 4.12 to 4.14.

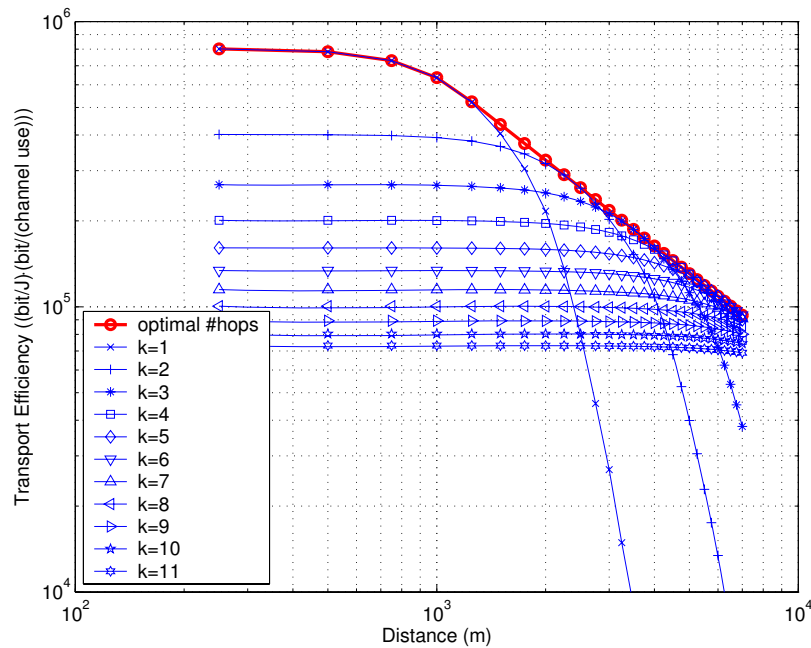


Figure 4.4: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for threshold model.

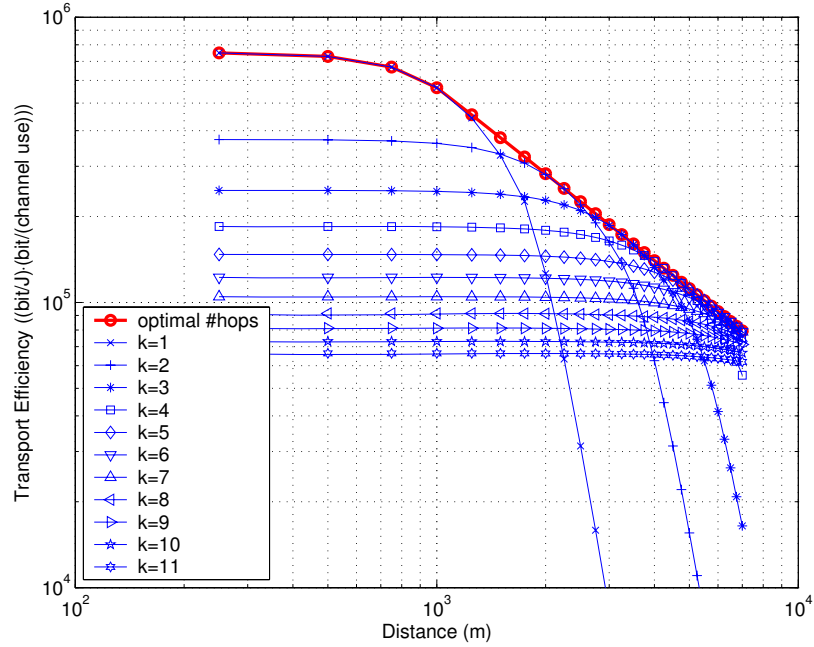


Figure 4.5: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for cutoff-rate model. The length of information sequence is 224 bits.

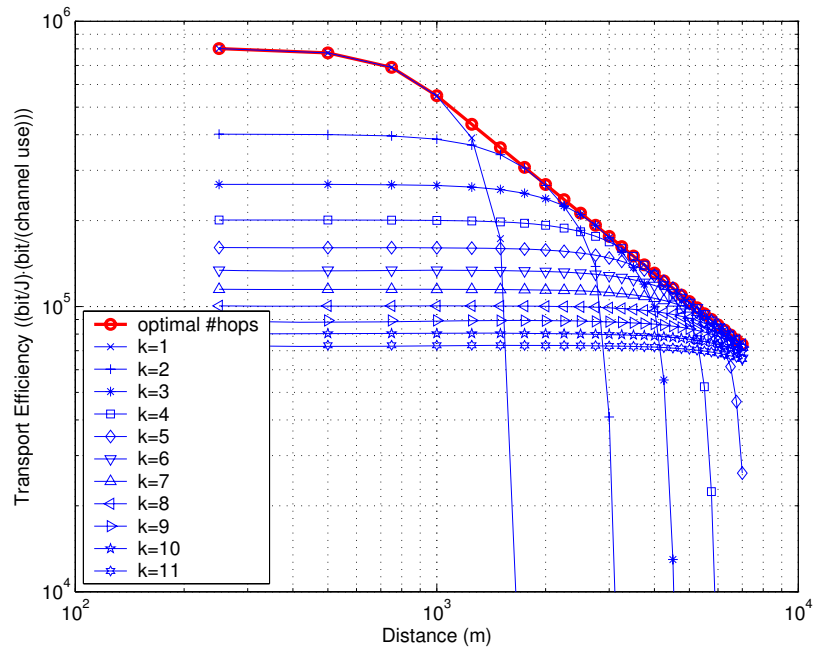


Figure 4.6: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for uncoded model. The length of information sequence is 224 bits.

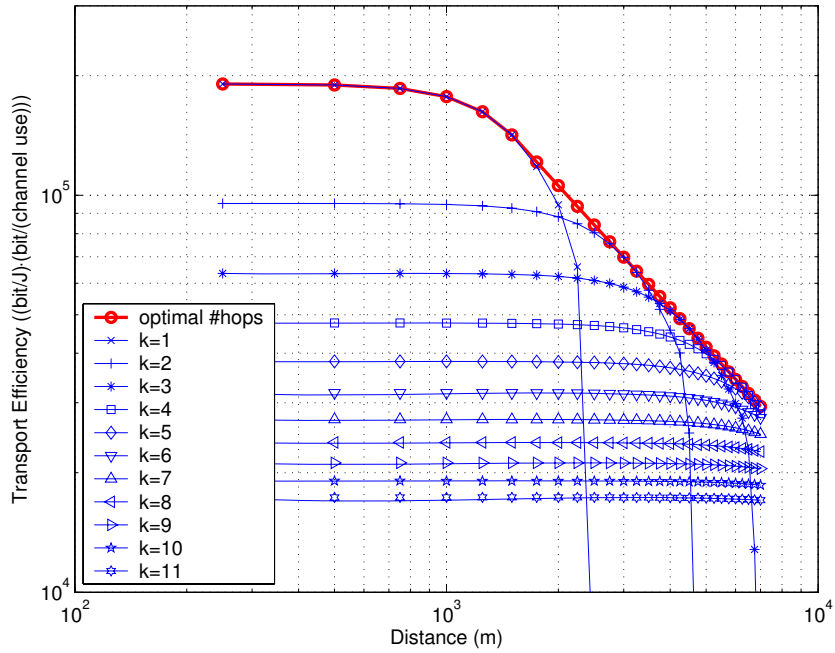


Figure 4.7: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for convolutional-coded model. The length of information sequence is 224 bits.

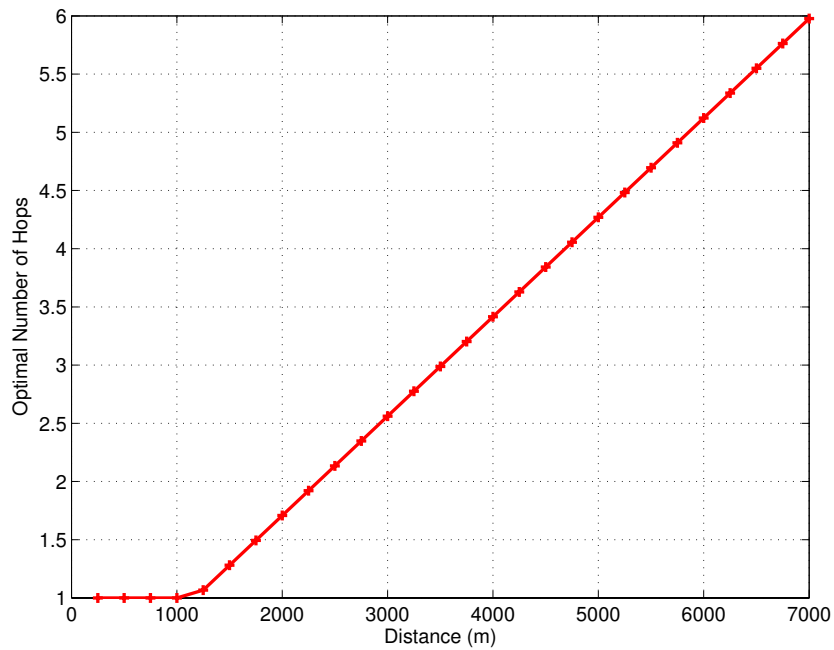


Figure 4.8: Optimal number of hops for threshold model.

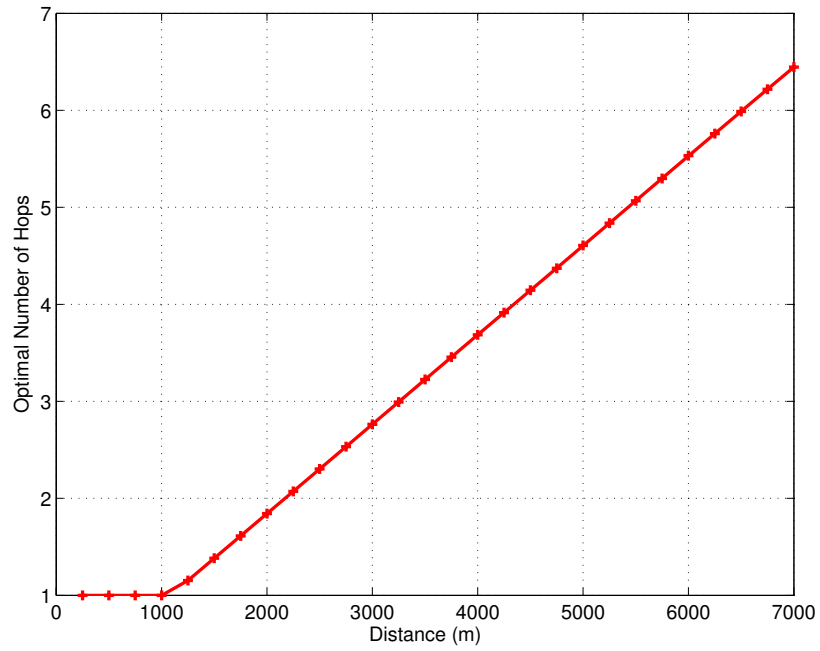


Figure 4.9: Optimal number of hops for cutoff-rate model. The length of information sequence is 224 bits.

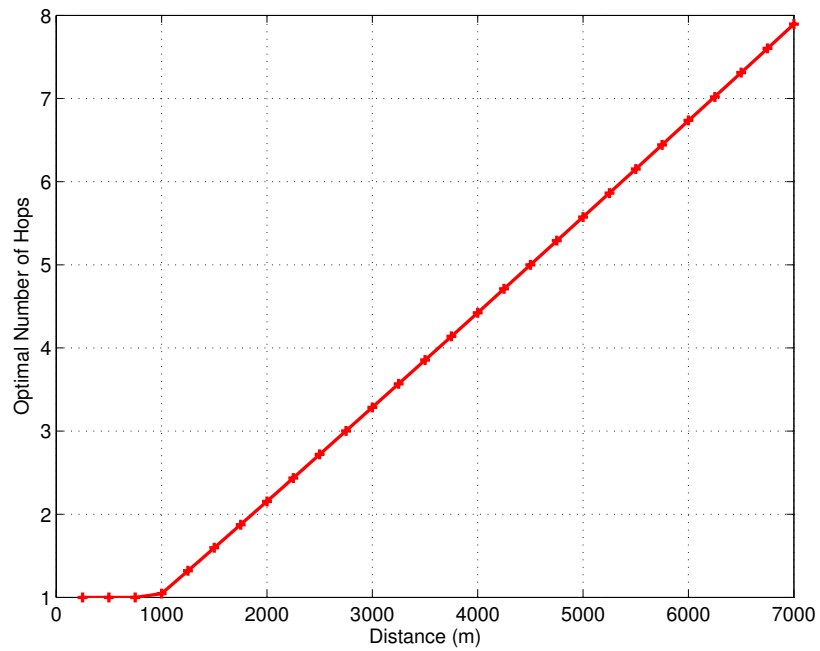


Figure 4.10: Optimal number of hops for uncoded model. The length of information sequence is 224 bits.

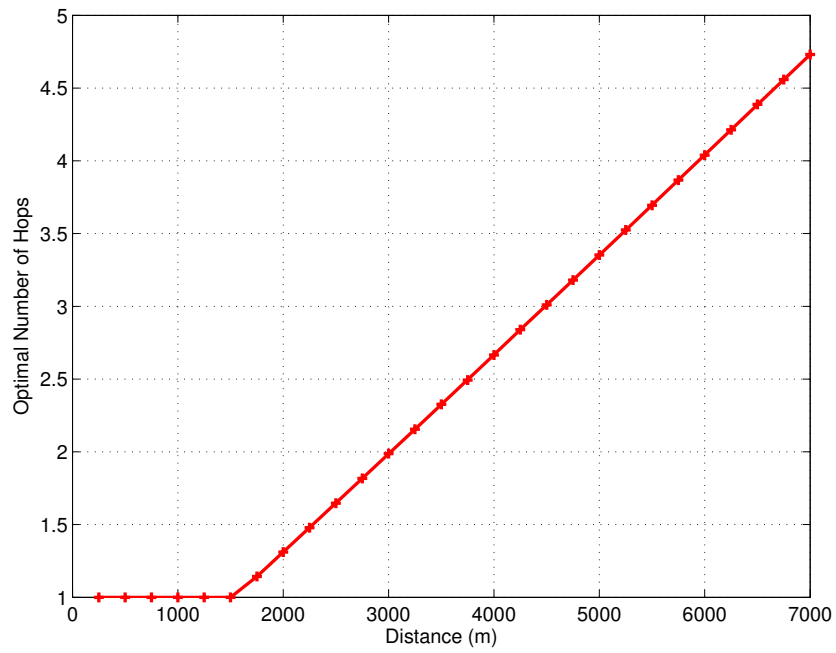


Figure 4.11: Optimal number of hops for convolutional-coded model. The length of information sequence is 224 bits.

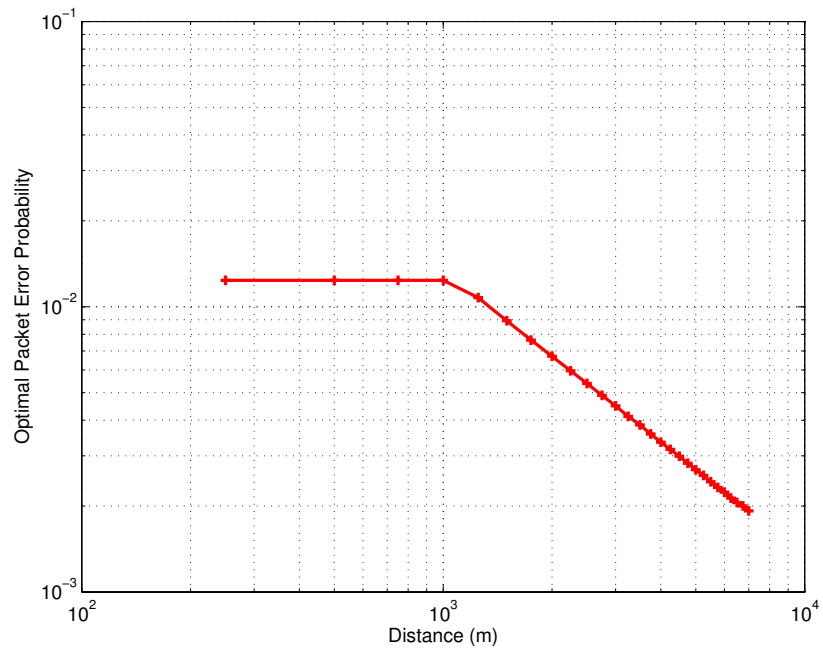


Figure 4.12: Optimal packet error probability for cutoff-rate model. The length of information sequence is 224 bits.

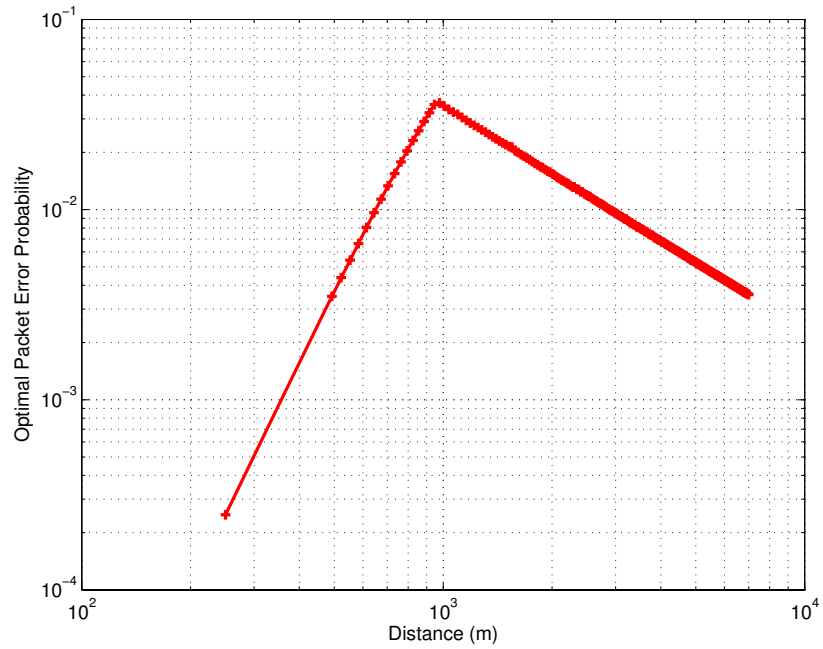


Figure 4.13: Optimal packet error probability for uncoded model. The length of information sequence is 224 bits.

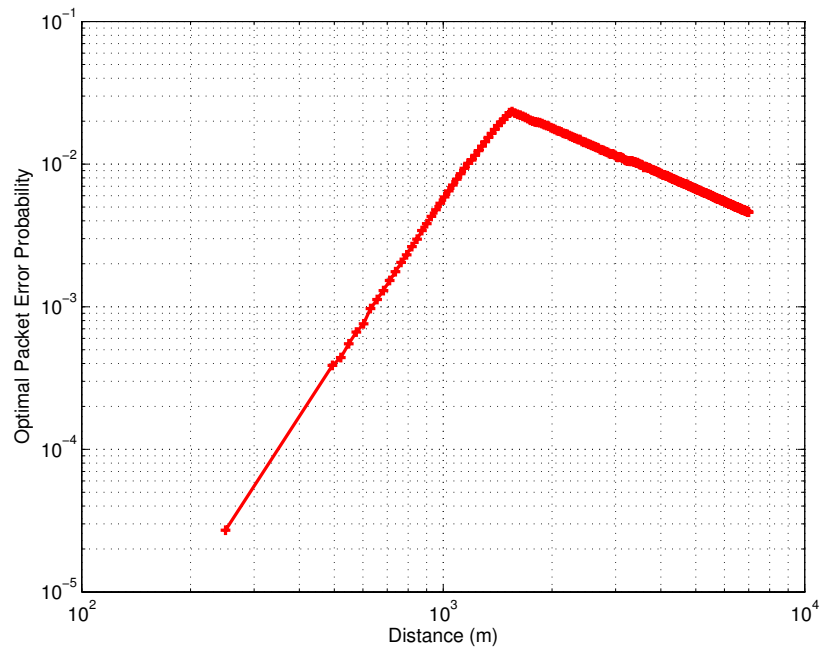


Figure 4.14: Optimal packet error probability for convolutional-coded model. The length of information sequence is 224 bits.

4.4.2 Transport Efficiency for Packet Size 1000 Bits

We now change the length of information sequence from 224 bits to 1000 bits and re-evaluate the transport efficiency for the cutoff-rate model, uncoded model, and convolutional-coded model. The transport efficiency for the threshold model does not depend on the length of information sequence. Therefore the transport efficiency for the threshold model with 1000-bit long information sequence is the same as that for the threshold model with 224-bit long information sequence. The optimized transport efficiency for each fixed k and optimal k (where k is allowed to be a non-integer greater than or equal to 1) for the cutoff-rate model, uncoded model, and convolutional-coded model is given in Figures 4.15 to 4.17. The optimal number of hops is shown in Figures 4.18 to 4.20. The optimal packet error probability is shown in Figures 4.21 to 4.23.

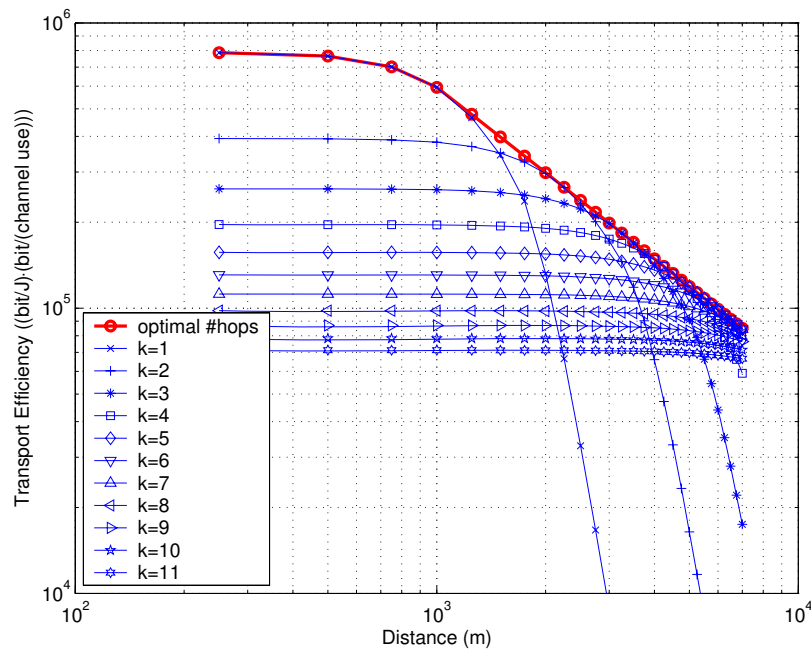


Figure 4.15: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for cutoff-rate model. The length of information sequence is 1000 bits.

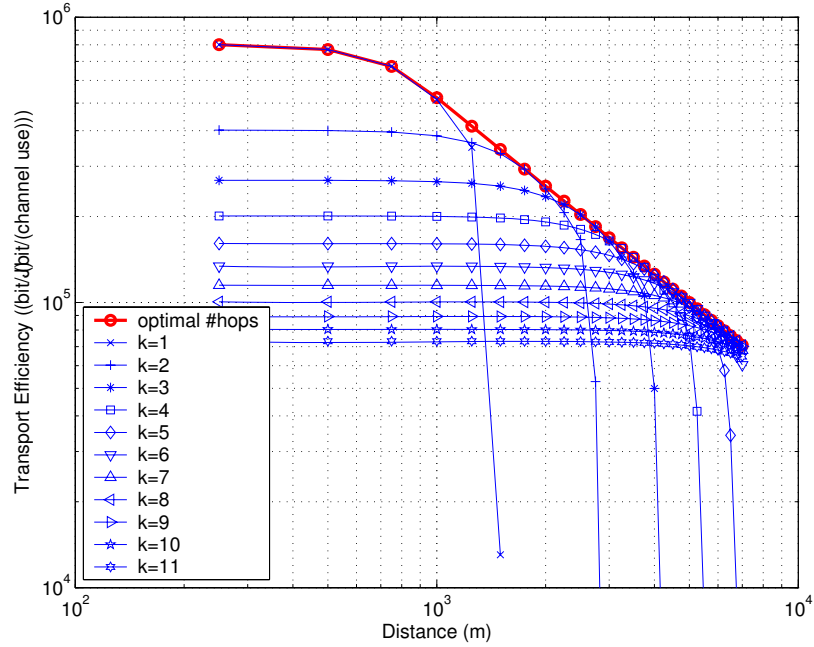


Figure 4.16: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for uncoded model. The length of information sequence is 1000 bits.

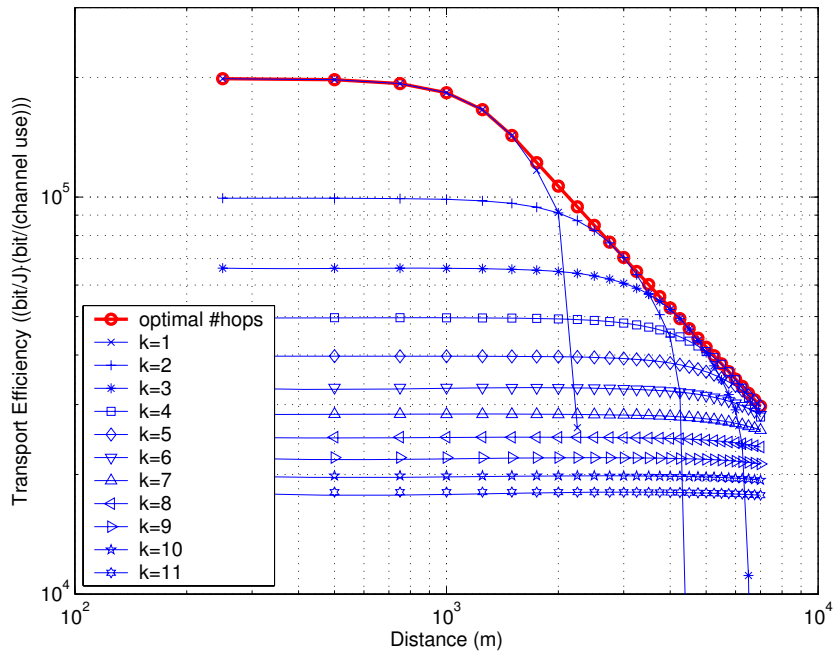


Figure 4.17: Optimized transport efficiency for each fixed k and optimized transport efficiency for optimal k for convolutional-coded model. The length of information sequence is 1000 bits.

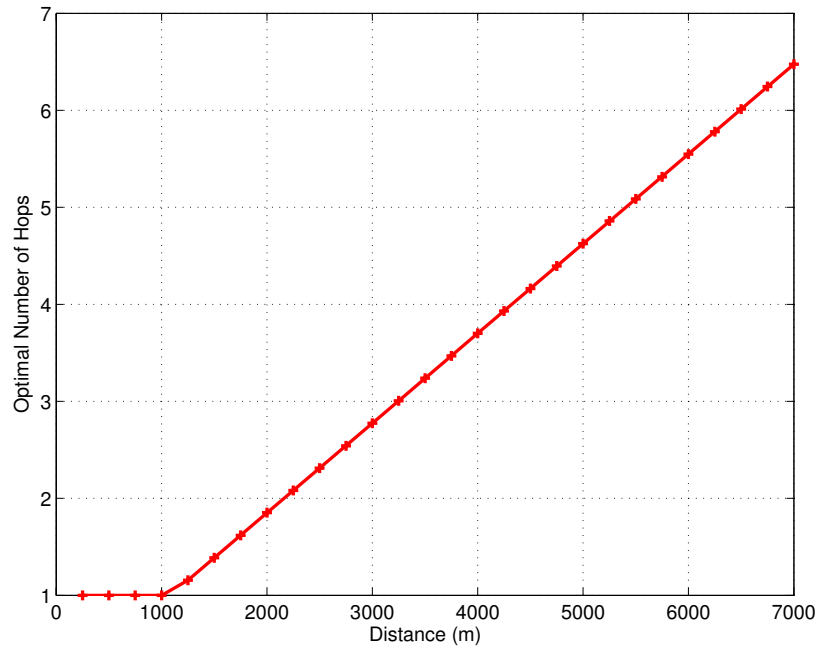


Figure 4.18: Optimal number of hops for cutoff-rate model. The length of information sequence is 1000 bits.

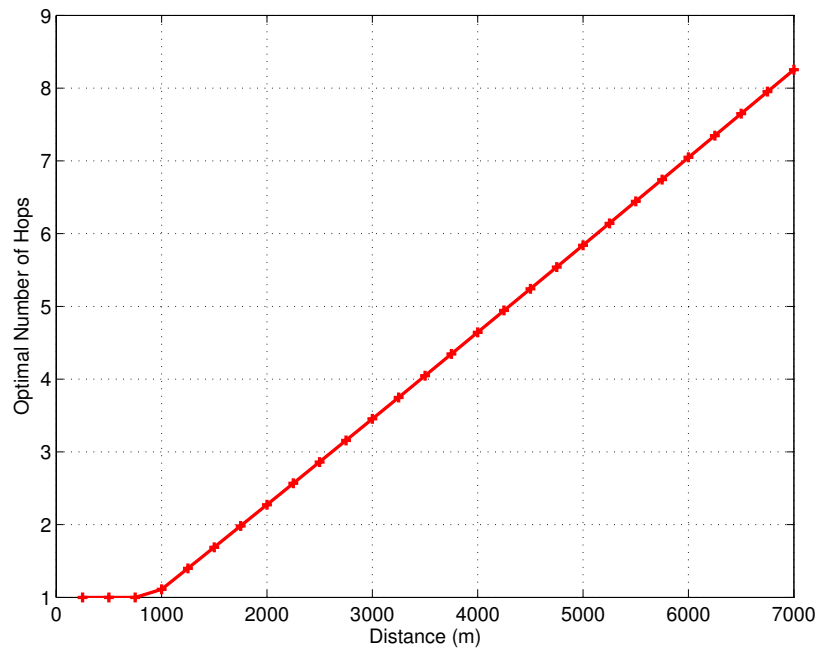


Figure 4.19: Optimal number of hops for uncoded model. The length of information sequence is 1000 bits.

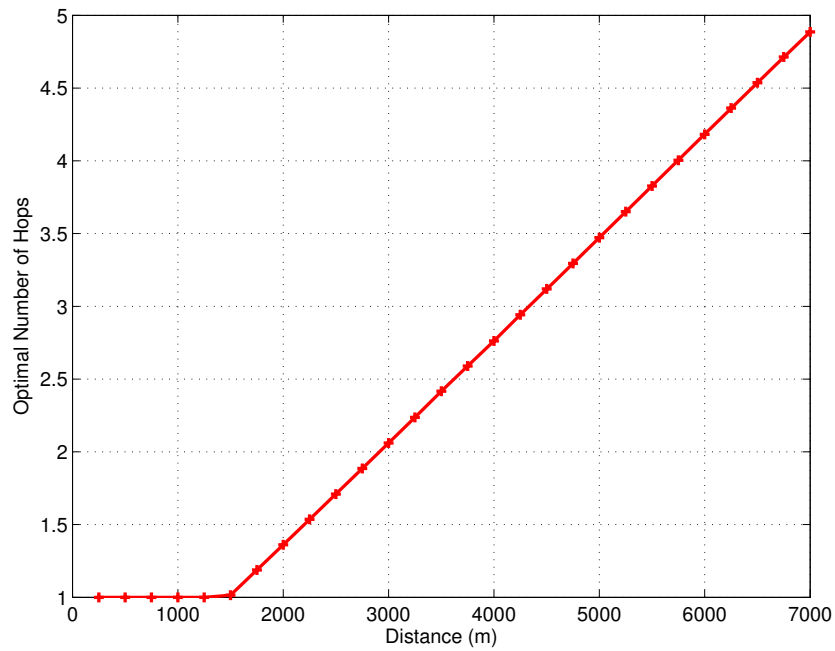


Figure 4.20: Optimal number of hops for convolutional-coded model. The length of information sequence is 1000 bits.

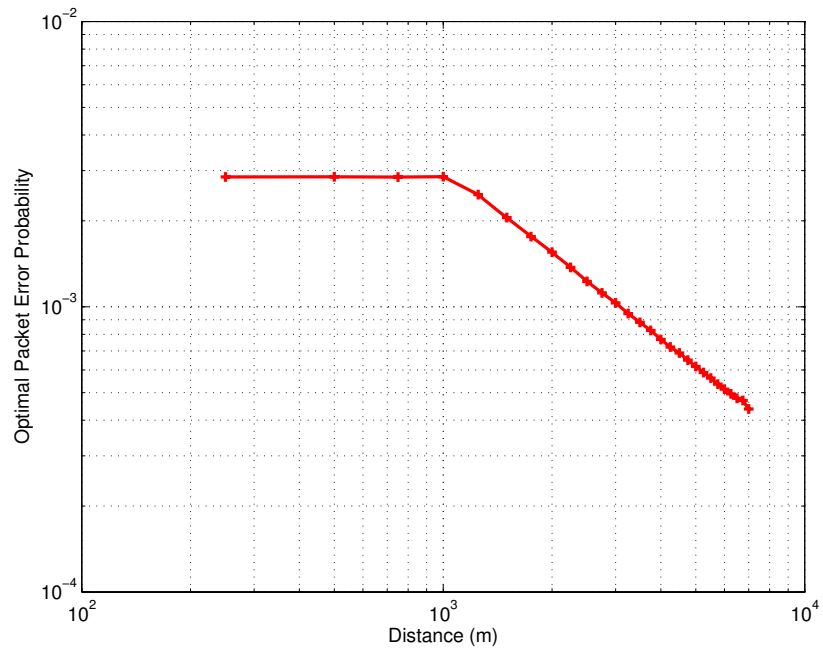


Figure 4.21: Optimal packet error probability for cutoff-rate model. The length of information sequence is 1000 bits.

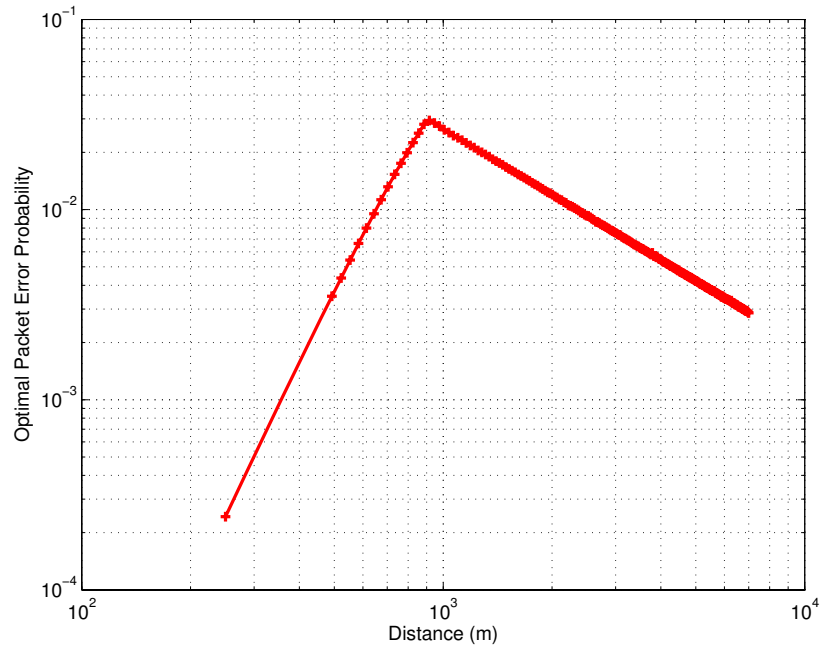


Figure 4.22: Optimal packet error probability for uncoded model. The length of information sequence is 1000 bits.

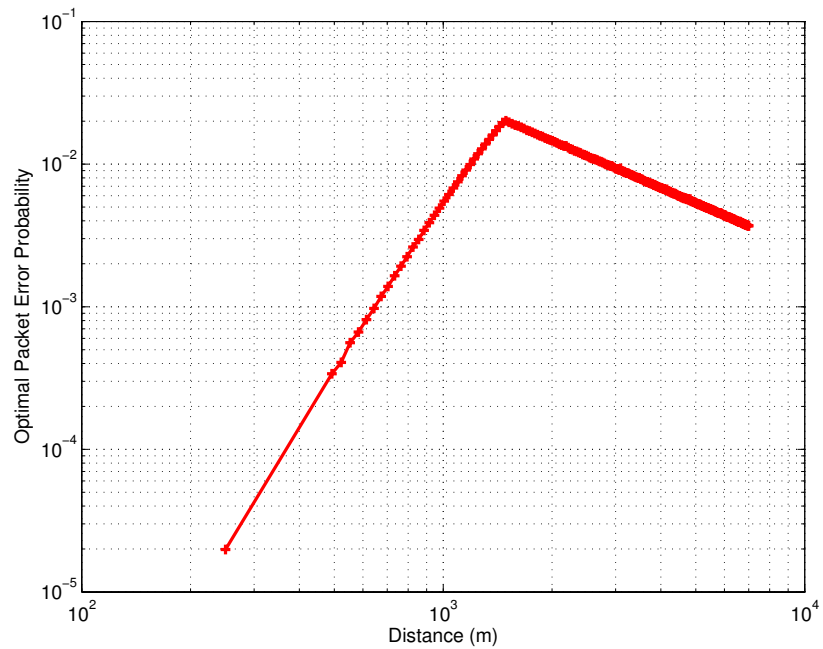


Figure 4.23: Optimal packet error probability for convolutional-coded model. The length of information sequence is 1000 bits.

4.4.3 Comparison of All Models

We now show the optimal transport efficiency for the threshold model, cutoff-rate model, uncoded model, and convolutional-coded model on the same graph in Figure 4.24. We observe that the length of information sequence plays a negligible role in the transport efficiency.

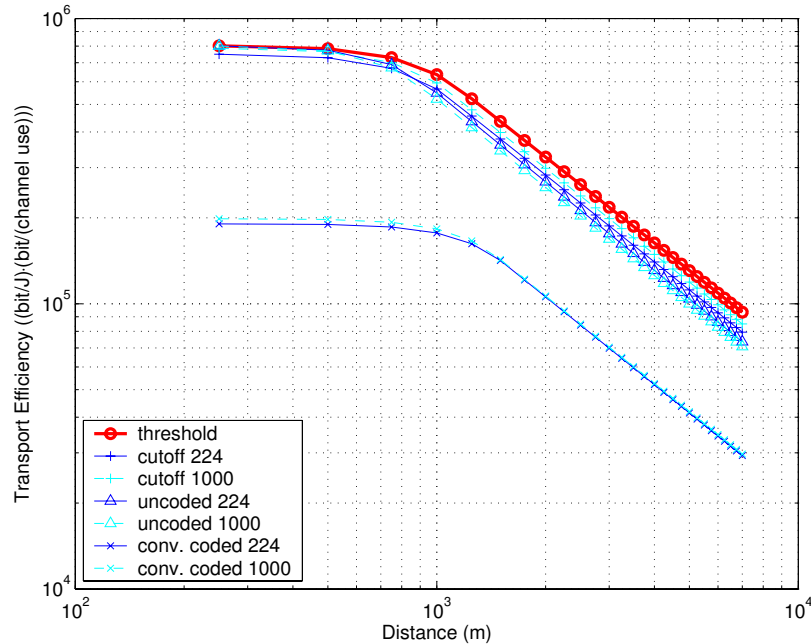


Figure 4.24: Optimal transport efficiency in log-log scale for the threshold model, cutoff-rate model, uncoded model, and convolutional-coded model.

4.4.4 Observation

For a linear network model given in Figure 4.1, where the source and the destination are at both ends, and when relay nodes are placed evenly on a straight line, we observe from our numerical results that $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \sup_{R \in [0,1]} \mu(k, P_{in}, R, d_e)$ for the cutoff-rate model, $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \mu(k, P_{in}, 1, d_e)$ for the uncoded model, and $\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \mu(k, P_{in}, R, d_e)$ for the convolutional-coded model, are approximately inversely proportional to the end-to-end distance d_e when d_e is large,

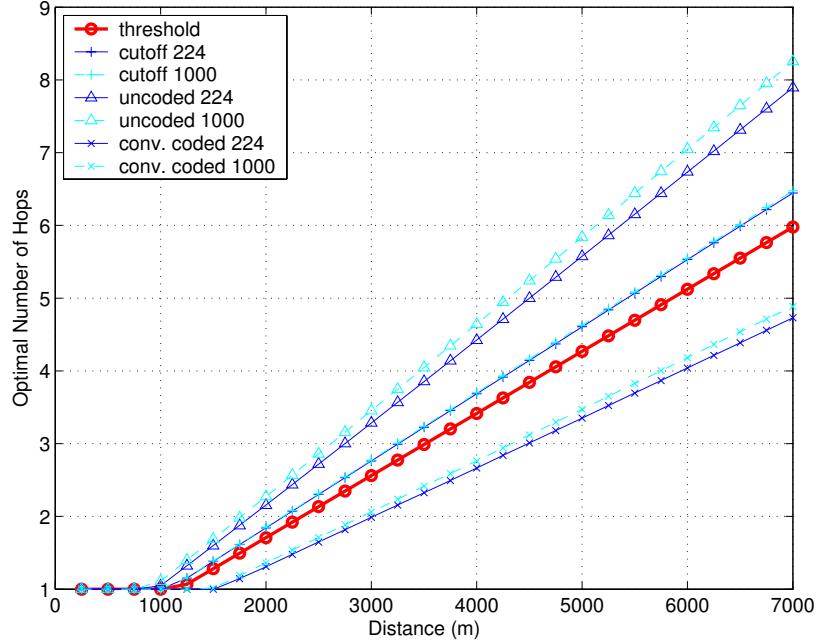


Figure 4.25: Optimal number of hops for the threshold model, cutoff-rate model, uncoded model, and convolutional-coded model.

i.e.,

$$\sup_{k \in \mathcal{K}(d_e)} \sup_{P_{in} \in \mathcal{P}} \sup_{R \in [0,1]} \mu(k, P_{in}, R, d_e) \approx \frac{A}{d_e}, \quad (4.71)$$

where A does not depend on d_e . The supremum over R in (4.71) is not needed for the uncoded model and convolutional-coded model.

4.4.5 Discussion

The reason that the linear network with the uncoded model has better transport efficiency than with the convolutional-coded model is that the optimal adjacent distance is small enough to make high code rate a better choice, but for the convolutional-coded model, the rate is only about $\frac{1}{2}$. We also assumed that the processing energy of an received channel symbol is the same for the uncoded model and convolutional-coded model. Normally the processing energy for the uncoded model is less than that for the convolutional-coded model. This makes the transport

efficiency for the uncoded model given in the above figures a lower bound for the actual transport efficiency for the uncoded model. As we can see in Figure 4.25, in order to achieve high transport efficiency, the uncoded model requires more nodes to relay packets, which may result in higher density requirement for the deployment of a network. This is a drawback to the uncoded model compared with the convolutional-coded model.

4.5 Conclusion and Future Research

We introduced the transport efficiency as a network performance measure that captures both bandwidth efficiency and energy efficiency. Our numerical results showed that the transport efficiency is not very sensitive to packet size. We saw that even though rate $\frac{1}{2}$ convolutional coding saves about 5.5 dB in energy compared with no coding, it does not give high transport efficiency due to its low code rate. Therefore, in order to optimize a communication network based on transport efficiency, we should consider both code rate and energy saving. We proved that the optimal transport efficiency is inversely proportional to d_e for the threshold model and observed from our numerical results that the optimal transport efficiency is approximately inversely proportional to d_e for the cutoff-rate model, uncoded model and convolutional-coded model when d_e is large.

For future research, we may consider turbo codes and other coding techniques. We may also consider fading channels, such as log-normal and Rayleigh fading channels. We may then investigate the transport efficiency under these assumptions and compare it with the results we obtained in this chapter.

CHAPTER V

TRANSPORT EFFICIENCY OF LINEAR NETWORKS WITH SPACE-TIME CODING AND ORDINARY END-FIRE ANTENNA ARRAYS

5.1 Introduction

With the fast development of wireless devices and networks, higher and higher information rates are desired. The IEEE 802.11b standards specify an information rate up to 11 Mb/s, while the newly developed IEEE 802.11g standards specify an information rate up to 54 Mb/s. In order to achieve even higher information rates, spectrally efficient wireless coding and modulation technologies are desired. It's very likely that some form of multiple-input and multiple-output (MIMO) technology will form the basis for future IEEE 802.11n standards. Multiple transmitting and receiving antennae as well as orthogonal frequency division multiplexing (OFDM) are two techniques that are being considered for high-throughput wireless networks. MIMO communication systems have attracted considerable interest due to its potential for high capacity. The combination of error control coding with multiple antennae, such as space-time coding, allows large data rates to be achieved. The use of an antenna array is another approach to achieving high data rates. An antenna array uses

multiple radiating elements to increase the intensity of an electric field in a given direction. In this chapter, we study how space-time coding and an antenna array affect the transport efficiency of linear wireless networks.

Much of the study on the throughput of wireless networks focuses on the performance of particular protocols. In [76] the effect of space-time block codes on the performance of 802.11a wireless local area network was studied via simulation. It was shown that the gain of utilizing space-time block codes at the physical layer gives rise to significant improvement in network layer performance of 801.11a wireless local area networks.

The transport capacity of a network was defined in [31, 32] to measure the performance of a wireless network. The transport capacity has a unit bit·m/s and it measures the average capability of a network to communicate bits from sources to destinations. In [31, 32] the transport capacity was given for a network where n nodes are arbitrarily located on a disk of unit area under two communication models, namely, the protocol model and the physical model, defined by the authors. However, the authors did not consider the bandwidth efficiency and energy efficiency explicitly.

Since in general the performance and energy tradeoff is a very complicated function of network topology and the traffic pattern, we will constrain ourselves to find performance limits for a simple linear network topology, where nodes lie on a straight line. We further assume that there is only unicast traffic between two end nodes on the linear networks. Normally when several transmissions happen at the same time, they cause interference to each other. We are going to investigate the effect of interference on the transport efficiency introduced in Chapter IV, when either space-time coding or an antenna array is used at each node. Space-time coding, even though

having high capacity, may incur large interference to nearby nodes. On the other hand, an antenna array, not possessing as high a capacity as space-time coding, causes less interference to neighboring nodes due to its properly designed amplitude pattern. We will show that space-time coding and an antenna array each yield a better transport efficiency under different circumstances.

The remainder of this chapter is organized as follows. In Section 5.2 we explain how space-time coding and antenna arrays work. In Section 5.3 we introduce an ordinary end-fire antenna array. In Section 5.4 we introduce the interfering network model and analyze the transport efficiency of a linear network for space-time coding and the ordinary end-fire antenna array. In Section 5.5 we present numerical results. We conclude the discussion in Section 5.6.

5.2 Space-Time Coding and Beam-Forming

In this section, we conduct a brief survey of space-time coding and beam-forming from both the information theoretic perspective and practical design point of view. Consider a point-to-point single-hop communication with a frequency non-selective slowly fading channel as given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (5.1)$$

where $\mathbf{x} \in \mathbb{C}^t$ is the transmitted vector, $\mathbf{y} \in \mathbb{C}^r$ is the received vector, $\mathbf{n} \in \mathbb{C}^r$ is the additive noise at the receiving antennae, and $\mathbf{H} \in \mathbb{C}^{r \times t}$ is the channel matrix. Let $Q = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ be the correlation matrix of \mathbf{x} . The energy constraint for \mathbf{x} is $\text{tr}(Q) \leq E$. Note that for each given Q , when the receiver knows the realization of \mathbf{H} and the transmitter does not know the realization of \mathbf{H} , the distribution of \mathbf{x} that maximizes the mutual information between \mathbf{x} and \mathbf{y} is a proper complex Gaussian distribution with mean $\mathbf{0}$.

The purpose of space-time coding is to design signaling \mathbf{x} to exploit the diversity given by \mathbf{H} . Normally, the higher the $\text{rank}(Q)$, the higher the diversity that can be achieved. Even though space-time coding provides very high information capacity by exploiting the diversity of channel matrix \mathbf{H} , the signal processing complexity can be very high at the transmitter and at the receiver. Therefore it is desirable to find some alternative scheme that provides high enough capacity with much less complexity. A transmission strategy is said to have δ -**fold diversity** if $\text{rank}(Q) = \delta$. A transmission strategy is said to be **beam-forming** if $\text{rank}(Q) = 1$. Note that the definition of diversity (or beam-forming) itself does not depend on the number of receiving antennae. Similar definitions can be found in [56, 57, 86, 40]. Recall that for space-time coding, there is no restriction on $\text{rank}(Q)$. Therefore beam-forming is a special case of space-time coding. Consequently, beam-forming may not achieve as high a capacity as that of space-time coding. In order for beam-forming to be effective, the transmitter usually has to utilize some knowledge about the status of the channel matrix \mathbf{H} , so beam-forming generally requires some form of feedback about the channel in order to achieve a capacity close to that for space-time coding. On the other hand, for space-time coding, the transmitter does not need the knowledge about the channel, even though an even higher capacity is possible if the transmitter has some feedback about the channel.

There are two categories of feedback in general, namely, *perfect feedback* and *imperfect feedback*. The perfect feedback is the case when the transmitter knows exactly the realization of \mathbf{H} , while the imperfect feedback is the case when the transmitter has partial knowledge of the channel. Most commonly, partial knowledge refers to some statistics of the channel. By utilizing feedback about the channel from the receiver to the transmitter, the transmitter can apply appropriate distributions

of \mathbf{x} , i.e., adjust the transmitted signal on each of the transmitting antennae.

We first describe imperfect feedback. There are in general two types of imperfect feedback, one is called *mean feedback*, the other is called *covariance feedback*. In order to describe these two types of imperfect feedback, we need several notations. Suppose now there is only one receiving antenna. Let the row vector \mathbf{H} be a proper complex Gaussian random vector with mean μ and covariance matrix K . Let $K = U_K \Lambda_K (U_K)^\dagger$, where U_K is the eigenvector matrix of K and Λ_K is a diagonal matrix containing the eigenvalues of K on the diagonal in descending order.

For the *mean feedback*, we have that $K = \alpha^2 I$, and both μ and K are known to the transmitter. The feedback information is completely contained in μ . It is shown in [86] that in this case the distribution of \mathbf{x} maximizing the mutual information between \mathbf{x} and \mathbf{y} is that \mathbf{x} is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Q such that (i) $Q = U_Q \Lambda_Q (U_Q)^\dagger$, where U_Q is a unitary matrix with the first column vector being $\mu^\dagger / \|\mu^\dagger\|$ and the other column vectors being arbitrarily chosen; and (ii) Λ_Q is a real diagonal matrix with the first entry on the diagonal being λ_1 and all other entries on the diagonal being $(E - \lambda_1)/(t - 1)$. The exact value of λ_1 for Q to be optimal has to be determined numerically. The beam-forming scheme under mean feedback corresponds to setting $\lambda_1 = E$ and is shown to be optimal when the feedback signal-to-noise ratio $\|\mu\|^2/\alpha^2$ is greater than a threshold.

For the *covariance feedback*, we have that $\mu = 0$, and both μ and K are known to the transmitter. The feedback information is completely contained in K . It is shown in [86] that the distribution of \mathbf{x} maximizing the mutual information between \mathbf{x} and \mathbf{y} is that \mathbf{x} is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Q such that (i) $Q = U_Q \Lambda_Q (U_Q)^\dagger$, where U_Q is a unitary matrix satisfying

$U_Q = U_K$ and Λ_Q is a real diagonal matrix; and (ii) the diagonal entries of Λ_Q are arranged in descending order, the same order for the diagonal entries of Λ_K . The exact values of diagonal entries of Λ_Q have to be determined numerically. In [40] the authors extend the capacity result for the covariance feedback to the case when there is more than one receiving antenna, with the assumption that all row vectors of \mathbf{H} are independent and identically distributed proper complex Gaussian random vectors with mean $\mathbf{0}$ and covariance matrix K . The optimal Q when there is more than one receiving antenna is shown to be the same as the optimal Q when there is only one receiving antenna. The beam-forming scheme under covariance feedback corresponds to choosing the first entry on the diagonal of Λ_Q to be E and all other entries on the diagonal of Λ_Q to be 0. The necessary and sufficient conditions for beam-forming to be optimal are also given in [40].

In general, for both *mean feedback* and *covariance feedback*, if there is a moderate disparity between the strengths of different paths from the transmitter to the receiver, it is nearly optimal to employ the simple beam-forming strategy of transmitting all available power in the direction given by the eigenvector of Q which the mean or covariance feedback indicates the strongest.

We now consider *perfect feedback*. Suppose there is only one receiving antenna, i.e., $r = 1$, and there is perfect feedback from the receiver to the transmitter, i.e., the transmitter knows the realization of \mathbf{H} , it is shown in [56, 57] that beam-forming is the optimal communication strategy and is equivalent to maximum ratio combining at the transmitter, i.e., the vector of the transmitted energy assigned to transmitting antennae is proportional to $\mathbf{H}^\dagger / \|\mathbf{H}^\dagger\|$.

We finally consider the case when there is no feedback from the receiver to the transmitter. Suppose that $r > 1$ and all entries of \mathbf{H} are independent and identically

distributed proper complex Gaussian random variables with mean 0 and variance α^2 . Then all row vectors of \mathbf{H} are independent and identically distributed proper complex Gaussian random vectors with mean $\mathbf{0}$ and covariance matrix $K = \alpha^2 I_t$. This corresponds to the case where there is no feedback from the receiver to the transmitter. In [26, 80] the authors derived the capacity of the channel given in (5.1) under this condition. It was shown that the optimal distribution for \mathbf{x} that maximizes mutual information between \mathbf{x} and \mathbf{y} is that \mathbf{x} is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $Q = \frac{E}{t} I_t$. Consequently, $\text{rank}(Q) = t$. According to Theorem C.3.1, we conclude that the beam-forming scheme, which requires $\text{rank}(Q) = 1$, can not be optimum under the above assumptions for \mathbf{H} . Since we are going to analyze network transport efficiency based on the above assumptions for \mathbf{H} , we will not consider beam-forming for the rest of the chapter.

Besides the above information-theoretic study of the multiple-antenna communication, the practical space-time coding schemes have also been designed and their performance has been analyzed in literature. In [3] a simple transmission scheme, now known as the ‘‘Alamouti code’’, for two transmitting antennae and one receiving antenna communication system using maximum likelihood detector is described. It is shown that the ‘‘Alamouti code’’ provides the same the diversity order as a communication system with one transmitting antenna and two receiving antennae. In [78, 79] the authors show that the pair-wise sequence error rate between two distinct code sequences is determined by matrices constructed from these two distinct code sequences. The authors use the minimum rank and minimum determinant among the matrices constructed from all pairs of distinct sequences to design trellis codes for the multiple transmitting antennae and multiple receiving antennae communication systems.

We will use the capacity result when there is no feedback from the receiver to the transmitter in our analysis of network transport efficiency in later sections. We are more interested in the information-theoretic aspect of results rather the practical design issues.

5.3 Ordinary End-Fire Antenna Array

Given t transmitting antennae, communication system designers can either place the radiating elements with enough separation to exploit the spatial diversity and use space-time coding or configure the radiating elements in an electrical and geometrical way to form an *antenna array* to enhance directivity. The simplest and most practical antenna array is a *linear array* where all radiating elements lie on a straight line. We are interested in a linear antenna array which consists of identical radiating elements with uniform spacing as shown in Figure 5.1. Let d be this uniform spacing. Appropriate current excitation can be applied to the radiating elements to make each of them have different amplitude and phase. The progressive phase is defined to be the phase lead of a radiating element over its immediate predecessor. An array of identical elements all of identical magnitude and each with a constant progressive phase is referred to as a uniform array [6]. Let γ be this constant progressive phase.

We can characterize the electric field radiated by a linear uniform antenna array through the *array factor* of the antenna array. The **array factor** is defined such that the total electric field at a far zone can be formed by multiplying the array factor with the electric field radiated by a single element at a reference point of the antenna array. Let \vec{E}_1 be the electric field radiated by a single element at a reference point of the antenna array (usually the origin). Then the total electric field \vec{E}_{total} radiated by the antenna array, assuming no coupling between the radiating elements, can be

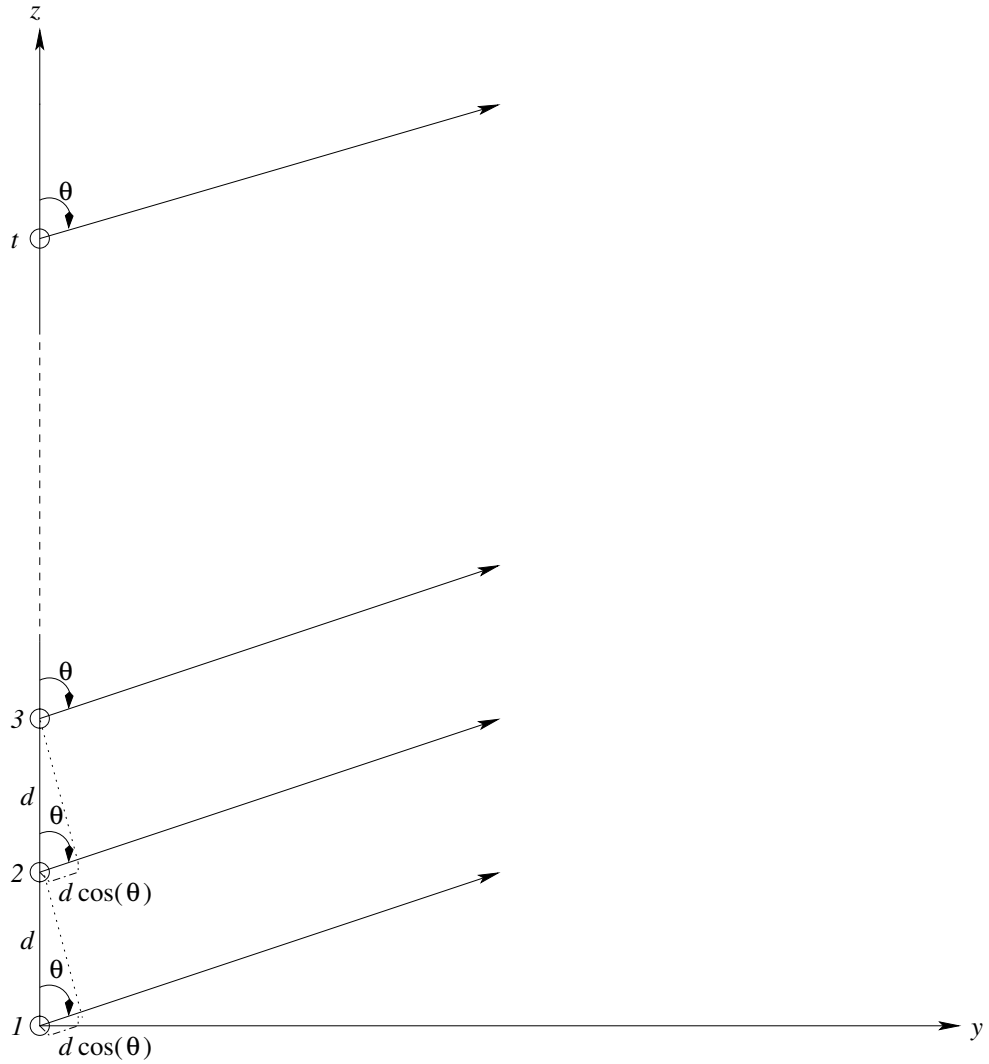


Figure 5.1: Far-zone geometry for a linear antenna array, where d is the distance between adjacent elements in the antenna array, θ is the angle between the line connecting the reference point of the antenna array (usually the origin) with the point of field observation and the line where radiating elements lie.

obtained via pattern multiplication of the array factor (AF) with \vec{E}_1 , i.e.,

$$\vec{E}_{total} = \vec{E}_1 \times (AF). \quad (5.2)$$

The line connecting the reference point of the antenna array with the point of field observation and the line where radiating elements lie form an angle θ as shown in Figure 5.1. The array factor (AF) of a linear uniform array when the reference point is at the origin is given by [6]

$$(AF) = \sum_{l=1}^t e^{i(l-1)\left(\frac{2\pi}{\lambda}d \cos(\theta) + \gamma\right)}, \quad (5.3)$$

where λ is the wavelength of the electromagnetic wave. Note that (AF) is a function of θ , γ and d . If the reference point is chosen to be the physical center of the linear uniform antenna array, the array factor in (5.3) reduces to

$$(AF) = \frac{\sin\left(\frac{t}{2}\left(\frac{2\pi}{\lambda}d \cos(\theta) + \gamma\right)\right)}{\sin\left(\frac{1}{2}\left(\frac{2\pi}{\lambda}d \cos(\theta) + \gamma\right)\right)}. \quad (5.4)$$

Many kinds of linear uniform arrays can be formed by different choices of γ and d . The linear uniform array we are interested in is called the *ordinary end-fire array*. To have only one maximum in the array factor of the end-fire array and to avoid any grating lobes, the maximum spacing d_{max} between the radiating elements should satisfy $d_{max} < \lambda/2$. To direct this maximum toward $\theta = 0$, we have

$$\left(\frac{2\pi}{\lambda}d \cos(\theta) + \gamma\right)\Big|_{\theta=0^\circ} = \frac{2\pi}{\lambda}d + \gamma = 0 \quad \Rightarrow \quad \gamma = -\frac{2\pi}{\lambda}d. \quad (5.5)$$

The array factor is then given by

$$(AF) = \frac{\sin\left(\frac{t}{2}\frac{2\pi}{\lambda}d(\cos(\theta) - 1)\right)}{\sin\left(\frac{1}{2}\frac{2\pi}{\lambda}d(\cos(\theta) - 1)\right)}. \quad (5.6)$$

The normalized array factor is given by

$$\overline{(AF)} = \frac{1}{t} \frac{\sin\left(\frac{t}{2}\frac{2\pi}{\lambda}d(\cos(\theta) - 1)\right)}{\sin\left(\frac{1}{2}\frac{2\pi}{\lambda}d(\cos(\theta) - 1)\right)}. \quad (5.7)$$

A typical amplitude pattern $\left| \overline{(AF)} \right|$ of the ordinary end-fire array is shown in Figure 5.2, where we show the amplitude pattern in a two dimensional space. If we rotate the amplitude pattern given in Figure 5.2 around the z -axis, we will have the amplitude pattern in a three dimensional space. The length of the vector from the origin to the point on the curve in Figure 5.2 represents the amplitude of $\overline{(AF)}$ in that given direction.

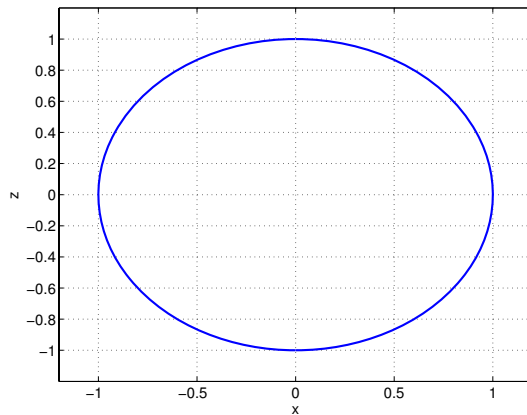
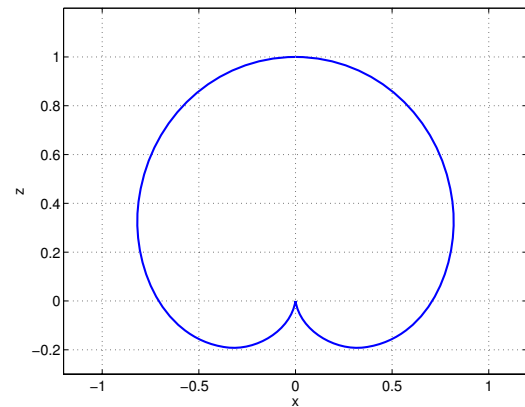
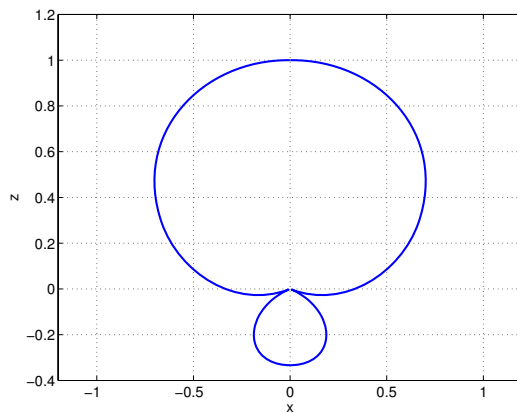
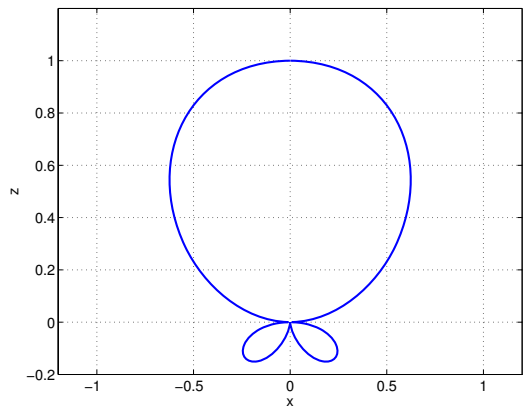
(a) $t = 1$ (b) $t = 2$ (c) $t = 3$ (d) $t = 4$

Figure 5.2: Amplitude pattern $\left| \overline{(AF)} \right|$ of the ordinary end-fire antenna for t radiating elements when $d = \frac{\lambda}{4}$.

In essence, we can treat the ordinary end-fire antenna array as a single antenna that has an amplitude pattern given in (5.6). The benefit of utilizing transmitting antenna array is that if the transmitter only knows the direction of the receiver, it can form an amplitude pattern of electric field in favor of the receiver. Usually as a bonus, it will cause less interference to other nearby receivers as well.

The array factor (AF) given in (5.3) can be viewed as the inner product of two vectors. The first vector

$$v = \left[1, e^{i\frac{2\pi}{\lambda}d\cos(\theta)}, \dots, e^{i(t-1)\frac{2\pi}{\lambda}d\cos(\theta)} \right]^T \quad (5.8)$$

relates to channel and the second vector

$$u = \left[1, e^{i\gamma}, \dots, e^{i(t-1)\gamma} \right]^T. \quad (5.9)$$

relates to the excitation. We can write (AF) as

$$(AF) = v^T u. \quad (5.10)$$

From the total electric field \vec{E}_{total} given in (5.2), we can derive a channel model for communication purposes. Let s be the distance between the reference point of the antenna array and the point of field observation. Define

$$\alpha(s) = \sqrt{\beta(s)}, \quad (5.11)$$

where $\beta(\cdot)$ is the propagation loss function for power given in (4.9). For a deterministic channel, we have a channel matrix H (now a scalar)

$$H = (\alpha)(AF) = \alpha v^T u. \quad (5.12)$$

Thus if \mathbf{x} is the input random signal to the antenna array and \mathbf{y} is the received signal at the point of field observation, we have that

$$\mathbf{y} = H\mathbf{x}. \quad (5.13)$$

Recall that for the channel matrix \mathbf{H} given in (5.1) for space-time coding, we assumed that all entries of \mathbf{H} are independent and identically distributed proper complex Gaussian random variables with mean 0 and variance α^2 . In order to make fair comparisons between the channel for the ordinary end-fire antenna array and the channel for space-time coding, we assume that for the ordinary end-fire antenna array the electric signal from each radiating element of the antenna array experiences a common multi-path fading due to the fact that the radiating elements are close to each other. We assume that there are enough local scatters around the point of field observation to make the received signal Rayleigh faded. The channel matrix \mathbf{H} for the ordinary end-fire antenna array with Rayleigh fading is

$$\mathbf{H} = (\alpha) |(AF)| \mathbf{z}(\theta) = \alpha |v^T u| \mathbf{z}(\theta), \quad (5.14)$$

where $\mathbf{z}(\theta)$ is a proper complex Gaussian random variable with mean 0 and variance 1. The received signal is

$$\mathbf{y} = \mathbf{H}\mathbf{x}. \quad (5.15)$$

Note that $u\mathbf{x}$ is the transmitted vector out of the radiating elements of the antenna array. Therefore we should impose the same energy constraint on $u\mathbf{x}$ as we have done for the transmitted vector for space-time coding. Thus we have

$$\mathbb{E} [(u\mathbf{x})^\dagger (u\mathbf{x})] \leq E, \quad (5.16)$$

where E is the energy constraint on the transmitter. Therefore

$$\mathbb{E} [\mathbf{x}^* \mathbf{x}] u^\dagger u = t \mathbb{E} [\mathbf{x}^* \mathbf{x}] \leq E. \quad (5.17)$$

Equivalently,

$$\mathbb{E} [\mathbf{x}^* \mathbf{x}] \leq \frac{E}{t}. \quad (5.18)$$

For the ordinary end-fire antenna array with maximum amplitude directed to 0° , we have

$$v = \left[1, e^{i\frac{2\pi}{\lambda}d\cos(\theta)}, \dots, e^{i(t-1)\frac{2\pi}{\lambda}d\cos(\theta)} \right]^T \quad (5.19)$$

and

$$u = \left[1, e^{-i\frac{2\pi}{\lambda}d}, \dots, e^{-i\frac{2\pi}{\lambda}(t-1)d} \right]^T. \quad (5.20)$$

Furthermore, when $d = \frac{\lambda}{4}$, as is assumed in our analysis, we have

$$v = \left[1, e^{i\pi/2\cos(\theta)}, \dots, e^{i(t-1)\pi/2\cos(\theta)} \right]^T \quad (5.21)$$

and

$$u = \left[1, e^{-i\pi/2}, \dots, e^{-i(t-1)\pi/2} \right]. \quad (5.22)$$

Consequently,

$$\mathbf{H} = \frac{\sin\left(\frac{t\pi}{4}(\cos(\theta) - 1)\right)}{\sin\left(\frac{\pi}{4}(\cos(\theta) - 1)\right)} \alpha \mathbf{z}(\theta). \quad (5.23)$$

Our analysis of network transport efficiency when each node uses the ordinary end-fire antenna array is based on the channel matrix given in (5.23). For the same number of radiating elements at the transmitter, we can compare the network transport efficiency for space-time coding and for the ordinary end-fire antenna array and draw conclusion about the optimal operating conditions for the two schemes.

5.4 Interfering Communication for Linear Networks

5.4.1 Interfering Linear Networks

Consider a network with topological model shown in Figure 5.3. For the scenario considered in this chapter, each node on linear network A treats the signal it receives from the nodes on linear network B as noise, and vice versa. The network model we present here can be two actual linear networks in the real world or they correspond topologically to two paths within some larger wireless network containing them. For

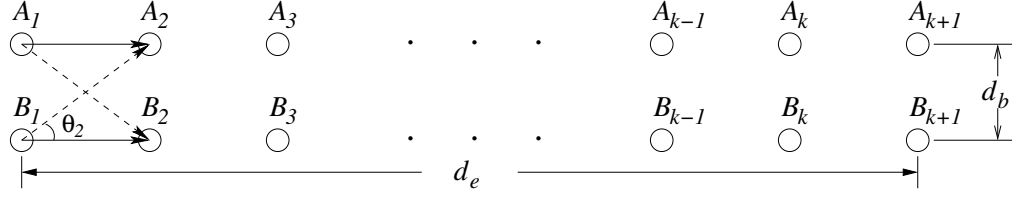


Figure 5.3: Two parallel linear wireless networks that are separated by distance d_b . The upper linear network is referred to as linear network A and the lower linear network is referred to as linear network B .

linear network A and B , we have very similar assumptions as the ones given in Chapter IV. In addition, we have the following assumptions.

- Nodes are evenly spaced on each linear network with distance between adjacent nodes being d_a . Two linear networks are separated by d_b .
- Each node on linear network A and B has t transmitting antennae.
- All nodes on both linear networks use the maximum input power P_{max} and consume the same amount of energy to transmit a channel symbol. Let the symbol duration be T_s . Therefore the consumed energy for a transmitted symbol is

$$E_c = f_c(P_{max})T_s. \quad (5.24)$$

The amplifier output energy for a transmitted symbol is

$$E_o = f_o(P_{max})T_s. \quad (5.25)$$

- The transmitted signal from each node experiences large-scale propagation loss and fading. The propagation loss model when the transmitter and the receiver are separated by distance s is

$$P_r = \beta(s)P_t, \quad (5.26)$$

where $\beta(\cdot)$ is the large-scale propagation loss depending on s .

- We assume that each receiver consumes E_p amount of energy to process a received channel symbol. In practice, the processing energy per symbol is related to the packet error rate, but we will not consider this factor in our analysis. We assume that E_p is a constant regardless of the coding and modulation techniques used.
- Each node A_i on linear network A sends its information to its immediate downstream node A_{i+1} , meanwhile, each node B_i on linear network B , sends its information to its immediate downstream node B_{i+1} , where $i = 1, \dots, k$.
- The transmission from B_i to B_{i+1} causes interference at A_{i+1} when A_{i+1} is receiving information from A_i . Similarly, the transmission from A_i to A_{i+1} causes interference at B_{i+1} when B_{i+1} is receiving information from B_i .
- For each linear network, only one node on that network can transmit at a time.

Because of the symmetry of linear network A and linear network B , we will only analyze the transport efficiency of linear network A , with the interference from linear network B . Let \mathbf{y} be the received signal by A_2 on linear network A , $\mathbf{x}^{(1)}$ be the transmitted signal from A_1 to node A_2 , $\mathbf{x}^{(2)}$ be the transmitted signal from B_1 to node B_2 , $\mathbf{H}^{(1)}$ be the channel matrix from A_1 to node A_2 , $\mathbf{H}^{(2)}$ be the channel matrix from B_1 to node A_2 , and \mathbf{n} is the additive noise at the receiver. The received signal \mathbf{y} is given by

$$\mathbf{y} = \mathbf{H}^{(1)}\mathbf{x}^{(1)} + \mathbf{H}^{(2)}\mathbf{x}^{(2)} + \mathbf{n}. \quad (5.27)$$

For different transmission strategies, such as space-time coding and an ordinary end-fire antenna array, we have different expressions for $\mathbf{H}^{(1)}$, $\mathbf{H}^{(2)}$, $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, and \mathbf{n} . We

will give specific evaluation of $\mathbf{H}^{(1)}$ and $\mathbf{H}^{(2)}$ in the following sections. Define

$$\alpha_1 = \sqrt{\beta (d_e/k)}; \quad (5.28)$$

$$\alpha_2 = \sqrt{\beta \left(\sqrt{d_b^2 + (d_e/k)^2} \right)}. \quad (5.29)$$

Then α_1^2 is the propagation loss of power from A_1 to A_2 and α_2^2 is the propagation loss of power from B_1 to A_2 . We assume threshold model (capacity-achieving transmitter and receiver) for each node in the network for the evaluation of transport efficiency.

5.4.2 Transport Efficiency for Space-Time Coding

When the spacing between the transmitting elements and the spacing between the receiving elements are large so that the channel matrix has independent and identically distributed entries, we utilize space-time coding for communication. As shown in Figure 5.3, when node A_1 wants to communicate with node A_2 , the propagation loss of power for the desired signal is α_1^2 . Due to many local scatters at the receiver, the channel with space-time coding can be modeled by the channel matrix

$$\mathbf{H}^{(1, stc)} = \alpha_1 \mathbf{Z}^{(1)}, \quad (5.30)$$

where $\mathbf{Z}^{(1)} = [\mathbf{z}_{i,j}^{(1)}]$ for $i = 1, \dots, r$ and $j = 1, \dots, t$. Each entry $\mathbf{z}_{i,j}^{(1)}$ is a proper complex Gaussian random variable with mean 0 and variance 1. All $\mathbf{z}_{i,j}^{(1)}$, $i = 1, \dots, r$ and $j = 1, \dots, t$ are independent. We assume that the input vector $\mathbf{x}^{(1, stc)}$ to the transmitting antennae is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $C_{\mathbf{x}^{(1, stc)}, \mathbf{x}^{(1, stc)}}$, where

$$C_{\mathbf{x}^{(1, stc)}, \mathbf{x}^{(1, stc)}} = \frac{E_o}{t} \mathbf{I}_t. \quad (5.31)$$

Then

$$\text{tr} \left(C_{\mathbf{x}^{(1, stc)}, \mathbf{x}^{(1, stc)}} \right) = E_o. \quad (5.32)$$

When node B_1 wants to communicate with B_2 , it causes interference to A_2 . The propagation loss of power for the interference from node B_1 to A_2 is α_2^2 . Due to local scatters at the receiver, the interference channel can be modeled by the channel matrix

$$\mathbf{H}^{(2, stc)} = \alpha_2 \mathbf{Z}^{(2)}, \quad (5.33)$$

where $\mathbf{Z}^{(2)} = [\mathbf{z}_{i,j}^{(2)}]$ for $i = 1, \dots, r$ and $j = 1, \dots, t$. Each entry $\mathbf{z}_{i,j}^{(2)}$ is a proper complex Gaussian random variable with mean 0 and variance 1. All $\mathbf{z}_{i,j}^{(2)}$, $i = 1, \dots, r$ and $j = 1, \dots, t$ are independent. We assume that the input vector $\mathbf{x}^{(2, stc)}$ to the transmitting antennae is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $C_{\mathbf{x}^{(2, stc)}, \mathbf{x}^{(2, stc)}}$, where

$$C_{\mathbf{x}^{(2, stc)}, \mathbf{x}^{(2, stc)}} = \frac{E_o}{t} I_t. \quad (5.34)$$

Then

$$\text{tr} (C_{\mathbf{x}^{(2, stc)}, \mathbf{x}^{(2, stc)}}) = E_o. \quad (5.35)$$

Define

$$\mathbf{v}^{(stc)} = \mathbf{H}^{(2, stc)} \mathbf{x}^{(2, stc)}. \quad (5.36)$$

Then we have

$$\begin{aligned} \mathbb{E} [\mathbf{v}^{(stc)}] &= \mathbb{E} [\mathbf{H}^{(2, stc)} \mathbf{x}^{(2, stc)}] \\ &= \mathbb{E} [\mathbf{H}^{(2, stc)}] \mathbb{E} [\mathbf{x}^{(2, stc)}] \\ &= \mathbf{0} \end{aligned} \quad (5.37)$$

and the covariance matrix

$$\begin{aligned} C_{\mathbf{v}^{(stc)}, \mathbf{v}^{(stc)}} &= \mathbb{E} \left[(\mathbf{v}^{(stc)} - \mathbb{E} [\mathbf{v}^{(stc)}]) (\mathbf{v}^{(stc)} - \mathbb{E} [\mathbf{v}^{(stc)}])^\dagger \right] \\ &= \mathbb{E} \left[\mathbf{v}^{(stc)} (\mathbf{v}^{(stc)})^\dagger \right] \end{aligned}$$

$$\begin{aligned}
&= \mathbb{E} \left[\mathbf{H}^{(2, stc)} \mathbf{x}^{(2, stc)} (\mathbf{x}^{(2, stc)})^\dagger (\mathbf{H}^{(2, stc)})^\dagger \right] \\
&= \mathbb{E} \left[\mathbb{E} \left[\mathbf{H}^{(2, stc)} \mathbf{x}^{(2, stc)} (\mathbf{x}^{(2, stc)})^\dagger (\mathbf{H}^{(2, stc)})^\dagger \mid \mathbf{H}^{(2, stc)} \right] \right] \\
&= \mathbb{E} \left[\mathbf{H}^{(2, stc)} \mathbb{E} \left[\mathbf{x}^{(2, stc)} (\mathbf{x}^{(2, stc)})^\dagger \mid \mathbf{H}^{(2, stc)} \right] (\mathbf{H}^{(2, stc)})^\dagger \right] \quad (5.38)
\end{aligned}$$

$$= \mathbb{E} \left[\mathbf{H}^{(2, stc)} \mathbb{E} \left[\mathbf{x}^{(2, stc)} (\mathbf{x}^{(2, stc)})^\dagger \right] (\mathbf{H}^{(2, stc)})^\dagger \right] \quad (5.39)$$

$$\begin{aligned}
&= \mathbb{E} \left[\mathbf{H}^{(2, stc)} \frac{E_o}{t} I_t (\mathbf{H}^{(2, stc)})^\dagger \right] \\
&= \frac{E_o}{t} \mathbb{E} \left[\mathbf{H}^{(2, stc)} (\mathbf{H}^{(2, stc)})^\dagger \right] \\
&= \frac{E_o}{t} \mathbb{E} \left[\alpha_2^2 \mathbf{Z}^{(2)} (\mathbf{Z}^{(2)})^\dagger \right] \\
&= \frac{\alpha_2^2 E_o}{t} \mathbb{E} \left[\mathbf{Z}^{(2)} (\mathbf{Z}^{(2)})^\dagger \right] \\
&= \frac{\alpha_2^2 E_o}{t} t I_r \\
&= \alpha_2^2 E_o I_r, \quad (5.40)
\end{aligned}$$

where from (5.38) to (5.39), we used the fact that $\mathbf{H}^{(2, stc)}$ and $\mathbf{x}^{(2, stc)}$ are independent.

Define

$$\begin{aligned}
\gamma^{(stc)} &= \sqrt{\alpha_2^2 E_o} \\
&= \sqrt{\beta \left(\sqrt{d_b^2 + (d_e/k)^2} \right) E_o}. \quad (5.41)
\end{aligned}$$

Then the mean of $\mathbf{v}^{(stc)}$ is

$$\mathbb{E} \left[\mathbf{v}^{(stc)} \right] = \mathbf{0} \quad (5.42)$$

and the covariance matrix of $\mathbf{v}^{(stc)}$ is

$$C_{\mathbf{v}^{(stc)}, \mathbf{v}^{(stc)}} = (\gamma^{(stc)})^2 I_r. \quad (5.43)$$

The channel model can now be modeled as

$$\mathbf{y}^{(stc)} = \mathbf{H}^{(stc)} \mathbf{x}^{(stc)} + \mathbf{v}^{(stc)} + \mathbf{n}^{(stc)}, \quad (5.44)$$

where $\mathbf{H}^{(stc)} = \mathbf{H}^{(1, stc)}$, $\mathbf{x}^{(stc)} = \mathbf{x}^{(1, stc)}$ and $\mathbf{n}^{(stc)}$ is a proper complex Gaussian vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$. The computation of exact joint probability

density function for $\mathbf{v}^{(stc)}$ involves finding the probability density function of the inner product of two proper complex Gaussian random vectors and is difficult to compute. From the information theoretic point of view, the worst case of channel (5.44) is when $\mathbf{v}^{(stc)}$ is a proper complex Gaussian random vector. We assume the worst case for the rest of the analysis. Define

$$\mathbf{w}^{(stc)} = \mathbf{v}^{(stc)} + \mathbf{n}^{(stc)} \quad (5.45)$$

and let

$$(\zeta^{(stc)})^2 = (\gamma^{(stc)})^2 + \sigma^2. \quad (5.46)$$

Then according to Lemma B.2.2, we know that $\mathbf{w}^{(stc)}$ is a proper complex Gaussian random vector with

$$\mathbb{E} [\mathbf{w}^{(stc)}] = \mathbf{0}. \quad (5.47)$$

and covariance matrix

$$C_{\mathbf{w}^{(stc)}, \mathbf{w}^{(stc)}} = (\zeta^{(stc)})^2 I_r. \quad (5.48)$$

Thus the channel model for space-time coding can be rewritten as

$$\mathbf{y}^{(stc)} = \mathbf{H}^{(stc)} \mathbf{x}^{(stc)} + \mathbf{w}^{(stc)} \quad (5.49)$$

We can now apply the capacity result in (C.20) to obtain the capacity of channel given in (5.49) to get

$$C^{(stc)}(d_e/k, d_b) = \int_0^\infty \log_2 \left(1 + \frac{s(\alpha_1)^2 E_o}{(\zeta^{(stc)})^2 t} \right) \sum_{j=0}^{N_{min}-1} \frac{j!}{(j + N_{max} - N_{min})!} \cdot (L_j^{N_{max}-N_{min}}(s))^2 s^{N_{max}-N_{min}} e^{-s} ds \text{ bits}/(\text{channel use}). \quad (5.50)$$

Since we use the capacity achieving transmitter and receiver, the transport efficiency of linear network A when space-time coding is used is given by

$$\mu^{(stc)}(k, P_{max}, C^{(stc)}(d_e/k, d_b), d_e) = \left(\frac{1}{d_e} \right) (C^{(stc)}(d_e/k, d_b))^2 \left(\frac{d_e/k}{E_c + E_p} \right). \quad (5.51)$$

As we did in Chapter IV, we can replace d_e/k by d_a . Let

$$g^{(stc)}(d_a, d_b) = d_a (C^{(stc)}(d_a, d_b))^2. \quad (5.52)$$

Then

$$\mu^{(stc)}(k, P_{max}, C^{(stc)}(d_e/k, d_b), d_e) = \frac{g^{(stc)}(d_a, d_b)}{d_e(E_c + E_p)}. \quad (5.53)$$

We will use $g^{(stc)}$ as a performance measure and compare it with the performance of other transmission schemes. Note that the unit of $g^{(stc)}$ is $\text{bit}^2 \cdot \text{m}/(\text{channel use})$. For each d_b , we optimize $g^{(stc)}$ over d_a .

5.4.3 Transport Efficiency for the Ordinary End-Fire Antenna Array

In Section 5.3, we saw how the ordinary end-fire antenna array works when we have a deterministic channel. We also derived fading channel when the radiating elements are placed close together and there are many local scatters around each of the receiving antennae. From (5.14), we have a channel matrix

$$\mathbf{H} = \alpha |(AF)| \mathbf{z}(\theta), \quad (5.54)$$

where $\mathbf{z}(\theta)$ is a proper complex Gaussian random variable with mean 0 and variance 1. For the ordinary end-fire antenna array, the channel model given in (5.27) can be specified. In particular, we have

$$\mathbf{H}^{(1,aa)} = \alpha_1 |(AF)|_{\theta=0^\circ} \mathbf{z}(\theta)|_{\theta=0^\circ} = \alpha_1 t \mathbf{z}_1 \quad (5.55)$$

and

$$\mathbf{H}^{(2,aa)} = \alpha_2 |(AF)|_{\theta=\theta_2} \mathbf{z}_2(\theta)|_{\theta=\theta_2} = \alpha_2 |(AF)|_{\theta=\theta_2} \mathbf{z}_2, \quad (5.56)$$

where \mathbf{z}_1 and \mathbf{z}_2 are independent proper complex Gaussian random variables with mean 0 and variance 1. From Figure 5.3, we see that

$$\cos(\theta_2) = \frac{d_e/k}{\sqrt{d_b^2 + (d_e/k)^2}}, \quad (5.57)$$

where $0 \leq \theta_2 \leq \pi/2$. We also assume that all the receiving elements are placed close to each other and they receive identical signals. The j -th receiving antenna receives signal

$$\begin{aligned} \mathbf{y}_j &= \mathbf{H}^{(1,aa)} \mathbf{x}^{(1,aa)} + \mathbf{H}^{(2,aa)} \mathbf{x}^{(2,aa)} + \mathbf{n}_j \\ &= \alpha_1 t \mathbf{z}_1 \mathbf{x}^{(1,aa)} + \alpha_2 |(AF)|_{\theta=\theta_2} |\mathbf{z}_2 \mathbf{x}^{(2,aa)} + \mathbf{n}_j \quad \text{for } j = 1, \dots, r, \end{aligned} \quad (5.58)$$

where $\mathbf{x}^{(1,aa)}$ is the input random variable to the ordinary end-fire antenna array at A_1 , $\mathbf{x}^{(2,aa)}$ is the input random variable to the ordinary end-fire antenna array at A_2 , and \mathbf{n}_j , being a proper complex Gaussian random variable with mean 0 and variance σ^2 , is the additive noise at the j -th receiving antenna. We assume equal gain combining at the receiver. Let $\mathbf{y}^{(aa)}$ be the received signal after combining. Then

$$\begin{aligned} \mathbf{y}^{(aa)} &= \sum_{j=1}^r \mathbf{y}_j \\ &= r \mathbf{H}^{(1,aa)} \mathbf{x}^{(1,aa)} + r \mathbf{H}^{(2,aa)} \mathbf{x}^{(2,aa)} + \sum_{j=1}^r \mathbf{n}_j \\ &= \alpha_1 r t \mathbf{z}_1 \mathbf{x}^{(1,aa)} + \alpha_2 r |(AF)|_{\theta=\theta_2} |\mathbf{z}_2 \mathbf{x}^{(2,aa)} + \sum_{j=1}^r \mathbf{n}_j. \end{aligned} \quad (5.59)$$

Define

$$\mathbf{v}^{(aa)} = \alpha_2 r |(AF)|_{\theta=\theta_2} |\mathbf{z}_2 \mathbf{x}^{(2,aa)}. \quad (5.60)$$

Then

$$\begin{aligned} \mathbb{E} [\mathbf{v}^{(aa)}] &= \mathbb{E} [\alpha_2 r |(AF)|_{\theta=\theta_2} |\mathbf{z}_2 \mathbf{x}^{(2,aa)}] \\ &= r \alpha_2 |(AF)|_{\theta=\theta_2} |\mathbb{E} [\mathbf{z}_2] \mathbb{E} [\mathbf{x}^{(2,aa)}] \\ &= 0 \end{aligned} \quad (5.61)$$

and the covariance

$$C_{\mathbf{v}^{(aa)}, \mathbf{v}^{(aa)}} = \mathbb{E} \left[(\mathbf{v}^{(aa)} - \mathbb{E} [\mathbf{v}^{(aa)}]) (\mathbf{v}^{(aa)} - \mathbb{E} [\mathbf{v}^{(aa)}])^* \right]$$

$$\begin{aligned}
&= \mathbb{E} \left[\mathbf{v}^{(aa)} (\mathbf{v}^{(aa)})^* \right] \\
&= \mathbb{E} \left[\alpha_2^2 r^2 |(AF)|_{\theta=\theta_2}|^2 |\mathbf{z}_2|^2 |\mathbf{x}^{(2,aa)}|^2 \right] \\
&= \alpha_2^2 r^2 |(AF)|_{\theta=\theta_2}|^2 \mathbb{E} [|\mathbf{z}_2|^2] \mathbb{E} [|\mathbf{x}^{(2,aa)}|^2] \\
&= \alpha_2^2 r^2 ((AF)|_{\theta=\theta_2})^2 \frac{E_o}{t}. \tag{5.62}
\end{aligned}$$

Define

$$\begin{aligned}
\gamma^{(aa)} &= \sqrt{\frac{\alpha_2^2 r^2 ((AF)|_{\theta=\theta_2})^2 E_o}{t}} \\
&= \sqrt{\frac{\beta \left(\sqrt{d_b^2 + (d_e/k)^2} \right) r^2 ((AF)|_{\theta=\theta_2})^2 E_o}{t}}. \tag{5.63}
\end{aligned}$$

Then the mean of $\mathbf{v}^{(aa)}$ is

$$\mathbb{E} [\mathbf{v}^{(aa)}] = \mathbf{0} \tag{5.64}$$

and the covariance of $\mathbf{v}^{(aa)}$ is

$$C_{\mathbf{v}^{(aa)}, \mathbf{v}^{(aa)}} = (\gamma^{(aa)})^2. \tag{5.65}$$

The computation of exact joint probability density function for $\mathbf{v}^{(aa)}$ involves finding the probability density function of the product of two proper complex Gaussian random variables and is difficult to compute. From the information theoretic point of view, the worst case of channel (5.59) is when $\mathbf{v}^{(aa)}$ is a proper complex Gaussian random variable. We assume the worst case for the rest of the analysis. Define

$$\mathbf{n}^{(aa)} = \sum_{j=1}^r \mathbf{n}_j. \tag{5.66}$$

Then

$$\begin{aligned}
\mathbb{E} [\mathbf{n}^{(aa)}] &= \mathbb{E} \left[\sum_{j=1}^r \mathbf{n}_j \right] \\
&= \sum_{j=1}^r \mathbb{E} [\mathbf{n}_j] \\
&= 0 \tag{5.67}
\end{aligned}$$

and

$$\begin{aligned}
C_{\mathbf{n}^{(aa)}, \mathbf{n}^{(aa)}} &= \mathbb{E} \left[\left(\sum_{j=1}^r \mathbf{n}_j \right) \left(\sum_{j=1}^r \mathbf{n}_j \right)^* \right] \\
&= \sum_{j=1}^r \mathbb{E} [\mathbf{n}_j \mathbf{n}_j^*] \\
&= r\sigma^2.
\end{aligned} \tag{5.68}$$

Define

$$\alpha = \alpha_1 r t. \tag{5.69}$$

Thus the channel model for the ordinary end-fire antenna array can be written as

$$\mathbf{y} = \mathbf{H}^{(aa)} \mathbf{x}^{(aa)} + \mathbf{v}^{(aa)} + \mathbf{n}^{(aa)} \tag{5.70}$$

where

$$\mathbf{H}^{(aa)} = \alpha \mathbf{z}_1 \tag{5.71}$$

and $\mathbf{x}^{(aa)} = \mathbf{x}^{(1,aa)}$. The transmitted energy from the desired transmitting antennae is E_o . Therefore

$$\mathbb{E} \left[\mathbf{x}^{(aa)} (\mathbf{x}^{(aa)})^* \right] = \frac{E_o}{t}. \tag{5.72}$$

Define

$$\mathbf{w}^{(aa)} = \mathbf{v}^{(aa)} + \mathbf{n}^{(aa)} \tag{5.73}$$

and let

$$(\zeta^{(aa)})^2 = (\gamma^{(aa)})^2 + \sigma^2. \tag{5.74}$$

Then according to Lemma B.2.2, we know that $\mathbf{w}^{(aa)}$ is a proper complex Gaussian variable with mean 0 and variance $(\zeta^{(aa)})^2$. Therefore we can write the interference channel model for the ordinary end-fire antenna array as

$$\mathbf{y}^{(aa)} = \mathbf{H}^{(aa)} \mathbf{x}^{(aa)} + \mathbf{w}^{(aa)}. \tag{5.75}$$

We can now apply the capacity result in Theorem C.3.1 and get the capacity of the channel given in (5.75)

$$C^{(aa)}(d_e/k, d_b) = \int_0^\infty \log_2 \left(1 + s\alpha^2 E_o / (\zeta^{(aa)})^2 \right) e^{-s} ds \quad \text{bits}/(\text{channel use}). \quad (5.76)$$

The transport efficiency for the antenna array is given by

$$\mu^{(aa)}(k, P_{max}, C^{(aa)}(d_e/k, d_b), d_e) = \left(\frac{1}{d_e} \right) (C^{(aa)}(d_e/k, d_b))^2 \frac{d_e/k}{E_c + E_p}. \quad (5.77)$$

As we did in Chapter IV, we can replace d/k by d_a . Let

$$g^{(aa)}(d_a, d_b) = d_a (C^{(aa)}(d_a, d_b))^2. \quad (5.78)$$

Then

$$\mu^{(aa)}(k, P_{max}, C^{(aa)}(d_e/k, d_b), d_e) = \frac{g^{(aa)}(d_a, d_b)}{d_e(E_c + E_p)}. \quad (5.79)$$

We will use $g^{(aa)}$ as a performance measure and compare it with the performance of other transmission schemes. Note that the unit of $g^{(aa)}$ is $\text{bit}^2 \cdot \text{m}/(\text{channel use})$. For each d_b , we optimize $g^{(aa)}$ over d_a .

5.5 Numerical Results

We compute (5.52) and (5.78) based on the above derivation of the transport efficiency in order to understand in what situation space-time coding is better than the ordinary end-fire antenna array, or the other way. For the numerical calculation, we assume the numerical values for the parameters given in Section 4.4. For those values, we have $E_o = 1.5 \times 10^{-6}$ J and $E_c = 2.2 \times 10^{-6}$ J. We set $\sigma^2 = \frac{N_0}{2}$. We fix the number of receiving antennae, but vary the number of transmitting antennae. Doing so would allow us to see how many receiving antennae are needed for space-time coding to be better than the ordinary end-fire antenna array. The optimal g and

optimal adjacent distance for space-time coding and the ordinary end-fire antenna array are shown in Figures 5.4 to 5.7 and Figures 5.8 to 5.11, respectively, with solid lines for the performance of space-time coding and dashed lines for the performance of the ordinary end-fire antenna array.

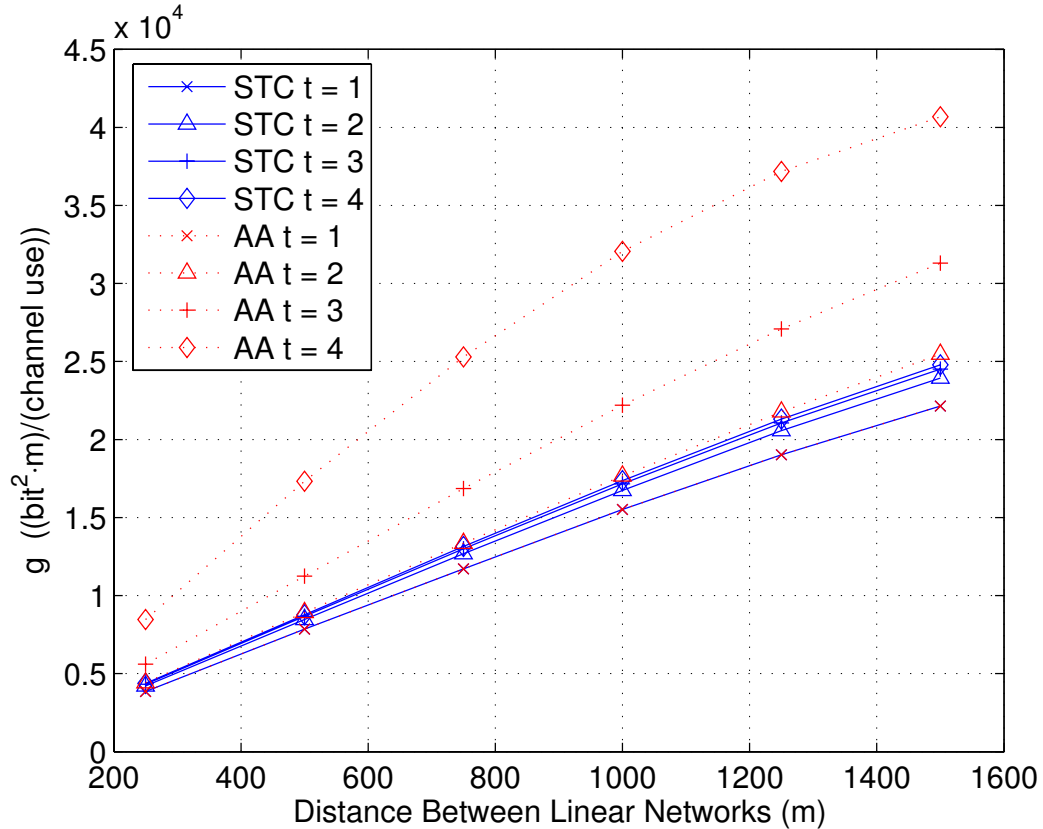


Figure 5.4: Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 1$ receiving antenna is used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

We see from Figures 5.4 to 5.7 that when $r = 1$, the ordinary end-fire antenna array always provides better transport efficiency than space-time coding. This is due to the fact that the ordinary end-fire antenna array has an amplitude pattern that causes less interference to the neighboring network and provides high signal-to-noise

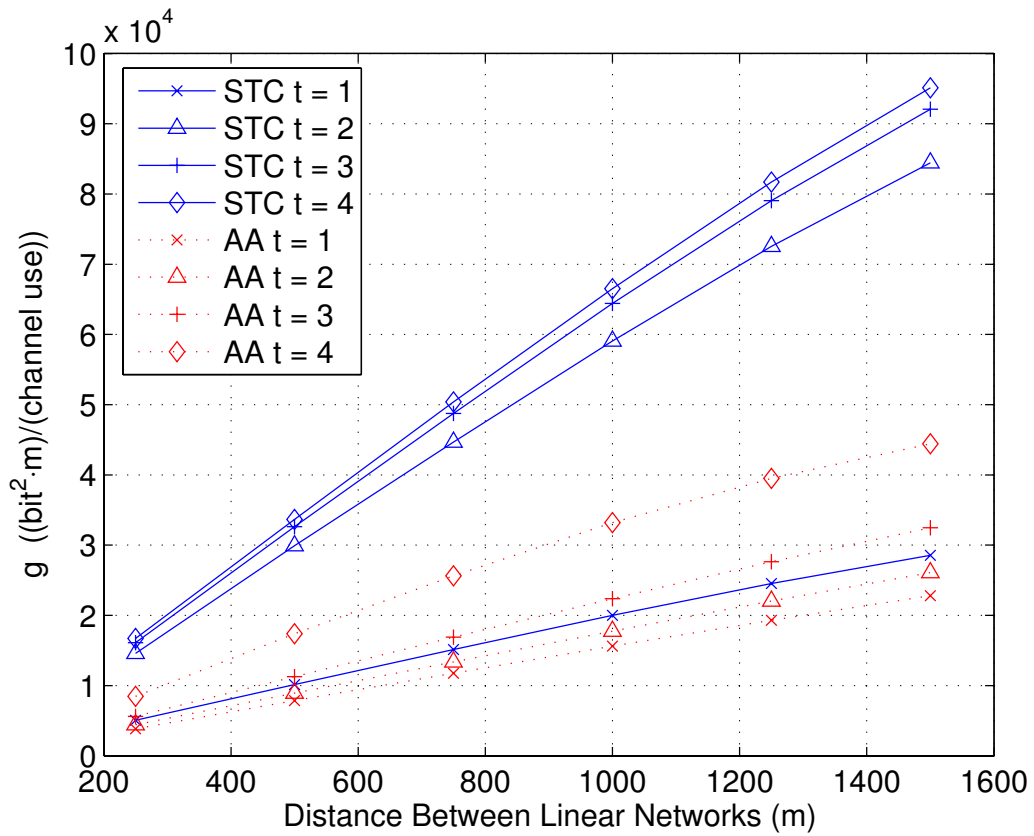


Figure 5.5: Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 2$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

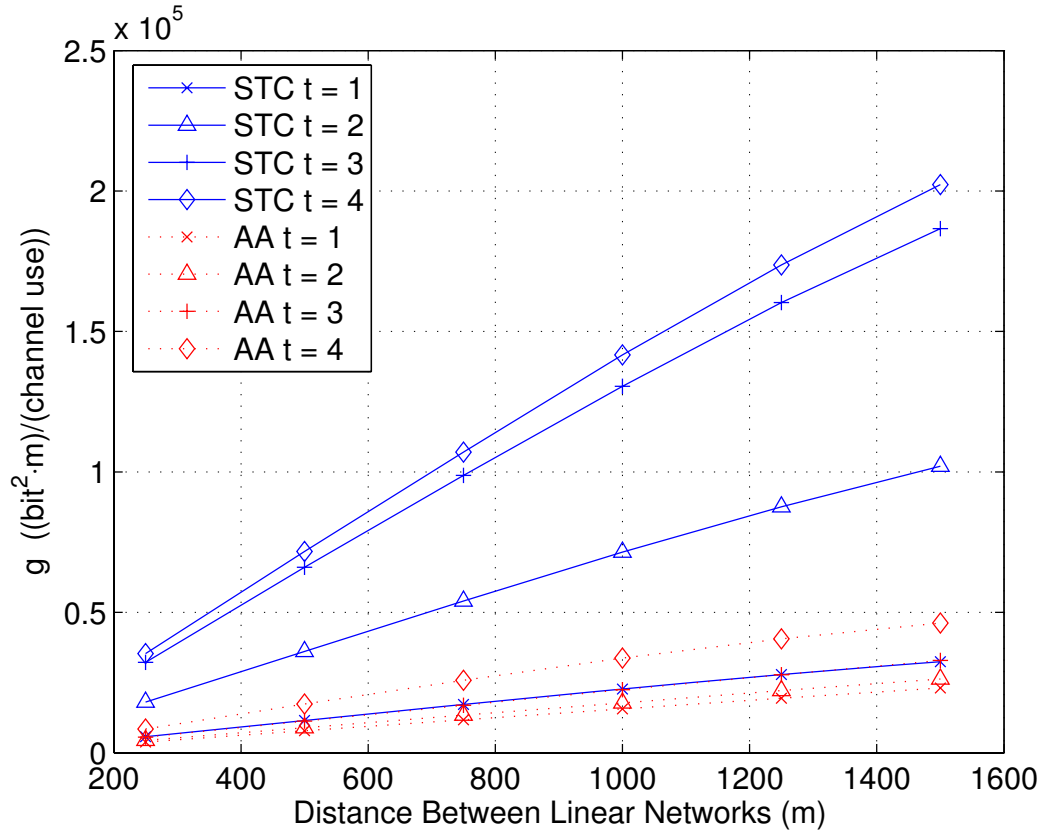


Figure 5.6: Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 3$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

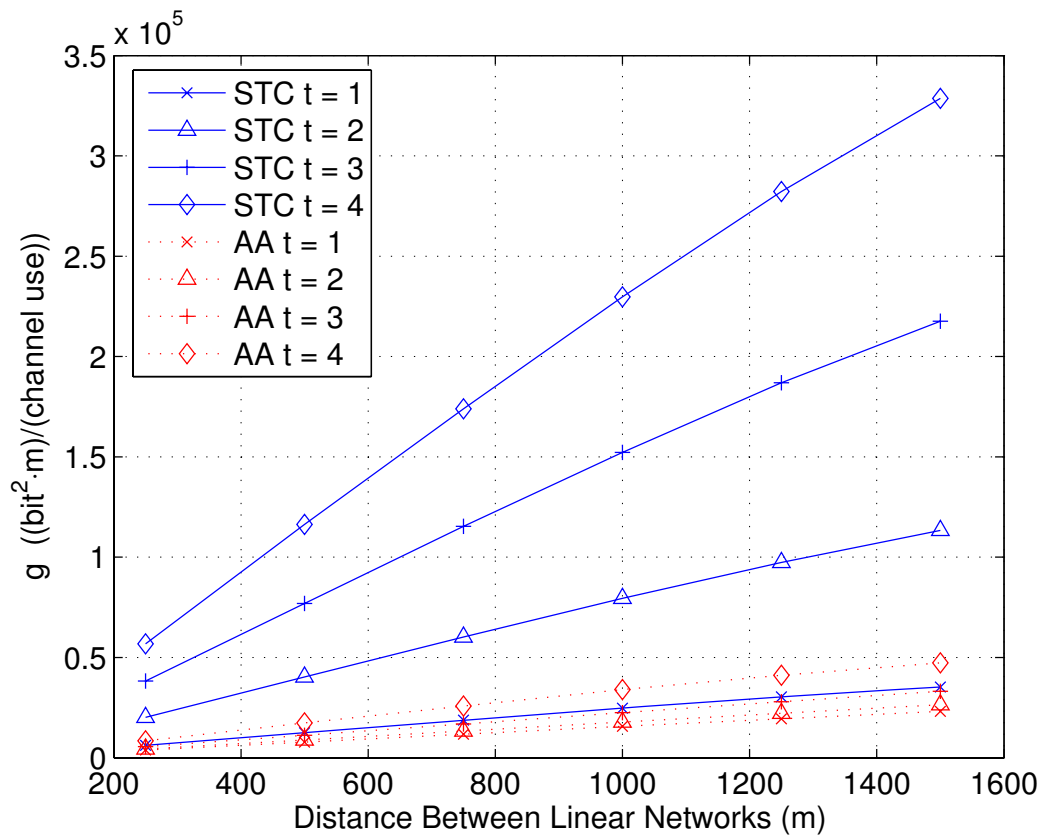


Figure 5.7: Optimal $g^{(stc)}$ and $g^{(aa)}$ with respect to the distance between linear networks when $r = 4$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

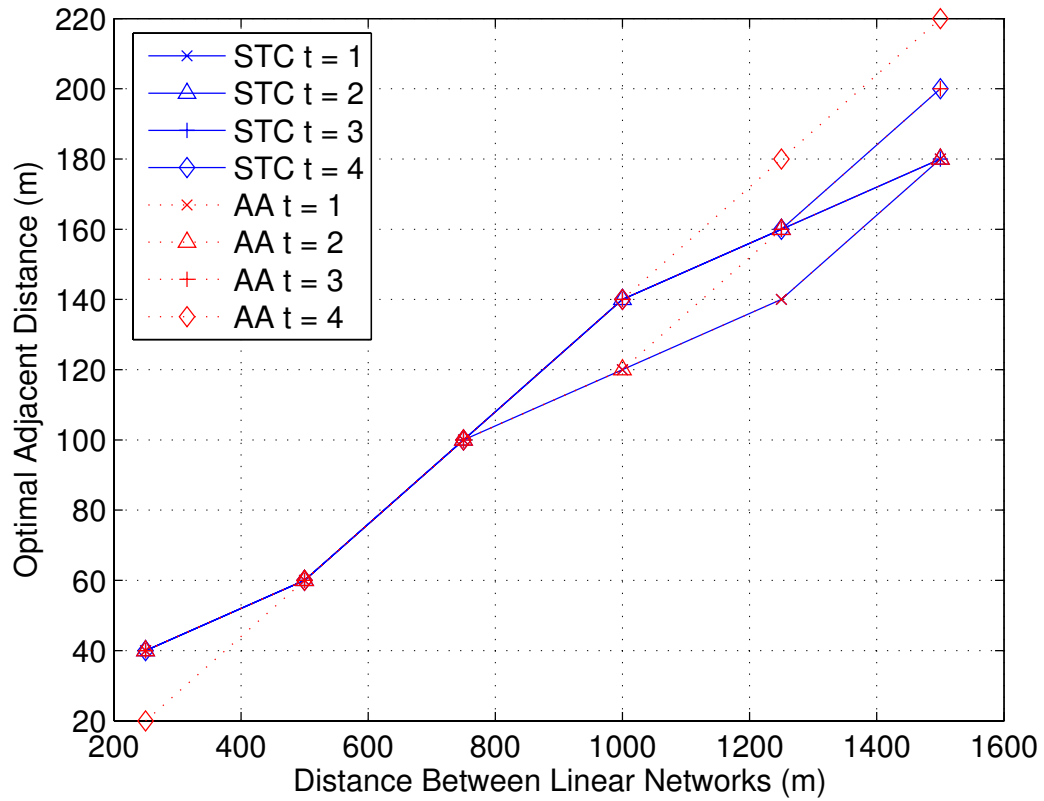


Figure 5.8: Optimal adjacent distance with respect to the distance between linear networks when $r = 1$ receiving antenna is used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

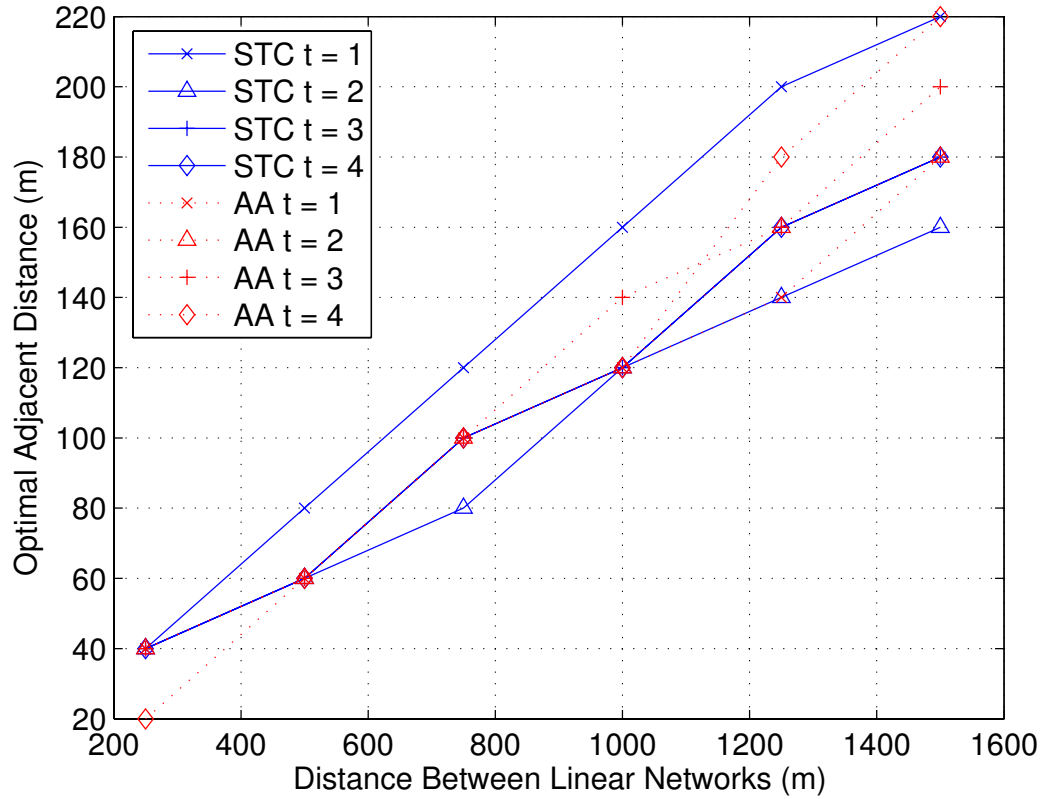


Figure 5.9: Optimal adjacent distance with respect to the distance between linear networks when $r = 2$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

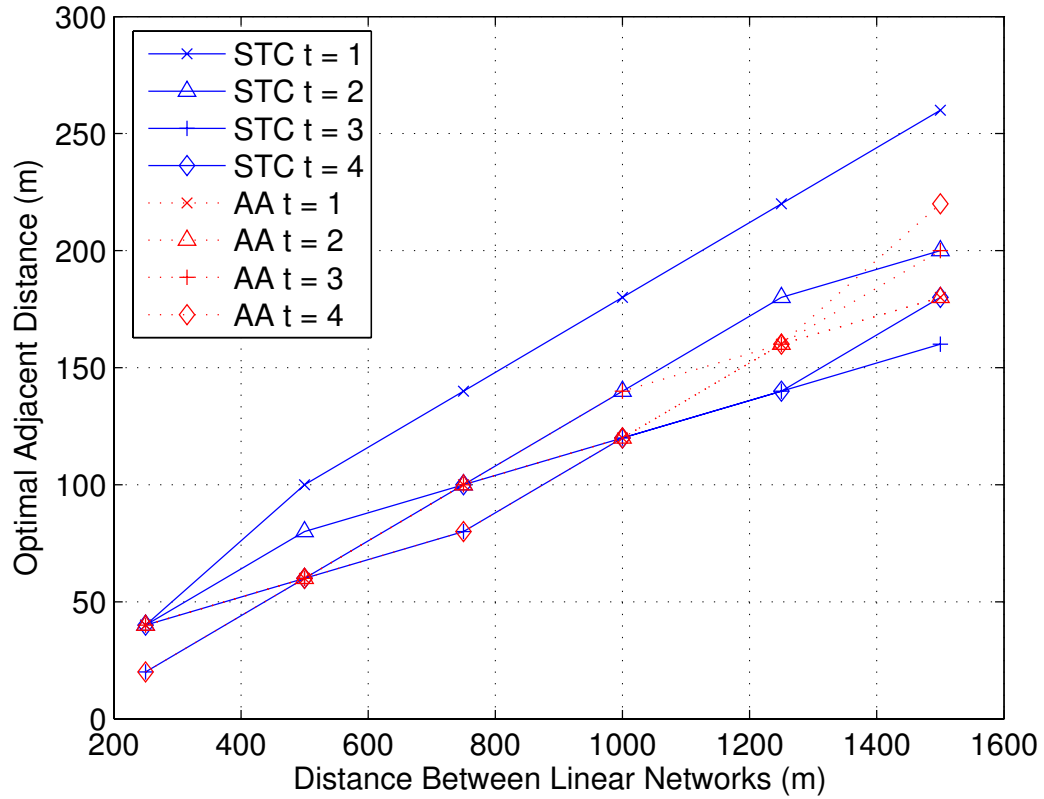


Figure 5.10: Optimal adjacent distance with respect to the distance between linear networks when $r = 3$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

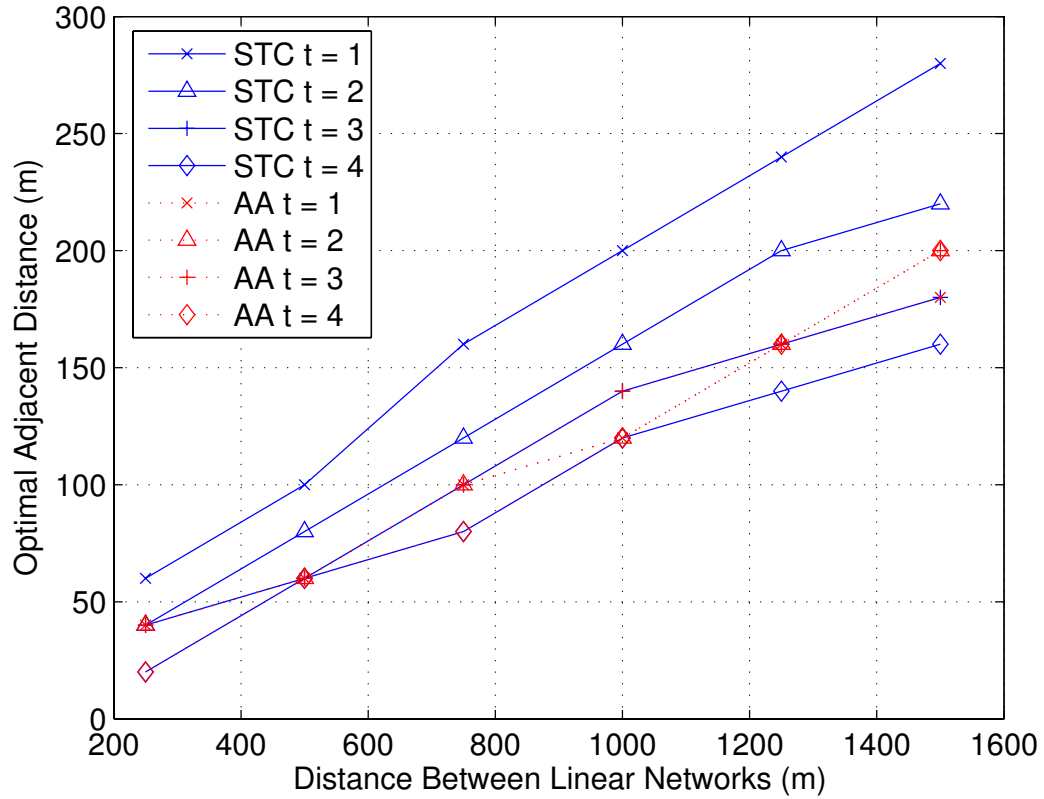


Figure 5.11: Optimal adjacent distance with respect to the distance between linear networks when $r = 4$ receiving antennae are used. Solid lines show the performance of space-time coding and dashed lines show the performance of the ordinary end-fire antenna array.

ratio in the desired direction. When $r \geq 2$, the potential of high channel capacity of space-time coding provides means to achieve higher transport efficiency. The drawback of space-time coding is its complexity, where extensive signal processing may be required at the receiver.

5.6 Conclusion

In this chapter we showed how space-time coding and the ordinary end-fire antenna array affect the transport efficiency of interfering networks. When the number of receiving antennae is small and when the distance between interfering networks is small, the antenna array provides higher transport efficiency. On the other hand, when the number of receiving antennae is large and the distance between interfering networks is large, space-time coding provides higher transport efficiency.

CHAPTER VI

DISTRIBUTED SPACE-TIME CODING AND COOPERATIVE COMMUNICATION

In Chapter V we investigated the transport efficiency of linear networks when nodes on different linear networks communicate without any cooperation, thus causing interference to the communications in nearby networks. In this chapter we will investigate a scenario where cooperative communication can occur. We use transport efficiency as a performance measure and show how cooperative communications can improve the transport efficiency when nodes use space-time coding.

6.1 Introduction

In wireless communication networks, an important question to ask is how much throughput a network can deliver from sources to destinations per unit bandwidth, per unit energy. With the implementation of multiple transmitting antennae and receiving antennae, it is possible to realize some form of cooperative communication to improve the performance of a wireless network.

Cooperative transmission for wireless networks is studied in [48]. The authors investigate a scenario where a source communicates with a destination via at most two hops. The destination has one receiving antenna and it receives signals from both

the source node and the relay nodes. If a relay node can successfully decode the received signal, it can either repeat the decoded signal to perform a repetition based cooperative communication or work with other relay nodes that have also successfully decoded the received signal to perform a space-time coding based cooperative communication. The authors use outage probability as a performance measure for both cooperative communication schemes and show that space-time coding based scheme achieves higher spectral efficiencies than repetition-based scheme.

We will measure the effectiveness of cooperative communication by the transport efficiency introduced in Chapter IV. Before we proceed to look into the cooperative communication of linear networks, we derive some capacity results for distributed space-time coding. We will use some basic results of space-time coding in Appendix C.

The remainder of this chapter is organized as follows. In Section 6.2 we derive the capacity for distributed space-time coding with individual energy constraint and covariance feedback. In Section 6.3 we calculate the capacity for distributed space-time coding when there is only one receiving antenna. In Section 6.4 we investigate the transport efficiency of cooperative communication for linear networks. We show numerical results and draw conclusions in Section 6.5.

6.2 Capacity for Distributed Space-Time Coding with Individual Constraint and Covariance Feedback

We are going to derive capacity for a channel where there are m geometrically distributed transmitters trying to communicate with a single receiver. The i -th transmitter has t_i number of transmitting antennae and the receiver has r receiving antennae. Let $\mathbf{x}^{(i)} \in \mathbb{C}^{t_i}$ be the transmitted vector from the i -th transmitter for

$i = 1, \dots, m$. We don't assume that $\mathbf{x}^{(i)} \in \mathbb{C}^{t_i}$, $i = 1, \dots, m$ are independent. Let $t = t_1 + \dots + t_m$ and

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \\ \vdots \\ \mathbf{x}^{(m)} \end{bmatrix}. \quad (6.1)$$

Then $\mathbf{x} \in \mathbb{C}^t$ is the transmitted vector from the m transmitters. Let $\mathbf{n} \in \mathbb{C}^r$ be the additive noise at the receiving antennae. We assume that $\mathbf{n} \in \mathbb{C}^r$ is a proper [58] complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$. Let $\mathbf{y} \in \mathbb{C}^r$ be the received signal at the receiving antennae. Let $\mathbf{H}^{(i)} \in \mathbb{C}^{r \times t_i}$, $i = 1, \dots, m$ be the channel matrix from the transmitter i to the receiver. We assume that $\mathbf{H}^{(i)}$, $i = 1, \dots, m$ are mutually independent. Furthermore, for each i , the row vectors of $\mathbf{H}^{(i)}$ are independent and identically distributed proper complex Gaussian random vectors with mean $\mathbf{0}$ and covariance matrix $K^{(i)}$. If row vectors of $\mathbf{H}^{(i)}$ are not proper, we have to consider the real part and imaginary part of the row vectors and reformulate the problem using real numbers. Define

$$\mathbf{H} = [\mathbf{H}^{(1)} \ \mathbf{H}^{(2)} \ \dots \ \mathbf{H}^{(m)}]. \quad (6.2)$$

Then $\mathbf{H} \in \mathbb{C}^{r \times t}$ and row vectors of \mathbf{H} are independent and identically distributed proper complex Gaussian random vectors with mean $\mathbf{0}$ and covariance matrix

$$K = \begin{bmatrix} K^{(1)} & & \\ & \ddots & \\ & & K^{(m)} \end{bmatrix}. \quad (6.3)$$

With \mathbf{x} , \mathbf{y} , \mathbf{n} , and \mathbf{H} defined above, we consider a frequency non-selective slowly fading channel model [65]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (6.4)$$

where each channel use by each transmitter is energy limited. More precisely,

$$\mathbb{E} \left[(\mathbf{x}^{(i)})^\dagger \mathbf{x}^{(i)} \right] \leq E_i \quad \text{for } i = 1, \dots, m, \quad (6.5)$$

or equivalently,

$$\text{tr} \left(\mathbb{E} \left[\mathbf{x}^{(i)} (\mathbf{x}^{(i)})^\dagger \right] \right) \leq E_i \quad \text{for } i = 1, \dots, m. \quad (6.6)$$

Let

$$Q^{(i,j)} = \mathbb{E} \left[\mathbf{x}^{(i)} (\mathbf{x}^{(j)})^\dagger \right] \quad \text{for } i = 1, \dots, m, \quad j = 1, \dots, m. \quad (6.7)$$

The constraint in (6.6) is equivalent to

$$\text{tr} (Q^{(i,i)}) \leq E_i \quad \text{for } i = 1, \dots, m. \quad (6.8)$$

Define

$$\begin{aligned} \mathcal{Q}^{(i)} &= \{ Q^{(i,i)} : Q^{(i,i)} \text{ is specified in (6.7) and satisfies (6.8)} \} \\ &\quad \text{for } i = 1, \dots, m. \end{aligned} \quad (6.9)$$

Let Q be the correlation matrix of the transmitted vector \mathbf{x} , then

$$Q = \begin{bmatrix} Q^{(1,1)} & \dots & Q^{(1,m)} \\ \vdots & \ddots & \vdots \\ Q^{(m,1)} & \dots & Q^{(m,m)} \end{bmatrix}. \quad (6.10)$$

Define

$$\mathcal{Q} = \{ Q : Q \text{ is specified in (6.10) and } Q^{(i,i)} \in \mathcal{Q}^{(i)} \text{ for } i = 1, \dots, m \}. \quad (6.11)$$

Recall that $K^{(i)}$ is the covariance matrix of any row vector of $\mathbf{H}^{(i)}$. Then $K^{(i)}$ is a positive semidefinite Hermitian matrix. Therefore there exists a unitary matrix $U^{(i)} \in \mathbb{C}^{t_i \times t_i}$ such that

$$(U^{(i)})^\dagger K^{(i)} U^{(i)} = \Lambda^{(i)}, \quad (6.12)$$

where

$$\Lambda^{(i)} = \begin{bmatrix} \lambda_{1,1}^{(i)} & & \\ & \ddots & \\ & & \lambda_{t_i,t_i}^{(i)} \end{bmatrix} \quad (6.13)$$

is a real positive semidefinite diagonal matrix and entries on the diagonal are arranged in descending order, i.e., $\lambda_{1,1}^{(i)} \geq \dots \geq \lambda_{t_i,t_i}^{(i)} \geq 0$. Let

$$U = \begin{bmatrix} U^{(1)} & & \\ & \ddots & \\ & & U^{(m)} \end{bmatrix} \quad (6.14)$$

and

$$\Lambda = \begin{bmatrix} \Lambda^{(1)} & & \\ & \ddots & \\ & & \Lambda^{(m)} \end{bmatrix}, \quad (6.15)$$

then we have

$$\Lambda = U^\dagger K U. \quad (6.16)$$

Let

$$D^{(i)} = \begin{bmatrix} d_{1,1}^{(i)} & & \\ & \ddots & \\ & & d_{t_i,t_i}^{(i)} \end{bmatrix} \quad (6.17)$$

be a real positive semidefinite diagonal matrix and entries on the diagonal are arranged in descending order, i.e., $d_{1,1}^{(i)} \geq \dots \geq d_{t_i,t_i}^{(i)} \geq 0$. Define

$$\mathcal{D}^{(i)} = \left\{ D^{(i)} : D^{(i)} \text{ is specified in (6.17) and } \text{tr} \left((\Lambda^{(i)})^{-\frac{1}{2}} D^{(i)} (\Lambda^{(i)})^{-\frac{1}{2}} \right) \leq E_i \right\} \\ \text{for } i = 1, \dots, m. \quad (6.18)$$

Let

$$D = \begin{bmatrix} D^{(1)} & & \\ & \ddots & \\ & & D^{(m)} \end{bmatrix}, \quad (6.19)$$

where $D^{(i)} \in \mathcal{D}^{(i)}$. Define

$$\mathcal{D} = \{D : D \text{ is specified in (6.19)}\}. \quad (6.20)$$

Theorem 6.2.1 gives the capacity for the channel model given by (6.4).

Theorem 6.2.1 (Capacity of Channel with Covariance Feedback) *Let the realization of \mathbf{H} be known to the receiver and $K^{(i)}$ be known to the transmitter i for $i = 1, \dots, m$. The capacity of distributed space-time coding for the channel model given in (6.4) is*

$$C = \max_{D^{(i)} \in \mathcal{D}^{(i)}, i=1, \dots, m} \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \frac{\mathbf{Z}^{(i)} D^{(i)} (\mathbf{Z}^{(i)})^\dagger}{\sigma^2} \right) \right) \right], \quad (6.21)$$

where for each i , all entries of $\mathbf{Z}^{(i)} \in \mathbb{C}^{r \times t_i}$ are independent and identically distributed proper complex Gaussian random variables with mean 0 and variance 1, and $\mathbf{Z}^{(i)} \in \mathbb{C}^{r \times t_i}$, $i = 1, \dots, m$ are independent.

Proof of Theorem 6.2.1:

We know that the capacity for channel (6.4) when \mathbf{H} is known to the receiver is given by [80, 88]

$$C = \max_{Q \in \mathcal{Q}} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{H} Q \mathbf{H}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.22)$$

For each fixed $Q \in \mathcal{Q}$, the mutual information $\mathcal{I}(\mathbf{x}; \mathbf{y})$ is maximized and

$$\mathcal{I}(\mathbf{x}; \mathbf{y}) = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{H} Q \mathbf{H}^\dagger}{\sigma^2} \right) \right) \right] \quad (6.23)$$

when the input vector \mathbf{x} is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Q . Define

$$\mathbf{Z} = \mathbf{H}U\Lambda^{-\frac{1}{2}}. \quad (6.24)$$

Then by Lemma B.2.2, we have that the row vectors of \mathbf{Z} are independent and identically distributed proper complex Gaussian random vectors with mean $\mathbf{0}$ and covariance matrix I_t . From (6.24), we have

$$\mathbf{H} = \mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger. \quad (6.25)$$

We can then rewrite (6.22) as

$$C = \max_{Q \in \mathcal{Q}} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.26)$$

We first partition Q into four blocks and have

$$Q = \begin{bmatrix} Q^{(1,1)} & A \\ A^\dagger & B \end{bmatrix}, \quad (6.27)$$

where

$$A = \begin{bmatrix} Q^{(1,2)} & \dots & Q^{(1,m)} \end{bmatrix} \quad (6.28)$$

and

$$B = \begin{bmatrix} Q^{(2,2)} & \dots & Q^{(2,m)} \\ \vdots & \ddots & \vdots \\ Q^{(m,2)} & \dots & Q^{(m,m)} \end{bmatrix}. \quad (6.29)$$

Define another matrix \tilde{Q} by

$$\tilde{Q} = \begin{bmatrix} Q^{(1,1)} & -A \\ -A^\dagger & B \end{bmatrix}. \quad (6.30)$$

Since Q is a positive semidefinite matrix, the matrix \tilde{Q} is also a positive semidefinite matrix by the observation that

$$\tilde{Q} = \begin{bmatrix} Q^{(1,1)} & -A \\ -A^\dagger & B \end{bmatrix} = \begin{bmatrix} I_{t_1} & 0 \\ 0 & -I_{t_2+\dots+t_m} \end{bmatrix} \begin{bmatrix} Q^{(1,1)} & A \\ A^\dagger & B \end{bmatrix} \begin{bmatrix} I_{t_1} & 0 \\ 0 & -I_{t_2+\dots+t_m} \end{bmatrix}. \quad (6.31)$$

We now show that

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger\tilde{Q}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.32)$$

Let

$$\mathbf{Z}^{(i)} = \mathbf{H}^{(i)}U^{(i)} (\Lambda^{(i)})^{-\frac{1}{2}}. \quad (6.33)$$

Then all entries of $\mathbf{Z}^{(i)}$ are independent and identically distributed proper complex Gaussian random variables with mean 0 and variance 1. From (6.24), we have

$$\mathbf{Z} = [\mathbf{Z}^{(1)} \ \mathbf{Z}^{(2)} \ \dots \ \mathbf{Z}^{(m)}] \quad (6.34)$$

Define $\mathbf{V}^{(i)} \in \mathbb{C}^{r \times t_i}$ for $i = 1, \dots, m$ as follows

$$\mathbf{V}^{(1)} = \mathbf{Z}^{(1)}; \quad (6.35)$$

$$\mathbf{V}^{(i)} = -\mathbf{Z}^{(i)} \quad \text{for } i = 2, \dots, m. \quad (6.36)$$

Let

$$\mathbf{V} = [\mathbf{V}^{(1)} \ \mathbf{V}^{(2)} \ \dots \ \mathbf{V}^{(m)}]. \quad (6.37)$$

Then \mathbf{V} has the same distribution as \mathbf{Z} and consequently

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{V}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{V}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.38)$$

Let

$$\begin{aligned}\mathbf{Z}_B &= [\mathbf{Z}^{(2)} \dots \mathbf{Z}^{(m)}]; \\ \mathbf{V}_B &= [\mathbf{V}^{(2)} \dots \mathbf{V}^{(m)}].\end{aligned}\tag{6.39}$$

Then

$$\mathbf{V}_B = -\mathbf{Z}_B.\tag{6.40}$$

Define

$$U_B = \begin{bmatrix} U^{(2)} & & \\ & \ddots & \\ & & U^{(m)} \end{bmatrix}\tag{6.41}$$

and

$$\Lambda_B = \begin{bmatrix} \Lambda^{(2)} & & \\ & \ddots & \\ & & \Lambda^{(m)} \end{bmatrix}.\tag{6.42}$$

We have that

$$\begin{aligned}& \mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger\tilde{Q}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger \\ &= \mathbf{Z}^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(U^{(1)})^\dagger Q^{(1,1)}U^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(\mathbf{Z}^{(1)})^\dagger - \\ & \quad \mathbf{Z}_B\Lambda_B^{\frac{1}{2}}(U_B)^\dagger A^\dagger U^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(\mathbf{Z}^{(1)})^\dagger - \\ & \quad \mathbf{Z}^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(U^{(1)})^\dagger AU_B\Lambda_B^{\frac{1}{2}}\mathbf{Z}_B^\dagger + \\ & \quad \mathbf{Z}_B\Lambda_B^{\frac{1}{2}}U_B^\dagger BU_B\Lambda_B^{\frac{1}{2}}\mathbf{Z}_B^\dagger \\ &= \mathbf{Z}^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(U^{(1)})^\dagger Q^{(1,1)}U^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(\mathbf{Z}^{(1)})^\dagger + \\ & \quad \mathbf{V}_B\Lambda_B^{\frac{1}{2}}(U_B)^\dagger A^\dagger U^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(\mathbf{Z}^{(1)})^\dagger + \\ & \quad \mathbf{Z}^{(1)}(\Lambda^{(1)})^{\frac{1}{2}}(U^{(1)})^\dagger AU_B\Lambda_B^{\frac{1}{2}}\mathbf{V}_B^\dagger + \\ & \quad \mathbf{V}_B\Lambda_B^{\frac{1}{2}}U_B^\dagger BU_B\Lambda_B^{\frac{1}{2}}\mathbf{V}_B^\dagger \\ &= \mathbf{V}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{V}^\dagger.\end{aligned}\tag{6.43}$$

Therefore

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger\tilde{Q}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{V}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{V}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.44)$$

It follows from (6.38) and (6.44) that

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger\tilde{Q}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.45)$$

Let

$$Q^{(d,1)} = \frac{1}{2} (Q + \tilde{Q}) = \begin{bmatrix} Q^{(1,1)} & 0 \\ 0 & B \end{bmatrix}. \quad (6.46)$$

We next show that

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger Q^{(d,1)}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] \\ & \geq \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \end{aligned} \quad (6.47)$$

Since $\log(\det(\cdot))$ is a strictly convex cap function on the set of positive definite Hermitian matrices [38](7.6.7) [13], we have

$$\begin{aligned} & \log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger Q^{(d,1)}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \\ & = \log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger \frac{1}{2} (Q + \tilde{Q}) U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \\ & \geq \frac{1}{2} \log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger QU\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) + \frac{1}{2} \log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger \tilde{Q}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right). \end{aligned} \quad (6.48)$$

Therefore

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\Lambda^{\frac{1}{2}}U^\dagger Q^{(d,1)}U\Lambda^{\frac{1}{2}}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]$$

$$\begin{aligned}
&\geq \frac{1}{2} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger Q U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] + \\
&\quad \frac{1}{2} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger \tilde{Q} U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] \\
&= \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger Q U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right], \tag{6.49}
\end{aligned}$$

where the equality in (6.49) follows from (6.45). We can now write

$$\begin{aligned}
&\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger Q^{(d,1)} U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \\
&\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}^{(1)} (\Lambda^{(1)})^{\frac{1}{2}} (U^{(1)})^\dagger Q^{(1,1)} U^{(1)} (\Lambda^{(1)})^{\frac{1}{2}} (\mathbf{Z}^{(1)})^\dagger}{\sigma^2} + \right. \right. \\
&\quad \left. \left. \frac{\mathbf{Z}_B (\Lambda_B)^{\frac{1}{2}} (U_B)^\dagger B U_B (\Lambda_B)^{\frac{1}{2}} (\mathbf{Z}_B)^\dagger}{\sigma^2} \right) \right) \right]. \tag{6.50}
\end{aligned}$$

Applying the above procedure for the proof of (6.47) iteratively to B in (6.50), we have

$$\begin{aligned}
&\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger Q U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] \\
&\leq \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} \Lambda^{\frac{1}{2}} U^\dagger Q^{(d)} U \Lambda^{\frac{1}{2}} \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \tag{6.51}
\end{aligned}$$

where $Q^{(d)}$ is formed from Q in (6.10) by taking the block matrices on the diagonal of Q , i.e.,

$$Q^{(d)} = \begin{bmatrix} Q^{(1,1)} & & \\ & \ddots & \\ & & Q^{(m,m)} \end{bmatrix}. \tag{6.52}$$

We conclude that in order to find the capacity C in (6.22), we only need to consider those matrices Q that are block diagonal, i.e.,

$$Q^{(i,j)} = O^{(i,j)} \quad \text{for } i \neq j, \tag{6.53}$$

where $O^{(i,j)}$ is a $t_i \times t_j$ matrix with all entries being 0. Therefore we only need to consider the covariance matrix

$$Q = \begin{bmatrix} Q^{(1,1)} & & \\ & \ddots & \\ & & Q^{(m,m)} \end{bmatrix}. \quad (6.54)$$

This implies that in order to find the capacity C in (6.22), we may assume that the transmitted vector $\mathbf{x}^{(i)}$, $i = 1, \dots, m$ are independent. Define

$$\mathcal{R} = \left\{ R : R = \Lambda^{\frac{1}{2}} U^\dagger Q U \Lambda^{\frac{1}{2}} \text{ for } Q \in \mathcal{Q} \text{ and } Q \text{ satisfies (6.54)} \right\}. \quad (6.55)$$

We can then rewrite (6.26) as

$$C = \max_{R \in \mathcal{R}} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} R \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.56)$$

We now show that for every $R \in \mathcal{R}$, there exists a corresponding $D \in \mathcal{D}$ such that

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} R \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} D \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.57)$$

For each $R \in \mathcal{R}$, let

$$R = \begin{bmatrix} R^{(1,1)} & \dots & R^{(1,m)} \\ \vdots & \ddots & \vdots \\ R^{(m,1)} & \dots & R^{(m,m)} \end{bmatrix}, \quad (6.58)$$

where $R^{(i,j)}$ is a $t_i \times t_j$ matrix for $i = 1, \dots, m$, $j = 1, \dots, m$. Then

$$R^{(i,j)} = O^{(i,j)} \quad \text{for } i \neq j \quad (6.59)$$

and

$$R^{(i,i)} = (\Lambda^{(i,i)})^{\frac{1}{2}} (U^{(i,i)})^\dagger Q^{(i,i)} U^{(i,i)} (\Lambda^{(i,i)})^{\frac{1}{2}} \quad \text{for } i = 1, \dots, m. \quad (6.60)$$

Consequently

$$\operatorname{tr} \left((\Lambda^{(i,i)})^{-\frac{1}{2}} R^{(i,i)} (\Lambda^{(i,i)})^{-\frac{1}{2}} \right) = \operatorname{tr} \left((U^{(i,i)})^\dagger Q^{(i,i)} U^{(i,i)} \right) \quad (6.61)$$

$$= \operatorname{tr} (Q^{(i,i)}) \quad (6.62)$$

$$\leq E_i \quad \text{for } i = 1, \dots, m. \quad (6.63)$$

Since $R^{(i,i)}$ is positive semidefinite, it follows that there exists a unitary matrix $W^{(i)}$ such that

$$(W^{(i)})^\dagger R^{(i,i)} W^{(i)} = D^{(i)} \quad \text{for } i = 1, \dots, m, \quad (6.64)$$

where $D^{(i)}$ is a real positive semidefinite diagonal matrix and entries on the diagonal are arranged in descending order, i.e., $d_{1,1}^{(i)} \geq \dots \geq d_{t_i,t_i}^{(i)} \geq 0$. Note that

$$\operatorname{tr} (D^{(i)}) = \operatorname{tr} (R^{(i,i)}) \quad \text{for } i = 1, \dots, m. \quad (6.65)$$

By a result from majorization theory proved in [40], we have that

$$\operatorname{tr} \left((\Lambda^{(i)})^{-\frac{1}{2}} R^{(i,i)} (\Lambda^{(i)})^{-\frac{1}{2}} \right) \geq \operatorname{tr} \left((\Lambda^{(i)})^{-\frac{1}{2}} D^{(i)} (\Lambda^{(i)})^{-\frac{1}{2}} \right) \quad \text{for } i = 1, \dots, m. \quad (6.66)$$

Then from (6.63), we have

$$\operatorname{tr} \left((\Lambda^{(i)})^{-\frac{1}{2}} D^{(i)} (\Lambda^{(i)})^{-\frac{1}{2}} \right) \leq E_i \quad \text{for } i = 1, \dots, m. \quad (6.67)$$

Let

$$W = \begin{bmatrix} W^{(1)} & & \\ & \ddots & \\ & & W^{(m)} \end{bmatrix}. \quad (6.68)$$

Then W is a unitary matrix. Let

$$D = \begin{bmatrix} D^{(1)} & & \\ & \ddots & \\ & & D^{(m)} \end{bmatrix}. \quad (6.69)$$

Then

$$D = W^\dagger R W, \quad (6.70)$$

and $D \in \mathcal{D}$. From (6.70), we have

$$R = W D W^\dagger. \quad (6.71)$$

Then

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} R \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} W D W^\dagger \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.72)$$

Since W is a unitary matrix, we have that $\mathbf{Z} W$ has the same distribution as that of \mathbf{Z} . Thus

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} W D W^\dagger \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} D \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.73)$$

From (6.72) and (6.73), we have

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} R \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} D \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.74)$$

Therefore, for each $R \in \mathcal{R}$, we have found a $D \in \mathcal{D}$ such that (6.57) holds. Since $\mathcal{D} \subset \mathcal{R}$, we have

$$C = \max_{D \in \mathcal{D}} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z} D \mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.75)$$

Equivalently,

$$C = \max_{D^{(i)} \in \mathcal{D}^{(i)}, i=1, \dots, m} \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \frac{\mathbf{z}^{(i)} D^{(i)} (\mathbf{z}^{(i)})^\dagger}{\sigma^2} \right) \right) \right].$$

■

In the proof of Theorem 6.2.1, we realize that for the calculation of (6.21) transmitter i only needs to know $K^{(i)}$, but does not need to know $K^{(j)}$ for $j \neq i$; furthermore, we only need to consider the case when $\mathbf{x}^{(i)} \in \mathbb{C}^{t_i}$, $i = 1, \dots, m$ are independent. As an application of Theorem 6.2.1, we derive the capacity for a special case

of \mathbf{H} . When the covariance matrix $K^{(i)}$ of any row vector of $\mathbf{H}^{(i)}$ satisfies

$$K^{(i)} = \alpha_i^2 I_{t_i} \quad \text{for } i = 1, \dots, m, \quad (6.76)$$

where $\alpha_i \geq 0$, we can find out optimal covariance matrix of transmitted vector explicitly. For $K^{(i)}$ given in (6.76), we have $\Lambda^{(i)} = \alpha_i^2 I_{t_i}$. The assumption of $K^{(i)}$ in (6.76) corresponds to many actual situations in communication. The capacity for the channel matrix \mathbf{H} satisfying (6.76) is given by the following corollary.

Corollary 6.2.1 *Let \mathbf{H} satisfy (6.76). The capacity for channel (6.4) is*

$$C = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\alpha_i^2 E_i}{t_i} \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right]. \quad (6.77)$$

Before we prove Corollary 6.2.1, we need several lemmas. We are going to state Lemma 6.2.1, which is proved in [38](4.3.3). Let A and B be two Hermitian matrices. We arrange the eigenvalues of A and $A + B$ in descending order¹. Let $\lambda_k(A)$ be k -th eigenvalue of A and $\lambda_k(A + B)$ be k -th eigenvalue of $A + B$. Then we have the following lemma.

Lemma 6.2.1 *Let A and B be two Hermitian matrices. Assume that B is positive semidefinite and that the eigenvalues of A and $A + B$ are arranged in descending order, then*

$$\lambda_k(A) \leq \lambda_k(A + B) \quad \text{for } k = 1, 2, \dots, n. \quad (6.78)$$

Lemma 6.2.2 *Let $\mathbf{Z}^{(i)} \in \mathbb{C}^{r \times t_i}$ be a complex random matrix and let $0 \leq a_i \leq b_i$ for $i = 1, \dots, m$. Then*

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \left(a_i \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right) \right] \\ & \leq \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \left(b_i \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right) \right]. \end{aligned} \quad (6.79)$$

¹The eigenvalues of Hermitian matrices are arranged in increasing order in [38](4.3.3).

Proof of Lemma 6.2.2:

Let $\mathbf{Z}^{(i)} = Z^{(i)}$ for $i = 1, \dots, m$. Since $a_i \leq b_i$, we have $b_i - a_i \geq 0$. Since $Z^{(i)} (Z^{(i)})^\dagger$ is positive semidefinite for $i = 1, \dots, m$, we have that $(b_i - a_i)Z^{(i)} (Z^{(i)})^\dagger$ is positive semidefinite for $i = 1, \dots, m$. Therefore $\sum_{i=1}^m \left((b_i - a_i)Z^{(i)} (Z^{(i)})^\dagger \right)$ is positive semidefinite. Let

$$A = I_r + \sum_{i=1}^m \left(a_i Z^{(i)} (Z^{(i)})^\dagger \right)$$

and

$$B = \sum_{i=1}^m \left((b_i - a_i) Z^{(i)} (Z^{(i)})^\dagger \right).$$

Then by Lemma 6.2.1, we have

$$\begin{aligned} & \det \left(I_r + \sum_{i=1}^m \left(a_i Z^{(i)} (Z^{(i)})^\dagger \right) \right) \\ & \leq \det \left(I_r + \sum_{i=1}^m \left(a_i Z^{(i)} (Z^{(i)})^\dagger \right) + \sum_{i=1}^m \left((b_i - a_i) Z^{(i)} (Z^{(i)})^\dagger \right) \right). \end{aligned} \quad (6.80)$$

Therefore

$$\det \left(I_r + \sum_{i=1}^m \left(a_i Z^{(i)} (Z^{(i)})^\dagger \right) \right) \leq \det \left(I_r + \sum_{i=1}^m \left(b_i Z^{(i)} (Z^{(i)})^\dagger \right) \right). \quad (6.81)$$

Since $\log : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a strictly increasing function, we have that

$$\log \left(\det \left(I_r + \sum_{i=1}^m \left(a_i Z^{(i)} (Z^{(i)})^\dagger \right) \right) \right) \leq \log \left(\det \left(I_r + \sum_{i=1}^m \left(b_i Z^{(i)} (Z^{(i)})^\dagger \right) \right) \right).$$

Therefore

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \left(a_i \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right) \right] \\ & \leq \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \left(b_i \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right) \right]. \end{aligned}$$

■

Define $\Psi : \mathcal{Q}^{(1)} \times \cdots \times \mathcal{Q}^{(m)} \rightarrow \mathbb{R}$ by

$$\Psi(Q^{(1)}, \dots, Q^{(m)}) = \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \frac{\mathbf{Z}^{(i)} Q^{(i)} (\mathbf{Z}^{(i)})^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.82)$$

Lemma 6.2.3 Ψ has the following property

$$\begin{aligned} & \Psi \left(\sum_{j^{(1)}=1}^{L_1} a_{j^{(1)}}^{(1)} B_{j^{(1)}}^{(1)}, \dots, \sum_{j^{(m)}=1}^{L_m} a_{j^{(m)}}^{(m)} B_{j^{(m)}}^{(m)} \right) \\ & \geq \sum_{j^{(1)}=1}^{L_1} \cdots \sum_{j^{(m)}=1}^{L_m} a_{j^{(1)}}^{(1)} \cdots a_{j^{(m)}}^{(m)} \Psi \left(B_{j^{(1)}}^{(1)}, \dots, B_{j^{(m)}}^{(m)} \right). \end{aligned} \quad (6.83)$$

where $\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} = 1$ for $i = 1, \dots, m$, $a_{j^{(i)}}^{(i)} \geq 0$ and $B_{j^{(i)}}^{(i)} \in \mathcal{Q}^{(i)}$ for $j^{(i)} = 1, \dots, L_i$, $i = 1, \dots, m$.

Proof of Lemma 6.2.3:

We note that $\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} B_{j^{(i)}}^{(i)} \in \mathcal{Q}_{t_i}$, for $i = 1, \dots, m$. Then

$$\begin{aligned} & \Psi \left(\sum_{j^{(1)}=1}^{L_1} a_{j^{(1)}}^{(1)} B_{j^{(1)}}^{(1)}, \dots, \sum_{j^{(m)}=1}^{L_m} a_{j^{(m)}}^{(m)} B_{j^{(m)}}^{(m)} \right) \\ & = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} \left(\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} B_{j^{(i)}}^{(i)} \right) (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \\ & = \mathbb{E} \left[\log \left(\det \left(\sum_{j^{(1)}=1}^{L_1} a_{j^{(1)}}^{(1)} \left(I_r + \frac{1}{\sigma^2} \sum_{i=2}^m \mathbf{Z}^{(i)} \left(\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} B_{j^{(i)}}^{(i)} \right) (\mathbf{Z}^{(i)})^\dagger + \right. \right. \right. \right. \\ & \quad \left. \left. \left. \frac{1}{\sigma^2} \mathbf{Z}^{(1)} B_{j^{(1)}}^{(1)} (\mathbf{Z}^{(1)})^\dagger \right) \right) \right) \right] \end{aligned} \quad (6.84)$$

$$\begin{aligned} & \geq \sum_{j^{(1)}=1}^{L_1} a_{j^{(1)}}^{(1)} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=2}^m \mathbf{Z}^{(i)} \left(\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} B_{j^{(i)}}^{(i)} \right) (\mathbf{Z}^{(i)})^\dagger + \right. \right. \right. \\ & \quad \left. \left. \left. \frac{1}{\sigma^2} \mathbf{Z}^{(1)} B_{j^{(1)}}^{(1)} (\mathbf{Z}^{(1)})^\dagger \right) \right) \right], \end{aligned} \quad (6.85)$$

where from (6.84) to (6.85) we used the property that $\log(\det(\cdot))$ is a strictly convex cap function on the set of positive definite Hermitian matrices [38](7.6.7) [13]. We

now carry out the above steps to $i = 2, \dots, m$ and have that

$$\begin{aligned} & \Psi \left(\sum_{j^{(1)}=1}^{L_1} a_{j^{(1)}}^{(1)} B_{j^{(1)}}^{(1)}, \dots, \sum_{j^{(m)}=1}^{L_m} a_{j^{(m)}}^{(m)} B_{j^{(m)}}^{(m)} \right) \\ & \geq \sum_{j^{(1)}=1}^{L_1} \cdots \sum_{j^{(m)}=1}^{L_m} a_{j^{(1)}}^{(1)} \cdots a_{j^{(m)}}^{(m)} \Psi \left(B_{j^{(1)}}^{(1)}, \dots, B_{j^{(m)}}^{(m)} \right). \end{aligned}$$

■

Proof of Corollary 6.2.1:

The proof procedure here is similar to that in [80]. Define

$$\mathcal{P}_{t_i} = \{ \Pi : \Pi \text{ is a } t_i \times t_i \text{ permutation matrix} \} \quad \text{for } i = 1, \dots, m. \quad (6.86)$$

and

$$L_i = \text{Cardinality of } \mathcal{P}_{t_i}. \quad (6.87)$$

Let $\Pi_{j^{(i)}}^{(i)} \in \mathcal{P}_{t_i}$ and

$$a_{j^{(i)}}^{(i)} = \frac{1}{L_i} \quad \text{for } j^{(i)} = 1, \dots, L_i, \quad i = 1, \dots, m. \quad (6.88)$$

From Lemma 6.2.3, we have that

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} \left(\sum_{j^{(i)}=1}^{L_i} a_{j^{(i)}}^{(i)} \Pi_{j^{(i)}}^{(i)} D^{(i)} \left(\Pi_{j^{(i)}}^{(i)} \right)^\dagger \right) (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \\ & \geq \sum_{j^{(1)}=1}^{L_1} \cdots \sum_{j^{(m)}=1}^{L_m} a_{j^{(1)}}^{(1)} \cdots a_{j^{(m)}}^{(m)} \cdot \\ & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} \Pi_{j^{(i)}}^{(i)} D^{(i)} \left(\Pi_{j^{(i)}}^{(i)} \right)^\dagger (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \end{aligned} \quad (6.89)$$

Note that for any permutation matrix Π , we have

$$\Pi = \Pi^\dagger$$

Since $\mathbf{Z}\mathbf{I}\mathbf{I}$ has the same distribution as that of \mathbf{Z} , we have that

$$\mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\mathbf{Q}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right] = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{\mathbf{Z}\mathbf{I}\mathbf{I}\mathbf{Q}\mathbf{I}\mathbf{I}\mathbf{Z}^\dagger}{\sigma^2} \right) \right) \right]. \quad (6.90)$$

Therefore

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \frac{1}{\sigma^2} \mathbf{Z}^{(i)} \Pi_{j^{(i)}}^{(i)} D^{(i)} \left(\Pi_{j^{(i)}}^{(i)} \right)^\dagger \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right] \\ &= \mathbb{E} \left[\log \left(\det \left(I_r + \sum_{i=1}^m \frac{1}{\sigma^2} \mathbf{Z}^{(i)} D^{(i)} \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right]. \end{aligned} \quad (6.91)$$

Substituting (6.88) and (6.91) into the right-hand side of (6.89), we have that

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} \left(\sum_{j^{(i)}=1}^{L_i} \frac{1}{L_i} \Pi_{j^{(i)}}^{(i)} D^{(i)} \left(\Pi_{j^{(i)}}^{(i)} \right)^\dagger \right) \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right] \\ & \geq \sum_{j^{(1)}=1}^{L_1} \cdots \sum_{j^{(m)}=1}^{L_m} \frac{1}{L_1} \cdots \frac{1}{L_m} \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} D^{(i)} \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right]. \end{aligned} \quad (6.92)$$

Note that

$$\sum_{j^{(i)}=1}^{L_i} \frac{1}{L_i} \Pi_{j^{(i)}}^{(i)} D^{(i)} \left(\Pi_{j^{(i)}}^{(i)} \right)^\dagger = \frac{\text{tr}(D^{(i)})}{t_i} I_{t_i} \quad \text{for } i = 1, \dots, m. \quad (6.93)$$

Therefore we have

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\text{tr}(D^{(i)})}{t_i} \mathbf{Z}^{(i)} I_{t_i} \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right] \\ & \geq \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} D^{(i)} \left(\mathbf{Z}^{(i)} \right)^\dagger \right) \right) \right]. \end{aligned} \quad (6.94)$$

Recall that

$$\Lambda^{(i)} = \alpha_i^2 I_{t_i}$$

for the $K^{(i)}$ given in (6.76). Since

$$\text{tr} \left(\left(\Lambda^{(i)} \right)^{-\frac{1}{2}} D^{(i)} \left(\Lambda^{(i)} \right)^{-\frac{1}{2}} \right) \leq E_i \quad \text{for } i = 1, \dots, m,$$

we have that

$$\text{tr} \left(D^{(i)} \right) \leq \alpha_i^2 E_i \quad \text{for } i = 1, \dots, m.$$

Thus by Lemma 6.2.2, we have

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\alpha_i^2 E_i}{t_i} \mathbf{Z}^{(i)} I_{t_i} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \\ & \geq \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\text{tr}(D^{(i)})}{t_i} \mathbf{Z}^{(i)} I_{t_i} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right]. \end{aligned} \quad (6.95)$$

Therefore from (6.94) and (6.95), we have

$$\begin{aligned} & \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\alpha_i^2 E_i}{t_i} \mathbf{Z}^{(i)} I_{t_i} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \\ & \geq \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \mathbf{Z}^{(i)} D^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right]. \end{aligned} \quad (6.96)$$

If we choose

$$D^{(i)} = \frac{\alpha_i^2 E_i}{t_i} I_{t_i} \quad \text{for } i = 1, \dots, m, \quad (6.97)$$

we will achieve the maximum in (6.96) and consequently from (6.21)

$$C = \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\alpha_i^2 E_i}{t_i} \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right].$$

■

From Corollary 6.2.1, we see that if \mathbf{H} satisfies (6.76), each transmitter just divides its energy evenly among its transmitting antennae and it does not need any feedback of the channel status. The capacity formula for distributed space-time coding with individual energy constraint is useful when we consider the transport efficiency for the network with cooperative communication.

6.3 Calculation of Capacity

For the calculation of (6.77), we see that we can not apply the method in [80], where the author simplified the calculation by only considering the eigenvalues of $\mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger$. Such a method would require that $\mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger$ be simultaneously diagonalizable for $i = 1, \dots, m$, which may not always be true. In general, when

$t_1 + \cdots + t_m > r$, $\mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger$ follows Wishart distribution [4], which can be thought of as a generalization of χ^2 distribution. We give the calculation of capacity for the simple case when $r = 1$. When $r \geq 2$ and $t \geq r$, the integral will then be multidimensional and the calculation of C is quite involved since $\mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger$ follows Wishart distribution.

When $r = 1$, we have that

$$\begin{aligned} C &= \mathbb{E} \left[\log \left(\det \left(I_r + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{\alpha_i^2 E_i}{t_i} \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right) \right] \\ &= \mathbb{E} \left[\log \left(1 + \sum_{i=1}^m \frac{\alpha_i^2 E_i}{\sigma^2 t_i} \mathbf{Z}^{(i)} (\mathbf{Z}^{(i)})^\dagger \right) \right] \\ &= \mathbb{E} \left[\log \left(1 + \sum_{i=1}^m \frac{\alpha_i^2 E_i}{\sigma^2 t_i} \mathbf{v}_i \right) \right] \end{aligned} \quad (6.98)$$

$$= \mathbb{E} \left[\log \left(1 + \sum_{i=1}^m \frac{\alpha_i^2 E_i}{2\sigma^2 t_i} \mathbf{u}_i \right) \right], \quad (6.99)$$

where $\mathbf{v}_i = \sum_{j=1}^{t_i} \left((\Re(\mathbf{Z}_j^{(i)}))^2 + (\Im(\mathbf{Z}_j^{(i)}))^2 \right)$ in (6.98), and \mathbf{u}_i in (6.99) is χ^2 -distributed with $2t_i$ degrees of freedom. The probability density function is given by [91]

$$f_{\mathbf{u}_i}(u_i) = \begin{cases} \frac{e^{-u_i/2} u_i^{t_i-1}}{2^{t_i} \Gamma(t_i)} & u_i \geq 0 \\ 0 & u_i < 0. \end{cases} \quad \text{for } i = 1, \dots, m. \quad (6.100)$$

Therefore

$$C = \int_0^\infty \cdots \int_0^\infty \log \left(1 + \sum_{i=1}^m \frac{\alpha_i^2 E_i}{2\sigma^2 t_i} u_i \right) \prod_{i=1}^m \frac{e^{-u_i/2} u_i^{t_i-1}}{2^{t_i} \Gamma(t_i)} du_1 \cdots du_m. \quad (6.101)$$

Let $s_i = \frac{u_i}{2}$, we have

$$C = \prod_{j=1}^m \frac{1}{\Gamma(t_j)} \int_0^\infty \cdots \int_0^\infty \log \left(1 + \frac{1}{\sigma^2} \sum_{i=1}^m \frac{s_i \alpha_i^2 E_i}{t_i} \right) \left(\prod_{i=1}^m s_i^{t_i-1} \right) e^{-\sum_{i=1}^m s_i} ds_1 \cdots ds_m. \quad (6.102)$$

From (6.102), we see that when $t_i = 1$ for $i = 1, \dots, m$, we have that

$$C = \int_0^\infty \cdots \int_0^\infty \log \left(1 + \frac{1}{\sigma^2} \sum_{i=1}^m s_i \alpha_i^2 E_i \right) e^{-\sum_{i=1}^m s_i} ds_1 \cdots ds_m. \quad (6.103)$$

This is the same as the capacity for the multiple access Gaussian channel given in [23].

6.4 Cooperative Communication for Linear Networks

In this section, we are going to investigate cooperative communication between two linear networks.

6.4.1 Cooperative Linear Networks

We consider two linear networks shown in Figure 5.3. The assumptions in this section for the two linear networks are almost the same as the assumptions in Section 5.4.1, except for the ones about interference. In this section, we assume that A_i and B_i on both linear networks have the same information source to transmit to A_{i+1} and B_{i+1} , respectively. Therefore it is possible for nodes A_i and B_i to cooperate in the form of distributed space-time coding. Under this situation, A_{i+1} treats the signals that it receives from both A_i and B_i as useful information and tries to decode both signals. Similarly, B_{i+1} performs the same operation. We assume threshold model for each node so that nodes in the linear networks are able to communicate at the rate of the channel capacity given in Section 6.2.

6.4.2 Transport Efficiency

In order to prevent confusion about notation, we will analyze the transport efficiency of linear network A when cooperative transmission is present. The analysis for the transport efficiency of linear network B can be done similarly. Let us consider node A_2 on linear network A . Let $\mathbf{y} \in \mathbb{C}^r$ be the received signal at A_2 , $\mathbf{x}^{(1)} \in \mathbb{C}^{t_1}$ be transmitted signal from A_1 on linear network A , $\mathbf{x}^{(2)} \in \mathbb{C}^{t_2}$ be transmitted signal from B_1 on linear network B , and $\mathbf{n} \in \mathbb{C}^r$ be the additive noise at A_2 . Let

$\mathbf{H}^{(1)} \in \mathbf{C}^{r \times t_1}$ be the channel matrix from A_1 to A_2 and $\mathbf{H}^{(2)} \in \mathbf{C}^{r \times t_2}$ be the channel matrix from B_1 to A_2 . We assume that

$$\mathbf{H}^{(j)} = \alpha_j \mathbf{Z}^{(j)} \quad \text{for } j = 1, 2, \quad (6.104)$$

where α_j captures the large scale propagation loss and $\mathbf{Z}^{(j)}$ captures the small scale variation of the channel. From the discussion of channel models given in Chapter V, we have that

$$\alpha_1 = \sqrt{\beta(d_e/k)} \quad (6.105)$$

and

$$\alpha_2 = \sqrt{\beta \left(\sqrt{d_b^2 + (d_e/k)^2} \right)}. \quad (6.106)$$

Furthermore, all entries in $\mathbf{Z}^{(j)}$ are independent and identically distributed proper Gaussian random variables with mean 0 and variance 1. Let \mathbf{n} be a proper complex Gaussian vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$. Define

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}^{(1)} \\ \mathbf{x}^{(2)} \end{bmatrix} \quad (6.107)$$

and

$$\mathbf{H} = [\mathbf{H}^{(1)} \quad \mathbf{H}^{(2)}]. \quad (6.108)$$

For a frequency non-selective slowly fading channel, we have the following channel model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (6.109)$$

with the constraint that $\text{tr} \left(\mathbb{E} \left[\mathbf{x}^{(i)} (\mathbf{x}^{(i)})^\dagger \right] \right) = E_o$, for $i = 1, 2$, where E_o is given in (5.25). From Corollary 6.2.1, we know that the mutual information between \mathbf{x} and \mathbf{y} is maximized when the sub-vector $\mathbf{x}^{(1)}$ is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\frac{E_o}{t_1} I_{t_1}$ and the sub-vector $\mathbf{x}^{(2)}$ is a proper

complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\frac{E_o}{t_2} I_{t_2}$. The capacity of the channel in (6.109) is given by (6.102), i.e.,

$$C^{(ct)}(d_e/k, d_b) = \frac{1}{\Gamma(t_1)} \frac{1}{\Gamma(t_2)} \int_0^\infty \int_0^\infty \log \left(1 + \frac{s_1 \alpha_1^2 E_o}{\sigma^2 t_1} + \frac{s_2 \alpha_2^2 E_o}{\sigma^2 t_2} \right) s_1^{t_1-1} s_2^{t_2-1} e^{-(s_1+s_2)} ds_1 ds_2. \quad (6.110)$$

The transport efficiency of the linear network A is given by

$$\mu^{(ct)}(k, P_{max}, C^{(ct)}(d_e/k, d_b), d_e) = \frac{1}{d_e} (C^{(ct)}(d_e/k, d_b))^2 \frac{d_e/k}{E_c + E_p}. \quad (6.111)$$

As we did in Chapter IV, we can replace d_e/k by d_a . Let

$$g^{(ct)}(d_a, d_b) = d_a (C^{(ct)}(d_a, d_b))^2. \quad (6.112)$$

Then

$$\mu^{(ct)}(k, P_{max}, C^{(ct)}(d_e/k, d_b), d_e) = \frac{g^{(ct)}(d_a, d_b)}{d_e(E_c + E_p)}. \quad (6.113)$$

We will use $g^{(ct)}$ as a performance measure to compare the performance of linear networks when different number of transmitting antennae is used. Note that the unit of $g^{(ct)}$ is bit²·m/(channel use). For each d_b , we optimize $g^{(ct)}$ over d_a .

6.5 Numerical Results and Discussion

For numerical calculation, we assume the numerical values for the parameters given in Section 5.5. We also set $t_1 = t_2 = t$ and $r = 1$. The optimized $g^{(ct)}$ and optimal adjacent distance are shown in Figure 6.1 and Figure 6.2, respectively.

Intuitively we know that the transport efficiency for both networks decreases as the distance between the two linear networks increases. From the result presented in Figure 6.1, we see if d_b is above 2000 m, the transport efficiency starts to flatten,

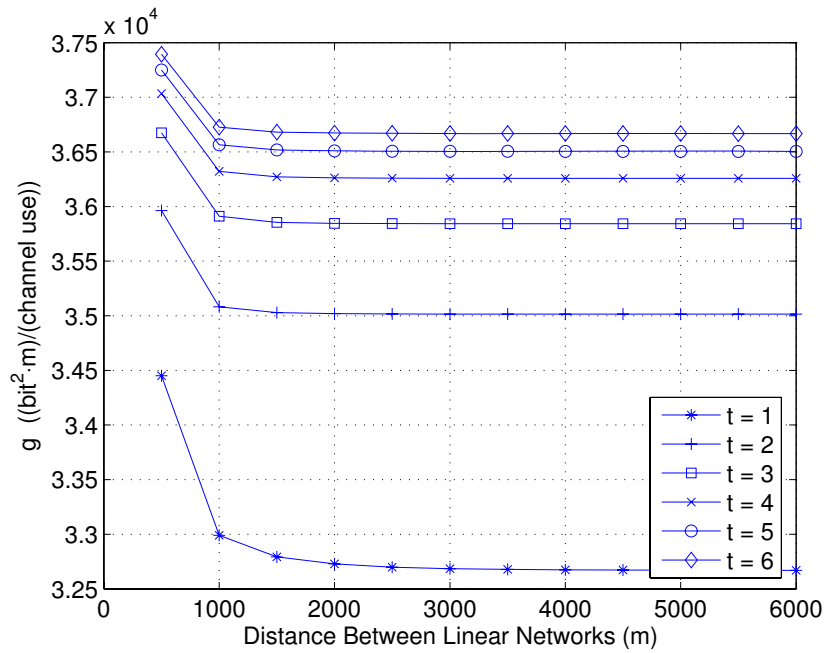


Figure 6.1: Optimal $g^{(ct)}$ versus the distance between two linear networks, where each node has $r = 1$ receiving antenna.

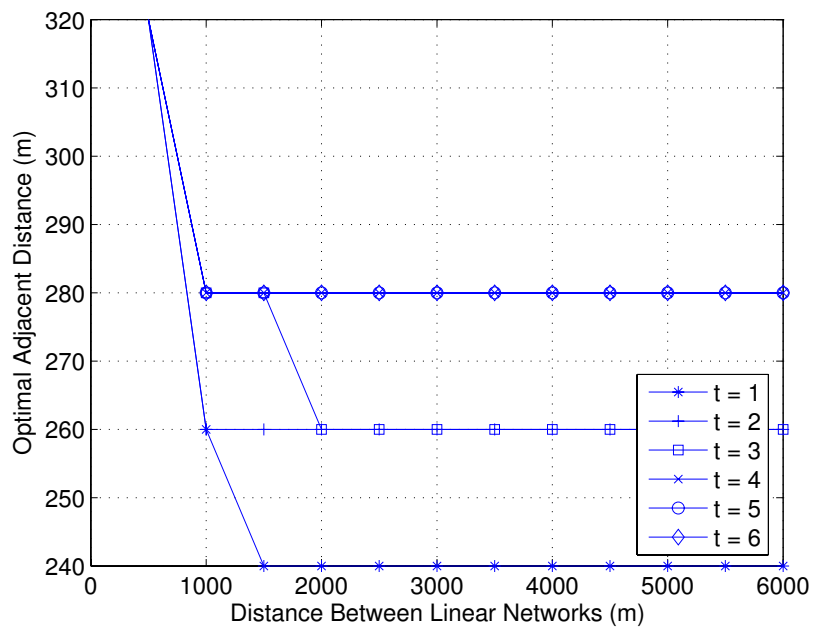


Figure 6.2: Optimal adjacent distance versus the distance between two linear networks, where each node has $r = 1$ receiving antenna.

and there is not much benefit of using cooperative communication. Also for large enough distance between the two linear networks, increasing the number of transmitting antennae above 3 does not generate too much gain in transport efficiency. Cooperative communication in our scenario is most useful when destination nodes on both linear networks want to receive the same information. In this case, each receiver on each linear network really has some bonus signal from the transmission on the other linear network. As a practical example, if there are some mirror sites in the Internet around several nodes which request the same file, it may be beneficial to allow more than one mirror site to communicate with these nodes simultaneously. From the energy consumption point of view, since distributed space-time coding demands very high signal processing complexity, it may not be desirable to perform cooperative communication when the distance between two linear networks is large.

CHAPTER VII

CONCLUSION AND FUTURE RESEARCH

7.1 Summary of Contributions and Conclusion

We presented a generic integrated design methodology that is suitable for many kinds of mobile systems. The integrated design methodology takes into account the coupling among the subsystems and simultaneously optimizes their operation under an energy constraint. Using our methodology, we were able to optimize a communication system from network level down to circuit level in a reasonable amount of time. We showed the improvement in performance that the integrated design methodology achieves over traditional design methodologies and the tradeoff between energy consumption and performance.

Routing is an efficient method for connectivity and low energy consumption of wireless networks. When each node is equipped with an omni-directional antenna, a point-to-multipoint connection is often available for routing purposes. When the design goal is to minimize the maximum power consumed by the nodes in a network, we described a polynomial-time complexity algorithm to assign power to each node for unicast, broadcast, and multicast sessions. When the design goal is to minimize the total power consumed by the nodes in a network, we described a polynomial-time complexity algorithm to assign power to each node for a unicast session and showed

that the computational complexity of routing algorithms for broadcast and multicast sessions is NP-hard. Our routing algorithms do not require the underlying graph we get after the power assignment be strongly connected.

We introduced transport efficiency as a network performance measure to capture both bandwidth efficiency and energy efficiency of a linear network. We showed that the optimal transport efficiency is inversely proportional to the end-to-end distance for the threshold model. We observed from our numerical results that the optimal transport efficiency is approximately inversely proportional to the end-to-end distance for the cutoff-rate model, uncoded model and convolutional-coded model. We demonstrated that amplifier characteristics and receiver processing energy has a direct effect on the optimal transport efficiency.

We investigated the interference caused by space-time coding and the ordinary end-fire antenna array to neighboring networks. Our analytical and numerical results suggested that the end-fire antenna array gives higher transport efficiency than space-time coding when the number of receiving antennae is small and gives transport efficiency close to that for space-time coding when two linear networks are close to each other. On the other hand, space-time coding gives higher transport efficiency than the ordinary end-fire antenna array when the number of receiving antennae is large and when two linear networks are far apart.

We showed that cooperative communication with space-time coding between linear networks can improve transport efficiency, but it only gives marginal benefit if the cooperating networks are separated too far apart. This suggests that cooperative communication among networks should be used carefully since it usually incurs significant signal processing at the receiver.

7.2 Future Research

It should be interesting to apply our integrated design methodology to more network scenarios. It will also be interesting to see how transport efficiency for linear networks can be extended to two-dimensional and three-dimensional networks. Of course, a general theory to measure network performance is always desirable.

Shannon's channel coding theory provides a long-term performance limit for point-to-point communication in the additive white Gaussian noise (AWGN) channel. This performance limit has essentially been reached with the invention of turbo codes, LDPC, and turbo-like coding techniques. There have been extensions to Shannon's theory to provide long-term performance limit for communications in broadcast channels and multiple-access channels. There have also been extensions to Shannon's theory to provide delay-limited capacity for communications with delay constraints. Network protocol information has been studied from an information-theoretic perspective [29]. Researchers also tried to combine queueing theory and information theory to study the multi-access channel [81]. However, we still lack a profound theory in the wireless or wired communication networks that can play the same role as Shannon's theory for point-to-point communications. Wireless communication networks do present a lot of challenges in many different aspects, most often researchers tried to answer the challenges in each aspect separately. The transit and burstiness nature of networks may require a whole new profound theory rather than extensions of Shannon's theory.

APPENDICES

APPENDIX A

SIMULATED ANNEALING ALGORITHM

Hide-and-Seek is a global continuous optimization algorithm of the simulated-annealing-type [70]. The algorithm can find the maximum of a continuous function (differentiability is not required) over an arbitrary closed and bounded feasible region (convexity is not required). It is guaranteed to converge (in probability) under mild assumptions on the cooling schedule used; the reader is referred to [70] for the precise technical details. *Hide-and-Seek* is an iterative algorithm where at each iteration a new candidate successor point in the feasible region is generated randomly according to the following strategy: from the current iteration point a search direction is chosen at random inside the feasible region; then the candidate successor is obtained according to a uniform distribution in the direction chosen within the feasible region. (In our implementation of the algorithm, the standard deviation of each component in the search direction vector is proportional to the difference of the corresponding upper and lower bounds on that component.) The initial candidate point can be any point in the feasible region. Candidate successors are “accepted” according to the familiar Metropolis criterion of simulated annealing algorithms [70]. A feature of *Hide-and-Seek* is that the cooling schedule parameterizing the Metropolis criterion is adaptive in the sense that it depends on all previously accepted points.

In our integrated design methodology, the function to optimize is evaluated by means of a simulation in *OPNET*. Hence, we cannot formally prove that it satisfies the required continuity assumption for convergence of *Hide-and-Seek*. In our simulation experiments, we let the algorithm run for 200 iterations. By the time we reached 200 iterations, the change in the best maximum found so far was negligible.

The detailed steps of the *Hide-and-Seek* algorithm used in the global optimization Step 1 in Section 2.4 are described below.

Step 1.1 Let M_1, M_2, M_3, M_4 and m_1, m_2, m_3, m_4 be the upper and lower bounds of the design variables T, q, E_{ct}, E_{cr} respectively. The “optimizer” module arbitrarily chooses an initial feasible design $\underline{x} = [T, q, E_{ct}, E_{cr}]^T$ whose performance, denoted by $g(\underline{x})$, is to be evaluated by the “network simulator” module. The “optimizer” module sets $\underline{x}_{bsf} = \underline{x}$ and $f_{bsf} = f(\underline{x}) = -g(x)$, where *bsf* stands for “best so far”. It also sets the loop counter $k = 0$, sets the initial temperature τ to be a very large value, and sets the total number of iterations *numiter* to the desired value.

Step 1.2 The “optimizer” module generates N zero-mean and unit-variance Gaussian random variables ξ_1, \dots, ξ_N , where N is the total number of design variables ($N = 4$ in our example). It then generates a search direction $\underline{\zeta} = [\zeta_1, \dots, \zeta_N]^T$ according to

$$\zeta_i = \xi_i(M_i - m_i), i = 1, \dots, N. \quad (\text{A.1})$$

Let $\underline{r} = \underline{\zeta} / \|\underline{\zeta}\|$ and let d_n and d_p be the largest distances from \underline{x} to the boundary of the feasible region along the direction $-\underline{r}$ and \underline{r} respectively. The “optimizer” module generates a new candidate feasible design $\underline{y}_k =$

$[T, q, E_{ct}, E_{cr}]^T$ according to

$$\underline{y}_k = \underline{x} + (-d_n(1 - u) + u * d_p)\underline{r}, \quad (\text{A.2})$$

where u is uniformly distributed random variable on $[0, 1]$. The corresponding performance $f(\underline{y}_k)$ is evaluated by the “network simulator” module as in Step 1.1.

Step 1.3 The “optimizer” module accepts \underline{y}_k with probability $\beta(\tau, f(\underline{x}), f(\underline{y}_k))$, where τ stands for current temperature and

$$\beta(\tau, f(\underline{x}), f(\underline{y}_k)) = \min(1, e^{(f(\underline{y}_k) - f(\underline{x}))/\tau}). \quad (\text{A.3})$$

This is called *Metropolis acceptance criterion*. If the “optimizer” module accepts \underline{y}_k , it sets $\underline{x} = \underline{y}_k$ and $f(\underline{x}) = f(\underline{y}_k)$.

Step 1.4 If $f(\underline{x}) > f_{bsf}$, the “optimizer” module sets

$$f_1 = \min(f_{bsf} + (f(\underline{x}) - f_{bsf})/\alpha, 0.0), \quad (\text{A.4})$$

$$\tau = \frac{2(f_1 - f(\underline{x}))}{\eta(N)}, \quad (\text{A.5})$$

$$f_{bsf} = f(\underline{x}), \quad (\text{A.6})$$

$$\underline{x}_{bsf} = \underline{x}, \quad (\text{A.7})$$

where α is a fixed scaling factor and is set to

$$\alpha = \frac{1}{0.9^{N/2.0}} - 1. \quad (\text{A.8})$$

The denominator $\eta(N)$ in (A.5) is chosen such that the CDF of the χ^2 distribution with N degrees of freedom evaluated at $\eta(N)$ is 0.99.

Next, the “optimizer” module increments k . If k does not exceed *numiter*, go to Step 1.2. Otherwise the “optimizer” module stops and returns the best design found \underline{x}_{bsf} and its corresponding value f_{bsf} .

APPENDIX B

PROPER COMPLEX RANDOM VARIABLES, VECTORS, AND PROCESSES

We introduce some properties of proper complex random variables, vectors, and processes which are commonly used in communication literature, but not always explicitly explained. We assume that all the properties of *real* random variables, vectors, and processes have already been well defined. For example, a real random variable has a cumulative distribution function; a real random vector has a joint cumulative distribution function; a continuous real random variable has a probability density function; etc. For ease of reference, we state definitions and theorems relevant to this dissertation without any proof because most of them can be found in [58].

We assume the reader has some basic understanding of measure theory [10, 14, 64]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with σ -algebra \mathcal{F} being a collection of measurable sets of Ω and \mathbb{P} being a probability measure. Recall that $\mathbf{x} : \Omega \rightarrow \mathbb{R}$ is a real random variable if \mathbf{x} is Borel-measurable.

B.1 Proper Complex Random Variables

Definition B.1.1 $\mathbf{x} : \Omega \rightarrow \mathbb{C}$ is a **complex random variable** if

$$\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}, \tag{B.1}$$

where $\mathbf{x}_{\Re} : \Omega \rightarrow \mathbb{R}$ and $\mathbf{x}_{\Im} : \Omega \rightarrow \mathbb{R}$ are both real random variables. \mathbf{x}_{\Re} is called the real part of \mathbf{x} and \mathbf{x}_{\Im} is called the imaginary part of \mathbf{x} .

Definition B.1.2 The **probability density function of a complex random variable** $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ is

$$p_{\mathbf{x}}(x_{\Re} + ix_{\Im}) = p_{\mathbf{x}_{\Re}, \mathbf{x}_{\Im}}(x_{\Re}, x_{\Im}). \quad (\text{B.2})$$

Definition B.1.3 The **expectation of a complex random variable** $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ is

$$\mathbb{E}[\mathbf{x}] = \mathbb{E}[\mathbf{x}_{\Re}] + i\mathbb{E}[\mathbf{x}_{\Im}]. \quad (\text{B.3})$$

Let $x \in \mathbb{C}$ be a complex number and let x^* be the complex conjugate of x .

Definition B.1.4 The **variance of a complex random variable** \mathbf{x} is

$$\text{Var}(\mathbf{x}) = \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^*]. \quad (\text{B.4})$$

Definition B.1.5 A complex random variable \mathbf{x} is **proper** if

$$\mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])^2] = 0. \quad (\text{B.5})$$

Note that (B.5) implies

$$\text{Var}(\mathbf{x}_{\Re}) = \text{Var}(\mathbf{x}_{\Im}) \quad \text{and} \quad \text{Cov}(\mathbf{x}_{\Re}, \mathbf{x}_{\Im}) = 0. \quad (\text{B.6})$$

Therefore the real part and imaginary part of X are uncorrelated. Let

$$\text{Image}(\mathbf{x}) = \mathbf{x}(\Omega) = \{\mathbf{x}(\omega) : \omega \in \Omega\}. \quad (\text{B.7})$$

Let $\mathbf{x}(\Omega) \subseteq \mathcal{D} \subseteq \mathbb{C}$ and let $F : \mathcal{D} \rightarrow \mathbb{C}$ be a measurable function.

Definition B.1.6 The expectation of F is defined

$$\mathbb{E}[F(\mathbf{x})] = \mathbb{E}[(F(\mathbf{x}_{\Re} + i\mathbf{x}_{\Im}))_{\Re}] + i\mathbb{E}[(F(\mathbf{x}_{\Re} + i\mathbf{x}_{\Im}))_{\Im}]. \quad (\text{B.8})$$

Equivalently,

$$\mathbb{E}[F(\mathbf{x})] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(x_{\Re} + ix_{\Im}) p_{\mathbf{x}}(x_{\Re} + ix_{\Im}) dx_{\Re} dx_{\Im} \quad (\text{B.9})$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_{\Re} + ix_{\Im}))_{\Re} p_{\mathbf{x}_{\Re}, \mathbf{x}_{\Im}}(x_{\Re}, x_{\Im}) dx_{\Re} dx_{\Im} + \\ & i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (F(x_{\Re} + ix_{\Im}))_{\Im} p_{\mathbf{x}_{\Re}, \mathbf{x}_{\Im}}(x_{\Re}, x_{\Im}) dx_{\Re} dx_{\Im}. \end{aligned} \quad (\text{B.10})$$

B.2 Proper Complex Random Vectors

Definition B.2.1 $\mathbf{x} : \Omega \rightarrow \mathbb{C}^n$ is a **complex random vector** if

$$\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}, \quad (\text{B.11})$$

where $\mathbf{x}_{\Re} : \Omega \rightarrow \mathbb{R}^n$ and $\mathbf{x}_{\Im} : \Omega \rightarrow \mathbb{R}^n$ are both real random vectors. \mathbf{x}_{\Re} is called the real part of \mathbf{x} and \mathbf{x}_{\Im} is called the imaginary part of \mathbf{x} . $\mathbf{Z} : \Omega \rightarrow \mathbb{C}^{m \times n}$ is a **complex random matrix** if

$$\mathbf{Z} = \mathbf{Z}_{\Re} + i\mathbf{Z}_{\Im}, \quad (\text{B.12})$$

where $\mathbf{Z}_{\Re} : \Omega \rightarrow \mathbb{R}^{m \times n}$ and $\mathbf{Z}_{\Im} : \Omega \rightarrow \mathbb{R}^{m \times n}$ are both real random matrices. \mathbf{Z}_{\Re} is called the real part of \mathbf{Z} and \mathbf{Z}_{\Im} is called the imaginary part of \mathbf{Z} .

Definition B.2.2 The **probability density function of a complex random vector** $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ is

$$p_{\mathbf{x}}(x_{\Re} + ix_{\Im}) = p_{\mathbf{x}_{\Re}, \mathbf{x}_{\Im}}(x_{\Re}, x_{\Im}). \quad (\text{B.13})$$

Definition B.2.3 Two complex random vectors \mathbf{x} and \mathbf{y} are **independent** if

$$\begin{bmatrix} \mathbf{x}_{\Re} \\ \mathbf{x}_{\Im} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{y}_{\Re} \\ \mathbf{y}_{\Im} \end{bmatrix}$$

are independent real random vectors.

Definition B.2.4 The expectation of a complex random vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_m \end{bmatrix}$$

is

$$m_{\mathbf{x}} = \mathbb{E}[\mathbf{x}] = \begin{bmatrix} \mathbb{E}[\mathbf{x}_1] \\ \vdots \\ \mathbb{E}[\mathbf{x}_m] \end{bmatrix}. \quad (\text{B.14})$$

The expectation of a complex random matrix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{z}_{1,1} & \cdots & \mathbf{z}_{1,n} \\ \vdots & \ddots & \vdots \\ \mathbf{z}_{m,1} & \cdots & \mathbf{z}_{m,n} \end{bmatrix}$$

is

$$m_{\mathbf{Z}} = \mathbb{E}[\mathbf{Z}] = \begin{bmatrix} \mathbb{E}[\mathbf{z}_{1,1}] & \cdots & \mathbb{E}[\mathbf{z}_{1,n}] \\ \vdots & \ddots & \vdots \\ \mathbb{E}[\mathbf{z}_{m,1}] & \cdots & \mathbb{E}[\mathbf{z}_{m,n}] \end{bmatrix}. \quad (\text{B.15})$$

Let $A \in \mathbb{C}^{m \times n}$ be a complex matrix. Let A^\dagger be the **conjugate transpose** of A and let A^T be the **transpose** of A .

Definition B.2.5 The **cross-correlation matrix** of complex random vectors \mathbf{x} and \mathbf{y} is

$$R_{\mathbf{x},\mathbf{y}} = \mathbb{E}[\mathbf{x}\mathbf{y}^\dagger]. \quad (\text{B.16})$$

The **pseudo-cross-correlation matrix** of complex random vectors \mathbf{x} and \mathbf{y} is

$$\tilde{R}_{\mathbf{x},\mathbf{y}} = \mathbb{E}[\mathbf{x}\mathbf{y}^T]. \quad (\text{B.17})$$

Definition B.2.6 The covariance matrix of complex random vectors of $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ and $\mathbf{y} = \mathbf{y}_{\Re} + i\mathbf{y}_{\Im}$ is

$$C_{\mathbf{x},\mathbf{y}} = \mathbb{E} \left[(\mathbf{x} - m_{\mathbf{x}}) (\mathbf{y} - m_{\mathbf{y}})^{\dagger} \right]. \quad (\text{B.18})$$

The pseudo-covariance matrix of complex random vectors of $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ and $\mathbf{y} = \mathbf{y}_{\Re} + i\mathbf{y}_{\Im}$ is

$$\tilde{C}_{\mathbf{x},\mathbf{y}} = \mathbb{E} \left[(\mathbf{x} - m_{\mathbf{x}}) (\mathbf{y} - m_{\mathbf{y}})^T \right]. \quad (\text{B.19})$$

Note that

$$\begin{aligned} C_{\mathbf{x},\mathbf{y}} &= \mathbb{E} \left[(\mathbf{x} - m_{\mathbf{x}}) (\mathbf{y} - m_{\mathbf{y}})^{\dagger} \right] \\ &= (\text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Re}) + \text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Im})) + \\ &\quad i (\text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Re}) - \text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Im})); \end{aligned} \quad (\text{B.20})$$

$$\begin{aligned} \tilde{C}_{\mathbf{x},\mathbf{y}} &= \mathbb{E} \left[(\mathbf{x} - m_{\mathbf{x}}) (\mathbf{y} - m_{\mathbf{y}})^T \right] \\ &= (\text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Re}) - \text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Im})) + \\ &\quad i (\text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Re}) + \text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Im})). \end{aligned} \quad (\text{B.21})$$

Conversely, we have

$$\text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Re}) = \frac{1}{2} \left(C_{\mathbf{x},\mathbf{y}} + \tilde{C}_{\mathbf{x},\mathbf{y}} \right)_{\Re}; \quad (\text{B.22})$$

$$\text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Im}) = \frac{1}{2} \left(C_{\mathbf{x},\mathbf{y}} - \tilde{C}_{\mathbf{x},\mathbf{y}} \right)_{\Re}; \quad (\text{B.23})$$

$$\text{Cov}(\mathbf{x}_{\Im}, \mathbf{y}_{\Re}) = \frac{1}{2} \left(C_{\mathbf{x},\mathbf{y}} + \tilde{C}_{\mathbf{x},\mathbf{y}} \right)_{\Im}; \quad (\text{B.24})$$

$$\text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Im}) = -\frac{1}{2} \left(C_{\mathbf{x},\mathbf{y}} - \tilde{C}_{\mathbf{x},\mathbf{y}} \right)_{\Im}. \quad (\text{B.25})$$

Definition B.2.7 Two complex random vectors $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ and $\mathbf{y} = \mathbf{y}_{\Re} + i\mathbf{y}_{\Im}$ are **uncorrelated** if

$$\text{Cov}(\mathbf{x}_{\Re}, \mathbf{y}_{\Re}) = 0, \quad (\text{B.26})$$

$$\text{Cov}(\mathbf{x}_{\mathfrak{R}}, \mathbf{y}_{\mathfrak{I}}) = 0, \quad (\text{B.27})$$

$$\text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{y}_{\mathfrak{R}}) = 0, \quad (\text{B.28})$$

$$\text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{y}_{\mathfrak{I}}) = 0. \quad (\text{B.29})$$

Lemma B.2.1 *The complex random vectors \mathbf{x} and \mathbf{y} are uncorrelated if and only if $C_{\mathbf{x},\mathbf{y}} = 0$ and $\tilde{C}_{\mathbf{x},\mathbf{y}} = 0$.*

Definition B.2.8 A complex random vector \mathbf{x} is **proper** if and only if $\tilde{C}_{\mathbf{x},\mathbf{x}} = 0$.

The complex random vectors \mathbf{x}_1 and \mathbf{x}_2 are **jointly proper** if the composite complex random vector

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$$

is proper.

From (B.21) we have that

$$\tilde{C}_{\mathbf{x},\mathbf{x}} = 0 \iff \text{Cov}(\mathbf{x}_{\mathfrak{R}}, \mathbf{x}_{\mathfrak{R}}) = \text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{I}}) \text{ and } \text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}}) = -\text{Cov}(\mathbf{x}_{\mathfrak{R}}, \mathbf{x}_{\mathfrak{I}}). \quad (\text{B.30})$$

Note that

$$\text{Cov}(\mathbf{x}_{\mathfrak{R}}, \mathbf{x}_{\mathfrak{I}}) = \text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}})^T. \quad (\text{B.31})$$

Therefore

$$\tilde{C}_{\mathbf{x},\mathbf{x}} = 0 \iff \text{Cov}(\mathbf{x}_{\mathfrak{R}}, \mathbf{x}_{\mathfrak{R}}) = \text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{I}}) \text{ and } \text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}}) = -\text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}})^T. \quad (\text{B.32})$$

The skew-symmetry of $\text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}})$ implies that the main diagonal of $\text{Cov}(\mathbf{x}_{\mathfrak{I}}, \mathbf{x}_{\mathfrak{R}})$ is zero, which means that the real and imaginary part of each component \mathbf{x}_i of \mathbf{x} are uncorrelated. However, $\tilde{C}_{\mathbf{x},\mathbf{x}} = 0$ does not imply that real part of \mathbf{x}_i and the imaginary part of \mathbf{x}_j are uncorrelated for $i \neq j$. If \mathbf{x} is a real random vector, we have that

$\mathbf{x}_{\Im} = 0$. Consequently, $\text{Cov}(\mathbf{x}_{\Im}, \mathbf{x}_{\Im}) = 0$ and $\text{Cov}(\mathbf{x}_{\Im}, \mathbf{x}_{\Re}) = \text{Cov}(\mathbf{x}_{\Re}, \mathbf{x}_{\Im}) = 0$. Therefore a real random vector \mathbf{x} is proper if and only if $\text{Cov}(\mathbf{x}_{\Re}, \mathbf{x}_{\Re}) = \text{Cov}(\mathbf{x}_{\Im}, \mathbf{x}_{\Im}) = 0$, i.e., if and only if \mathbf{x} is a constant with probability 1.

Lemma B.2.2 *Let \mathbf{x} be a proper complex n -dimensional random vector. Then any random vector obtained from \mathbf{x} by a linear or affine transformation, i.e., any random vector $\mathbf{y} = A\mathbf{x} + b$, where $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$ are constant, is also proper.*

Lemma B.2.3 *Let \mathbf{x}_1 and \mathbf{x}_2 be two independent complex random vectors and let \mathbf{x}_2 be proper. Then the linear combination $\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2$, where a_1 and a_2 are complex numbers and $a_1 \neq 0$, is proper, if and only if \mathbf{x}_1 is proper.*

Lemma B.2.4 *Two jointly proper, complex random vectors \mathbf{x}_1 and \mathbf{x}_2 are uncorrelated, if and only if their covariance matrix $C_{\mathbf{x}_1, \mathbf{x}_2} = 0$.*

Definition B.2.9 A complex random vector $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ is a **complex Gaussian random vector** if \mathbf{x}_{\Re} and \mathbf{x}_{\Im} are jointly Gaussian.

Lemma B.2.5 *Two jointly proper Gaussian random vectors \mathbf{x}_1 and \mathbf{x}_2 are independent, if and only if $C_{\mathbf{x}_1, \mathbf{x}_2} = 0$.*

Theorem B.2.1 *Let \mathbf{x} be a proper complex n -dimensional Gaussian random vector with mean $m = \mathbb{E}[\mathbf{x}]$ and nonsingular covariance matrix $C = \mathbb{E}[(\mathbf{x} - m)(\mathbf{x} - m)^\dagger]$. The probability density function of \mathbf{x} is*

$$\begin{aligned} p_{\mathbf{x}}(x_{\Re} + ix_{\Im}) &= p_{\mathbf{x}_{\Re}, \mathbf{x}_{\Im}}(x_{\Re}, x_{\Im}) \\ &= \frac{1}{\pi^n \det(C)} \exp \left\{ -(x - m)^\dagger C^{-1} (x - m) \right\}. \end{aligned} \quad (\text{B.33})$$

Definition B.2.10 The differential entropy of a complex random vector $\mathbf{x} = \mathbf{x}_{\Re} + i\mathbf{x}_{\Im}$ is the joint differential entropy of its real and imaginary parts.

$$h(\mathbf{x}) = h(\mathbf{x}_{\Re}, \mathbf{x}_{\Im}). \quad (\text{B.34})$$

Recall that if \mathbf{x} is a real, continuous, n -dimensional Gaussian random vector with mean m and nonsingular covariance matrix $C_{\mathbf{x}}$, we have that the differential entropy of \mathbf{x} is

$$h(\mathbf{x}) = \frac{1}{2} \log [(2\pi e)^n \det(C_{\mathbf{x}})]. \quad (\text{B.35})$$

Let \mathbf{x} be any real, continuous, n -dimensional random vector with nonsingular correlation matrix $R_{\mathbf{x}} = \mathbf{E}[\mathbf{x}\mathbf{x}^T]$. Then

$$h(\mathbf{x}) \leq \frac{1}{2} \log [(2\pi e)^n \det(R_{\mathbf{x}})], \quad (\text{B.36})$$

with equality if and only if \mathbf{x} is a real Gaussian random vector with zero mean.

Theorem B.2.2 Let \mathbf{x} be a proper, complex, continuous, n -dimensional Gaussian random vector with mean m and nonsingular covariance matrix $C_{\mathbf{x}}$. Then the differential entropy of \mathbf{x} is

$$h(\mathbf{x}) = \log [(\pi e)^n \det(C_{\mathbf{x}})]. \quad (\text{B.37})$$

Theorem B.2.3 Let \mathbf{x} be any complex, continuous, n -dimensional random vector with nonsingular correlation matrix $R_{\mathbf{x}} = \mathbf{E}[\mathbf{x}\mathbf{x}^\dagger]$. Then

$$h(\mathbf{x}) \leq \log [(\pi e)^n \det(R_{\mathbf{x}})], \quad (\text{B.38})$$

with equality if and only if \mathbf{x} is a proper complex Gaussian random vector with zero mean.

It is important to note that Theorem B.2.3 points out that \mathbf{x} must be proper to maximize the differential entropy. Based on our discussion (preceding Lemma B.2.2) of proper real random vectors, we conclude that no real random vector can maximize the differential entropy for a given correlation matrix $R_{\mathbf{x}}$ if complex random vectors are allowed.

In [80], the author uses the terminology “circularly symmetric n -dimensional complex Gaussian random vector” to mean “proper n -dimensional complex Gaussian random vector”. The terminology “circularly symmetric complex Gaussian random vector” also appears in other papers, either explicitly or implicitly.

B.3 Proper Complex Random Processes

Definition B.3.1 $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is a **continuous-time complex random process** if

$$\mathbf{x}(t) = \mathbf{x}_{\Re}(t) + i\mathbf{x}_{\Im}(t), \quad (\text{B.39})$$

where $\mathbf{x}_{\Re}(t) \in \mathbb{R}, t \in \mathcal{T}$ and $\mathbf{x}_{\Im}(t) \in \mathbb{R}, t \in \mathcal{T}$ are a pair of real continuous-time random processes. $\mathbf{x}_{\Re}(t), t \in \mathcal{T}$ is called the real part of $\mathbf{x}(t), t \in \mathcal{T}$ and $\mathbf{x}_{\Im}(t), t \in \mathcal{T}$ is called the imaginary part of $\mathbf{x}(t), t \in \mathcal{T}$. $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is a **discrete-time complex random process** if

$$\mathbf{x}(n) = \mathbf{x}_{\Re}(n) + i\mathbf{x}_{\Im}(n), \quad (\text{B.40})$$

where $\mathbf{x}_{\Re}(n) \in \mathbb{R}, n \in \mathcal{N}$ and $\mathbf{x}_{\Im}(n) \in \mathbb{R}, n \in \mathcal{N}$ are a pair of real discrete-time random processes. $\mathbf{x}_{\Re}(n), n \in \mathcal{N}$ is called the real part of $\mathbf{x}(n), n \in \mathcal{N}$ and $\mathbf{x}_{\Im}(n), n \in \mathcal{N}$ is called the imaginary part of $\mathbf{x}(n), n \in \mathcal{N}$.

Definition B.3.2 Two continuous-time complex random processes $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$

and $\mathbf{y}(t) \in \mathbb{C}, t \in \mathcal{T}$ are independent if real random vector processes

$$\begin{bmatrix} \mathbf{x}_{\Re}(t) \\ \mathbf{x}_{\Im}(t) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{y}_{\Re}(t) \\ \mathbf{y}_{\Im}(t) \end{bmatrix}$$

are independent.

Definition B.3.3 A complex random process is **wide-sense stationary (w.s.s.)** if its real and imaginary parts are jointly wide-sense stationary.

Definition B.3.4 The **mean** of a continuous-time complex random process $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is

$$m_{\mathbf{x}}(t) = \mathbb{E}[\mathbf{x}(t)]. \quad (\text{B.41})$$

The **mean** of a discrete-time complex random process $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is

$$m_{\mathbf{x}}(n) = \mathbb{E}[\mathbf{x}(n)]. \quad (\text{B.42})$$

Definition B.3.5 The **autocorrelation function** of a continuous-time complex random process $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is

$$r_{\mathbf{x}}(t, \tau) = \mathbb{E}[\mathbf{x}(t + \tau)\mathbf{x}^*(t)], \quad (\text{B.43})$$

and the **pseudo-autocorrelation function** of $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is

$$\tilde{r}_{\mathbf{x}}(t, \tau) = \mathbb{E}[\mathbf{x}(t + \tau)\mathbf{x}(t)]. \quad (\text{B.44})$$

The **autocorrelation function** of a discrete-time complex random process $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is

$$r_{\mathbf{x}}(n, k) = \mathbb{E}[\mathbf{x}(n + k)\mathbf{x}^*(n)] \quad (\text{B.45})$$

and the **pseudo-autocorrelation function** of $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is

$$\tilde{r}_{\mathbf{x}}(n, k) = \mathbb{E}[\mathbf{x}(n + k)\mathbf{x}(n)]. \quad (\text{B.46})$$

Definition B.3.6 The **covariance function** of a continuous-time complex random process $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is

$$c_{\mathbf{x}}(t, \tau) = \mathbb{E} [(\mathbf{x}(t + \tau) - m_{\mathbf{x}}(t + \tau)) (\mathbf{x}(t) - m_{\mathbf{x}}(t))^*], \quad (\text{B.47})$$

and the **pseudo-covariance function** of $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ is

$$\tilde{c}_{\mathbf{x}}(t, \tau) = \mathbb{E} [(\mathbf{x}(t + \tau) - m_{\mathbf{x}}(t + \tau)) (\mathbf{x}(t) - m_{\mathbf{x}}(t))]. \quad (\text{B.48})$$

The **covariance function** of a discrete-time complex random process $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is

$$c_{\mathbf{x}}(n, k) = \mathbb{E} [(\mathbf{x}(n + k) - m_{\mathbf{x}}(n + k)) (\mathbf{x}(n) - m_{\mathbf{x}}(n))^*], \quad (\text{B.49})$$

and the **pseudo-covariance function** of $\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$ is

$$\tilde{c}_{\mathbf{x}}(n, k) = \mathbb{E} [(\mathbf{x}(n + k) - m_{\mathbf{x}}(n + k)) (\mathbf{x}(n) - m_{\mathbf{x}}(n))]. \quad (\text{B.50})$$

Lemma B.3.1 *A continuous-time (discrete-time) complex random process $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ ($\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$) is w.s.s. if and only if $m_{\mathbf{x}}(t)$, $r_{\mathbf{x}}(t, \tau)$, $\tilde{r}_{\mathbf{x}}(t, \tau)$ are independent of t ($m_{\mathbf{x}}(n)$, $r_{\mathbf{x}}(n, k)$, $\tilde{r}_{\mathbf{x}}(n, k)$ are independent of n).*

Definition B.3.7 A continuous-time (discrete-time) complex random process $\mathbf{x}(t) \in \mathbb{C}, t \in \mathcal{T}$ ($\mathbf{x}(n) \in \mathbb{C}, n \in \mathcal{N}$) is **proper** if its pseudo-covariance function $\tilde{c}_{\mathbf{x}}(t, \tau) = 0$ for all t and τ ($\tilde{c}_{\mathbf{x}}(n, k) = 0$ for all n and k).

Lemma B.3.2 *Any linear or affine transformation of a proper complex random process is proper. Any linear combination of independent proper complex random processes is proper. Any vector of samples taken from a proper complex random process is proper.*

Proper complex random processes arise in equivalent baseband representations of bandpass communication systems and signals, the equivalent baseband signal representation of a passband signal is a proper complex random process. Interested readers are referred to [77, 90] for details.

APPENDIX C

CAPACITY OF MULTIPLE-ANTENNA COMPLEX GAUSSIAN CHANNELS

We state some important results about the capacities of channels between a single transmitter and a single receiver, where both the transmitter and the receiver may have multiple antennae. The proof for these results can be found in [80], probably with minor modification.

C.1 Multiple-Antenna Complex Gaussian Channels

We consider a transmitter with t transmitting antennae and a receiver with r receiving antennae. Let $\mathbf{x} \in \mathbb{C}^t$ be the transmitted signal, $\mathbf{y} \in \mathbb{C}^r$ be the received signal, and $\mathbf{n} \in \mathbb{C}^r$ be the additive noise at the receiving antennae. We assume that the channel under consideration is frequency non-selective and slow fading [9, 65, 68]. Under such an assumption, we have a **channel matrix** $\mathbf{H} \in \mathbb{C}^{r \times t}$ and a channel model given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (\text{C.1})$$

where \mathbf{H} , \mathbf{x} , \mathbf{n} are independent. The transmitted energy per channel use is constrained by

$$\mathbb{E} [\mathbf{x}^\dagger \mathbf{x}] \leq E, \quad (\text{C.2})$$

Define

$$Q = \mathbb{E} [\mathbf{x}\mathbf{x}^\dagger]. \quad (\text{C.3})$$

Then (C.2) is equivalent to

$$\text{tr}(Q) \leq E. \quad (\text{C.4})$$

Let

$$\mathcal{P} = \{p_{\mathbf{x}} : p_{\mathbf{x}} \text{ is a probability density function of } \mathbf{x} \text{ that satisfies (C.4)}.\}. \quad (\text{C.5})$$

When the realization of \mathbf{H} is known to the receiver, the capacity of the channel given in (C.1) with energy constraint given in (C.4) is [5, 63, 88]

$$C = \sup_{p_{\mathbf{x}} \in \mathcal{P}} \mathbb{E} [\mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H} = H)] \quad (\text{C.6})$$

$$= \sup_{p_{\mathbf{x}} \in \mathcal{P}} \mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H}), \quad (\text{C.7})$$

where the expectation in (C.6) is taken with respect to the cumulative distribution function of \mathbf{H} and $\mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H})$ in (C.7) is the conditional mutual information between \mathbf{x} and \mathbf{y} given \mathbf{H} [23].

When \mathbf{n} is a proper [58] complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$, the channel given in (C.1) is called a **multiple-antenna additive white complex Gaussian noise channel**. We will only consider such a channel for the remainder of this chapter. For any given correlation matrix Q of \mathbf{x} , we know that $\mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H} = H)$ is maximized if and only if \mathbf{x} is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix Q . The corresponding maximized mutual information between \mathbf{x} and \mathbf{y} when $\mathbf{H} = H$ is

$$\mathcal{I}(\mathbf{x}; \mathbf{y} | \mathbf{H} = H) = \log_2 \left(I_r + \frac{H Q H^\dagger}{\sigma^2} \right). \quad (\text{C.8})$$

Therefore, in order to evaluate C in (C.6), we may only consider proper complex Gaussian random vectors that satisfy (C.4), and the evaluation of C in (C.6) is

equivalent to

$$C = \sup_{\text{tr}(Q) \leq E} \mathbb{E} \left[\log_2 \left(I_r + \frac{\mathbf{H}Q\mathbf{H}^\dagger}{\sigma^2} \right) \right]. \quad (\text{C.9})$$

Different cumulative distribution functions of \mathbf{H} result in different channel capacities. For the multiple-antenna additive white complex Gaussian noise channel, we are interested in the capacities under two cases. One is when \mathbf{H} is deterministic. The other is when all entries of \mathbf{H} are independent and identically distributed proper complex Gaussian random variables with mean 0 and variance α^2 , where $\alpha > 0$.

C.2 Capacity of Deterministic Channels

When $\mathbf{H} = H$ is deterministic, the channel given by (C.1) is called a **deterministic channel**. By singular value decomposition [38], any $H \in \mathbb{C}^{r \times t}$ can be written as

$$H = UDV^\dagger,$$

where $U \in \mathbb{C}^{r \times r}$ and $V \in \mathbb{C}^{t \times t}$ are both unitary matrices, and $D \in \mathbb{R}^{r \times t}$ is a non-negative diagonal matrix. We can rewrite (C.1) as

$$\mathbf{y} = UDV^\dagger \mathbf{x} + \mathbf{n}. \quad (\text{C.10})$$

We consider the capacity of the channel given in (C.10) when \mathbf{n} is a proper [58] complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$. Let $\tilde{\mathbf{y}} = U^\dagger \mathbf{y}$, $\tilde{\mathbf{x}} = V^\dagger \mathbf{x}$, and $\tilde{\mathbf{n}} = U^\dagger \mathbf{n}$. Then

$$\tilde{\mathbf{y}} = D\tilde{\mathbf{x}} + \tilde{\mathbf{n}}, \quad (\text{C.11})$$

where $\tilde{\mathbf{n}}$ is a proper complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix

$$\mathbb{E} [\tilde{\mathbf{n}}\tilde{\mathbf{n}}^\dagger] = \mathbb{E} [U^\dagger \mathbf{n}\mathbf{n}^\dagger U] = U^\dagger \mathbb{E} [\mathbf{n}\mathbf{n}^\dagger] U = U^\dagger \sigma^2 I_r U = \sigma^2 I_r. \quad (\text{C.12})$$

Note that the channel given in (C.10) is equivalent to the channel given in (C.11) in terms of mutual information between the transmitted signal and received signal. Since H is of rank at most $\min\{r, t\}$, we have that at most $\min\{r, t\}$ of the singular values of H are non-zero. Let these singular values be $\lambda_i^{1/2}$, $i = 1, \dots, \min\{r, t\}$. We can write (C.11) component-wise to get

$$\tilde{\mathbf{y}}_i = \lambda_i^{1/2} \tilde{\mathbf{x}}_i + \tilde{\mathbf{n}}_i \quad \text{for } 1 \leq i \leq \min\{r, t\}. \quad (\text{C.13})$$

Let $a^+ = \max\{0, a\} \forall a \in \mathbb{R}$. The “water filling” argument (or Kuhn-Kucker Theorem) [23] can be applied here to show that the mutual information $\mathcal{I}(\tilde{\mathbf{x}}; \tilde{\mathbf{y}})$ is maximized when $\tilde{\mathbf{x}}_i, i = 1, \dots, \min\{r, t\}$ are independent and identically distributed proper complex Gaussian random variables with

$$\mathbb{E} [(\Re(\tilde{\mathbf{x}}_i))^2] = \mathbb{E} [(\Im(\tilde{\mathbf{x}}_i))^2] = \frac{1}{2} (\mu - \lambda_i^{-1})^+ \quad \text{for } i = 1, \dots, \min\{r, t\}, \quad (\text{C.14})$$

where $\mu \in \mathbb{R}$ satisfies

$$\sum_{i=1}^{\min\{r, t\}} (\mu - \lambda_i^{-1})^+ = E. \quad (\text{C.15})$$

The maximum mutual information between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ is given by

$$C = \sum_{i=1}^{\min\{r, t\}} \left(\log_2 \left(\frac{\mu \lambda_i}{\sigma^2} \right) \right)^+ \quad \text{bits/(channel use)}. \quad (\text{C.16})$$

For a deterministic channel, we have the following theorem [80].

Theorem C.2.1 (Capacity of Deterministic Channel) *Let $\mathbf{H} = H$ be deterministic and known to both the transmitter and the receiver. Then the capacity of the channel specified in (C.10) is given by (C.16).*

C.3 Capacity of Rayleigh Fading Channels

Let $\mathbf{H} = [\mathbf{h}_{j,k}]$, where $j = 1, \dots, r, k = 1 \dots, t$, be random. We assume that $\mathbf{h}_{j,k}, j = 1, \dots, r, k = 1 \dots, t$ are independent and identically distributed proper complex

Gaussian random variables with mean 0 and variance α^2 , where $\alpha > 0$, i.e.,

$$\mathbb{E} [\mathbf{h}_{j,k} \mathbf{h}_{j,k}^*] = \mathbb{E} [|\mathbf{h}_{j,k}|^2] = \alpha^2 \quad \text{for } j = 1, \dots, r, k = 1 \dots, t.$$

Then $|\mathbf{h}_{j,k}|$ is Rayleigh distributed and $|\mathbf{h}_{j,k}|^2$ is χ^2 -distributed with 2 degrees of freedom. (In this case, $|\mathbf{h}_{j,k}|^2$ is also exponentially distributed.) We assume that the channel is memoryless: for each use of the channel an independent realization of \mathbf{H} is drawn. We also assume that the receiver knows the realization of \mathbf{H} . Under the above assumptions, the channel given by (C.1) is called a **Rayleigh fading channel**. Let

$$\mathbf{Z} = \frac{1}{\alpha} \mathbf{H}. \quad (\text{C.17})$$

Then all entries of \mathbf{Z} are independent and identically distributed proper complex Gaussian random variables with mean 0 and variance 1. We may rewrite (C.1) as

$$\mathbf{y} = \alpha \mathbf{Z} \mathbf{x} + \mathbf{n}. \quad (\text{C.18})$$

We consider the capacity of the Rayleigh fading channel given in (C.18) when \mathbf{n} is a proper [58] complex Gaussian random vector with mean $\mathbf{0}$ and covariance matrix $\sigma^2 I_r$. For such a channel, we have the following theorem [80].

Theorem C.3.1 (Capacity of Rayleigh Fading Channel) *The mutual information between \mathbf{x} and \mathbf{y} of the Rayleigh fading channel specified in (C.18) is maximized if and only if \mathbf{x} is a proper complex Gaussian vector with mean $\mathbf{0}$ and covariance matrix $\frac{E}{t} I_t$. The capacity of such a channel is given by*

$$C = \mathbb{E} \left[\log_2 \left(\det \left(I_r + \frac{\alpha^2 E}{\sigma^2 t} \mathbf{Z} \mathbf{Z}^\dagger \right) \right) \right] \quad \text{bits}/(\text{channel use}). \quad (\text{C.19})$$

The following theorem [80] tells us how to evaluate (C.19).

Theorem C.3.2 (Evaluation of Capacity) (C.19) can be computed by

$$C = \int_0^\infty \log_2 \left(1 + \frac{\lambda \alpha^2 E}{\sigma^2 t} \right) \sum_{j=0}^{N_{min}-1} \frac{j!}{(j + N_{max} - N_{min})!} \cdot [L_j^{N_{max}-N_{min}}(\lambda)]^2 \lambda^{N_{max}-N_{min}} e^{-\lambda} d\lambda \quad \text{bits/(channel use)}, \quad (\text{C.20})$$

where $N_{min} = \min\{r, t\}$, $N_{max} = \max\{r, t\}$, and L_j^i 's are the associated Laguerre polynomials [1].

When $r = t$, we have $N_{min} = N_{max} = r = t$. Then (C.20) gives

$$C = \int_0^\infty \log_2 \left(1 + \frac{\lambda \alpha^2 E}{\sigma^2 t} \right) \sum_{j=0}^{r-1} [L_j(\lambda)]^2 e^{-\lambda} d\lambda \quad \text{bits/(channel use)}, \quad (\text{C.21})$$

where $L_j = L_j^0$ is the Laguerre polynomial of order j . In particular, when $t = r = 1$, we have $N_{min} = N_{max} = 1$ and (C.21) gives

$$C = \int_0^\infty \log_2 \left(1 + \frac{\lambda \alpha^2 E}{\sigma^2} \right) e^{-\lambda} d\lambda \quad \text{bits/(channel use)}. \quad (\text{C.22})$$

When $t = 1$, we have $N_{min} = 1$ and $N_{max} = r$. Then (C.20) gives

$$C = \frac{1}{\Gamma(r)} \int_0^\infty \log_2 \left(1 + \frac{\lambda \alpha^2 E}{\sigma^2} \right) \lambda^{r-1} e^{-\lambda} d\lambda \quad \text{bits/(channel use)}. \quad (\text{C.23})$$

When $r = 1$, we have $N_{min} = 1$ and $N_{max} = t$. Then (C.20) gives

$$C = \frac{1}{\Gamma(t)} \int_0^\infty \log_2 \left(1 + \frac{\lambda \alpha^2 E}{t \sigma^2} \right) \lambda^{t-1} e^{-\lambda} d\lambda \quad \text{bits/(channel use)}. \quad (\text{C.24})$$

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., 1970.
- [2] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung. Network information flow. *IEEE Transactions on Information Theory*, 46:1204–1216, 2000.
- [3] S. M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16, Oct. 1998.
- [4] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley and Sons Inc., 1984.
- [5] R. B. Ash. *Information Theory*. Dover Publications, Inc., 1990.
- [6] C. A. Balanis. *Antenna Theory: Analysis and Design*. John Wiley & Sons Inc, 2nd edition, 1996. Chap. 6.
- [7] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo codes. In *Proceedings of IEEE International Conference on Communications (ICC)*, pages 1064–1070, 1993.
- [8] D. Bertsekas and R. Gallager. *Data Networks*. Englewood Cliffs: Prentice Hall, 1992.
- [9] E. Biglieri, J. Proakis, and S. Shamai (Shitz). Fading channels: Information-theoretic and communications aspects. *IEEE Transactions on Information Theory*, 44:2619–2692, Oct. 1998.
- [10] P. Billingsley. *Probability and Measure*. John Wiley & Sons, Inc., third edition, 1995.
- [11] K. Bogart. *Introductory Combinatorics*. Harcourt Brace Jovanovitch, Orlando, second edition, 1990.
- [12] V. Borich, J.-H. Jong, J. East, and W. E. Stark. Nonlinear effects of power amplification on multicarrier spread spectrum systems. *International Microwave Symposium*, 1:323–26, 1998.

- [13] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [14] A. M. Bruckner, J. B. Bruckner, and B. S. Thomson. *Real Analysis*. Prentice-Hall, Inc., 1997.
- [15] J.-H. Chang and L. Tassiulas. Energy conserving routing in wireless ad-hoc networks. In *Proceedings of IEEE INFOCOM*, pages 22–31, 2000.
- [16] T.-W. Chen and M. Gerla. Global state routing: A new routing scheme for ad-hoc wireless networks. In *Proceedings of IEEE International Conference on Communications (ICC)*, volume 1, pages 171–175, 1998. <http://www.ics.uci.edu/~atm/adhoc/paper-collection/gerla-gsr-icc98.pdf>.
- [17] C.-C. Chiang. Routing in clustered multihop, mobile wireless networks with fading channel. In *Proceedings of IEEE Singapore International Conference on Networks (SICON)*, pages 197–211, Apr. 1997. <http://www.ics.uci.edu/~atm/adhoc/paper-collection/gerla-routing-clustered-sicon97.pdf>.
- [18] A. E. F. Clementi, P. Crescenzi, P. Penna, G. Rossi, and P. Vocca. On the complexity of computing minimum energy consumption broadcast subgraphs. *Proceedings of Symposium on Theoretical Aspects of Computer Science (STACS)*, LNCS(2010):121–131, 2001.
- [19] A. E. F. Clementi, P. Penna, and R. Silvestri. On the power assignment problem in radio networks. *Electronic Colloquium on Computational Complexity*, 2000. <http://citeseer.nj.nec.com/294834.html>.
- [20] M. S. Corson and A. Ephremides. A distributed routing algorithm for mobile wireless networks. *ACM/Baltzer Wireless Networks*, 1(1):61–81, Feb. 1995.
- [21] T. M. Cover. Broadcast channels. *IEEE Transactions on Information Theory*, 18(1):2–14, Jan. 1972.
- [22] T. M. Cover. Comments on broadcast channels. *IEEE Transactions on Information Theory*, 44(6):2524–2530, Oct. 1998.
- [23] T. M. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons Inc., 1991.
- [24] R. Dube et al. Signal stability based adaptive routing for ad hoc mobile networks. *IEEE Personal Communications*, pages 36–45, Feb. 1997. <http://www.cs.umd.edu/projects/mcml/papers/pcm97.ps>.
- [25] A. Ephremides and B. Hajek. Information theory and communication networks: an unconsummated union. *IEEE Transactions on Information Theory*, 44(6):2416–2434, Oct. 1998.

- [26] G. J. Foschini and M. J. Gans. On limits of wireless communications in a fading environment when using multi-element antennas. *Wireless Personal Communications*, 6:311–335, 1998.
- [27] R. G. Gallager. Low-density parity-check codes. *IRE Transactions on Information Theory*, 9(1):666–680, Jan. 1962.
- [28] R. G. Gallager. *Low-Density Parity Check Codes*. MIT Press, Cambridge, MA, 1963.
- [29] R. G. Gallager. Basic limits on protocol information in data communication networks. *IEEE Transactions on Information Theory*, 22(4):385–398, Jul. 1976.
- [30] M. R. Garey and D. S. Johnson. *Computers and Intractability : A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, San Francisco, 1979.
- [31] P. Gupta and P. R. Kumar. The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46(2):388–404, Mar. 2000.
- [32] P. Gupta and P. R. Kumar. Internets in the sky: The capacity of three dimensional wireless networks. *Communications in Information and Systems*, 1(1):33–49, Jan. 2001.
- [33] P. Gupta and P. R. Kumar. Towards an information theory of large networks: An achievable rate region. *IEEE Transactions on Information Theory*, 49(8):1877–1894, Aug. 2003.
- [34] Z. J. Haas and M. R. Pearlman. The performance of query control schemes for the zone routing protocol. *IEEE/ACM Transactions on Networking (TON)*, 9(4):427–438, Aug. 2001.
- [35] S. Hong, S. Kim, and W. E. Stark. Low power parallel-pipeline multiplier design for DSP applications through coefficient optimization. *International Journal of Custom-Chip Design, Simulation, and Testing*, 2001.
- [36] S. Hong and W. E. Stark. Power consumption vs. decoding performance relationship of VLSI decoders for low-energy wireless communication system design. *Proceedings of IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, pages 1593–1596, 1999.
- [37] S. Hong and W. E. Stark. Design and implementation of a low complexity VLSI turbo-code decoder architecture for low energy mobile wireless communications. *Journal of VLSI Signal Processing*, pages 43–57, Feb. 2000.
- [38] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

- [39] A. Iwata, C.-C. Chiang, G. Pei, M. Gerla, and T.-W. Chen. Scalable routing strategies for ad hoc wireless networks. *IEEE Journal on Selected Areas in Communications*, pages 1369–1379, Aug. 1999. <http://www.cs.ucla.edu/NRL/wireless/PAPER/jsac99.ps.gz>.
- [40] S. A. Jafar, S. Vishwanath, and A. J. Goldsmith. Channel capacity and beamforming for multiple transmit and receive antennas with covariance feedback. *Proceedings of IEEE International Conference on Communications (ICC)*, pages 2266–2270, Jun. 2001.
- [41] M. Jiang, J. Li, and Y.C. Tay. Cluster based routing protocol. *IETF Draft*, Aug. 1999. <http://www.ietf.org/internet-drafts/draft-ietf-manet-cbrp-spec-01.txt>.
- [42] M. Joa-Ng and I.-T. Lu. A peer-to-peer zone-based two-level link state routing for mobile ad hoc networks. *IEEE Journal on Selected Areas in Communications*, 17(8):1415–1425, Aug. 1999.
- [43] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc networks. In T. Imielinski and H. Korth, editors, *Mobile Computing*, pages 152–181. Kluwer Academic Publishers, 1996. <http://www.ics.uci.edu/~atm/adhoc/paper-collection/johnson-dsr.pdf>.
- [44] D. B. Johnson and D. A. Maltz. The dynamic source routing protocol for mobile ad hoc networks. *IETF Draft*, Oct. 1999. <http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-03.txt>.
- [45] L. M. Kirousis, E. Kranakis, D. Krizanc, and A. Pelc. Power consumption in packet radio networks. *Proceedings of Symposium on Theoretical Aspects of Computer Science (STACS)*, LNCS(1200):363–374, 1997.
- [46] L. Kleinrock and J. A. Silvester. Optimum transmission radii in packet radio networks or why six is a magic number. In *National Telecommunications Conference*, pages 4.3.1–4.3.5, Dec. 1978.
- [47] R. Koetter and M. Medard. Beyond routing: An algebraic approach to network coding. In *Proceedings of IEEE INFOCOM*, volume 1, pages 122–130, Jun. 2002.
- [48] J. N. Laneman and G. W. Wornell. Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks. *IEEE Transactions on Information Theory*, 49(10):2415–2425, Oct. 2003.
- [49] S.-Y. R. Li, R. W. Yeung, and N. Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49(2):371–381, Feb. 2003.
- [50] D. MacKay. Good error correcting codes based on very sparse matrices. *IEEE Transactions on Information Theory*, 45(2):399–431, Mar. 1999.

- [51] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng. Turbo decoding as an instance of pearl's "belief propagation" algorithm. *IEEE Journal on Selected Areas in Communications*, 16(2):140–152, Feb. 1998.
- [52] P. Misra. Routing protocols for ad hoc mobile wireless networks. 1999. ftp://ftp.netlab.ohio-state.edu/pub/jain/courses/cis788-99/adhoc_routing/index.html.
- [53] J. Moy. Multicast extensions to OSPF. *IETF RFC 1584*, 1994.
- [54] J. Moy. OSPF version 2. *IETF RFC 2328*, Apr. 1998.
- [55] S. Murthy and J.J. Garcia-Luna-Aceves. An efficient routing protocol for wireless networks. *ACM Mobile Networks and Applications Journal*, pages 183–197, Oct. 1996. <http://www.ics.uci.edu/~atm/adhoc/paper-collection/aceves-routing-winet.pdf>.
- [56] A. Narula, M. J. Lopez, M. D. Trott, and G. W. Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE Journal on Selected Areas in Communications*, 16(8):1423–1436, Oct. 1998.
- [57] A. Narula, M. D. Trott, and G. W. Wornell. Performance limits of coded diversity methods for transmitter antenna arrays. *IEEE Transactions on Information Theory*, 45(7):2418–2433, Nov. 1999.
- [58] F. Nesser and J. Massey. Proper complex random process with applications to information theory. *IEEE Transactions on Information Theory*, 39:1293–1302, Jul. 1993.
- [59] OPNET Technologies. *OPNET Modeler Manual*. Online Manual, 2000.
- [60] V. D. Park and M. S. Corson. A highly adaptive distributed routing algorithm for mobile wireless networks. In *Proceedings of IEEE INFOCOM*, volume 3, pages 1405–1413, Apr. 1997. <http://www.ics.uci.edu/atm/adhoc/paper-collection/corson-adaptive-routing-infocom97.pdf>.
- [61] C. E. Perkins and P. Bhagwat. Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers. *ACM SIGCOMM Computer Communication Review*, pages 234–244, Oct. 1994. <http://www.svrloc.org/~charliep/txt/sigcomm94/paper.ps>.
- [62] C. E. Perkins, E. M. Royer, and S. R. Das. Ad hoc on-demand distance vector routing. *IETF Draft*, Oct. 1999.
- [63] M. S. Pinsker. *Information and Information Stability of Random Variables and Processes*. Holden-Day, Inc., 1964.
- [64] D. Pollard. *A User's Guide to Measure Theoretic Probability*. Cambridge University Press, 2002.

- [65] J. G. Proakis. *Digital Communications*. McGraw-Hill Higher Education, 2000.
- [66] J. M. M. Rabaey. *Digital Integrated Circuits: A Design Perspective*. Prentice Hall PTR, 1995.
- [67] R. Ramanathan and R. Rosales-Hain. Topology control of multihop wireless networks using transmitpower adjustment. In *Proceedings of IEEE INFOCOM*, volume 2, pages 404–413, 2000.
- [68] T. Rappaport. *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice Hall, second edition, 2002.
- [69] V. Rodoplu and T. Meng. Minimum energy mobile wireless networks. *IEEE Journal on Selected Areas in Communications*, 17(8):1333–1344, Aug. 1999.
- [70] H. E. Romeijn and R. L. Smith. Simulated annealing for constrained global optimization. *Journal of Global Optimization*, 5:101–26, 1994.
- [71] E. M. Royer and C.-K. Toh. A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications*, 6(2):46–55, Apr. 1999. <http://users.ece.gatech.edu/~cktoh/royer.html>.
- [72] K. Scott and N. Bambos. Routing and channel assignment for low power transmission in PCS. In *IEEE ICUPC*, pages 498–502, 1996.
- [73] C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–423, 623–656, Jul.-Oct. 1948.
- [74] S. Singh, M. Woo, and C. S. Raghavendra. Power-aware routing in mobile ad hoc networks. In *Proceedings of ACM/IEEE MOBICOM*, pages 181–190, Oct. 1998.
- [75] D. Slepian and J. K. Wolf. A coding theorem for multiple access channels with correlated sources. *Bell Systems Technical Journal*, 52:1037–1076, 1973.
- [76] A. Stamoulis and N. Al-Dhahir. Impact of space-time block codes on 802.11 network throughput. *IEEE Transactions on Wireless Communications*, 2(5):1029–1039, Sep. 2003.
- [77] W. E. Stark. *Digital Communication Theory*. University of Michigan Course Pack, 1999.
- [78] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time block codes from orthogonal designs. *IEEE Transactions on Information Theory*, 45:1456–1466, Jul. 1999.
- [79] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, Mar. 1998.

- [80] Í. E. Telatar. Capacity of multi-antenna Gaussian channels. Technical report, AT&T-Bell Labs, Jun. 1995.
- [81] Í. E. Telatar and R. G. Gallager. Combining queueing theory with information theory for multiaccess. *IEEE Journal on Selected Areas in Communications*, 13:963–969, 1995.
- [82] C.-K. Toh. A novel distributed routing protocol to support ad hoc mobile computing. In *Proceedings of IEEE 15th Annual International Phoenix Conference Computing and Communications*, pages 480–486, Mar. 1996. <http://www.ics.uci.edu/~atm/adhoc/paper-collection/toh-distributed-routing-ipccc96.pdf>.
- [83] C.-K. Toh. Long-lived ad-hoc routing based on the concept of associativity. *IETF Draft*, 1999.
- [84] D. Tse and S. Hanly. Multi-access fading channels: Part I: Polymatroid structure, optimal resource allocation and throughput capacities. *IEEE Transactions on Information Theory*, 44(7):2796–2815, Nov. 1998.
- [85] D. Tse and S. Hanly. Multi-access fading channels: Part II: Delay-limited capacities. *IEEE Transactions on Information Theory*, 44(7):2816–2831, Nov. 1998.
- [86] E. Visotsky and U. Madhow. Space-time transmit precoding with imperfect feedback. *IEEE Transactions on Information Theory*, 47(6):2632–2639, Sep. 2001.
- [87] J. E. Wieselthier, G. D. Nguyen, and A. Ephremides. On the construction of energy-efficient broadcast and multicast trees in wireless networks. In *Proceedings of IEEE INFOCOM*, volume 2, pages 585–594, 2000.
- [88] J. Wolfowitz. *Coding Theorems of Information Theory*. Berlin:Springer-Verlag, third edition, 1978.
- [89] A. Worthen, S. Hong, R. Gupta, and W. E. Stark. Performance optimization of vlsi transceivers for low-energy communications systems. In *Proceedings of IEEE Military Communications Conference (MILCOM)*, volume 2, pages 1434–1438, Oct. 1999.
- [90] M. Zakai. The representation of narrow-band processes. *IRE Transactions on Information Theory*, 8:323–325, July 1962.
- [91] D. Zwillinger and S. Kokoska. *CRC Standard Probability and Statistics Tables and Formulae*. Chapman & Hall/CRC, 2000.