

Cross-Layer Design and Optimization of Short Range  
Wireless Networks

by

Tianqi Wang

Submitted in Partial Fulfillment

of the

Requirements for the Degree

Doctor of Philosophy

Supervised by

Professor Wendi B. Heinzelman

Professor Alireza Seyedi

Department of Electrical and Computer Engineering

Arts, Sciences and Engineering

Edmund A. Hajim School of Engineering and Applied Sciences

University of Rochester

Rochester, New York

2011

# Curriculum Vitae

The author was born in Weihai City, Shan dong, China on January 15th, 1983. He attended Beijing University of Posts and Telecommunications from 2000 to 2004 and graduated with a Bachelor of Engineering degree in 2004. After that, he attended Memorial University of Newfoundland from 2004 to 2007 and graduated with a Master of Engineering degree in 2007. He came to the University of Rochester in the Fall of 2007 and began graduate studies in Electrical and Computer Engineering. He pursued his research in the energy efficient wireless networks under the direction of Professor Wendi Heinzelman and Professor Alireza Seyedi, and received the Master of Science degree from the University of Rochester in 2010.

## List of Publications:

- Tianqi Wang, Wendi Heinzelman, Alireza Seyedi, and Azadeh Vosoughi, “Maximizing Gathered Samples in Wireless Sensor Networks with Slepian-Wolf Coding,” to appear, IEEE Transactions on Wireless Communications.
- Tianqi Wang, Wendi Heinzelman, and Alireza Seyedi, “Link Energy Minimization for Wireless Sensor Networks,” to appear, Ad Hoc Networks journal.
- Matthew Holland, Tianqi Wang, Bulent Tavli, Alireza Seyedi and Wendi Heinzelman, “Optimizing Physical Layer Parameters for Wireless Sensor Networks,” ACM Transactions on Sensor Networks, Vol.7, No.4, Nov. 2011.

- Tianqi Wang, Wendi Heinzelman, and Alireza Seyedi, “Energy Minimization for IR-UWB-based Wireless Networks,” *IEEE Transactions on Wireless Communications*, Vol. 9, Issue 9, Sept. 2010.
- Tianqi Wang, Cheng Li, and Hsiao-Hwa Chen, “On EM-Based Adaptive Frequency Domain Estimation of Doppler Shifts with CRLB Analysis for CDMA Systems,” *IEEE Transactions on Communications*, Vol. PP, Issue 99, Nov. 2011.
- Tianqi Wang, Cheng Li, and Hsiao-Hwa Chen, “An Iterative Expectation-Maximization Algorithm Based Joint Estimation Approach for CDMA/OFDM Composite Radios,” *IEEE Transactions on Wireless Communications*, Vol. 7, Issue 8, Aug. 2008.

# Acknowledgments

First and foremost, I would like to offer my uttermost gratitude to my advisors, Dr. Wendi Heinzelman and Dr. Alireza Seyedi, who have supported me throughout my thesis with their patience and knowledge. This dissertation would not have been possible without the guidance and invaluable insights from my supervisors. It has been a privilege for me to have this opportunity to work with Dr. Wendi Heinzelman and Dr. Alireza Seyedi. Their sincerity, encouragement and personality always inspired and motivated me as I worked towards my doctoral degree.

It is a great pleasure to express my sincere gratitude to my thesis committee members, Professors Azadeh Vosoughi and Kai Shen, for their generous offers of time, efforts, and valuable advice.

I am indebted to many of my brilliant colleagues in the Wireless Communications and Networking Group who have supported me during my studies. It is an honor for me to have the opportunity to work with my exceptional lab mates. I will always cherish the memories I have shared with you.

Last but not least, I would like to thank my parents, Shunfu Wang and Fengyuan Li, and my wife, Shenqiu Zhang, for their love, selfless support and encouragement to finish my studies. Your love keeps me warm in the long winter and makes me live my life with no regrets.

# Abstract

Short-range wireless networks, such as wireless sensor networks, have become an integral part of our modern lives and have been broadly applied in many fields such as industry, military and research to facilitate the gathering and distribution of information. Compared with traditional wireless networks, such as cellular networks, short-range wireless networks have the following unique characteristics. (i) *Dense deployment*: the network devices are often densely deployed to achieve better monitoring of the environment. (ii) *Circuit power consumption*: due to the short communication distances, the network devices communicate with each other using low transmit power that is comparable to the devices' circuit power consumption. Thus, circuit power consumption is a major contributor to the energy drain of the network devices. (iii) *Battery powered*: the network devices are usually battery powered and may be deployed in remote areas. Thus, it is difficult or even impossible to replace the energy supplies of many of the network devices in a short-range wireless network. Therefore, maximizing the energy efficiency of short-range wireless networks is of paramount importance.

In this dissertation, I explore the cross-layer design principle to improve the energy efficiency of energy constrained short-range wireless networks, while fully considering their unique characteristics as outlined above. In order to maximize energy efficiency, my research focuses on the cross-layer optimization of the physical layer, the data link layer, the multiple access layer, the network layer, and the application layer. In this dissertation, I (i) develop an energy efficient cross-layer design of the physical layer and the data link layer in a typical narrowband system over an additive white Gaussian noise

(AWGN) channel and a Rayleigh fading channel, as well as in a typical Impulse Radio Ultra Wideband (IR-UWB) system over a frequency selective channel; (ii) optimize the energy efficiency of a clustered wireless network by choosing the optimal transmit power, selecting the optimal cluster head, and deciding whether or not to use multi-hop routing within a cluster; and (iii) optimize the energy efficiency of a short-range wireless network with distributed source coding (DSC) and adaptive transmission, as well as with DSC over Gaussian multiple access channels.

Compared with existing work in the literature, I make unique contributions in this dissertation in the following aspects. First, the unique characteristics of short-range wireless networks, such as dense deployment and circuit power consumption, are considered in all of my cross-layer optimizations. Second, I focus on achieving a balance between cost and performance during the development of the cross-layer optimization schemes, due to the limited computational capacity of the network devices in short-range wireless networks. Third, throughout this dissertation, I develop universal optimal solutions that are highly parameterized and directly applicable in general scenarios. My research results in a large improvement in the energy efficiency of devices for short-range wireless networks.

# Table of Contents

<b>Curriculum Vitae</b>	<b>ii</b>
<b>Acknowledgments</b>	<b>iv</b>
<b>Abstract</b>	<b>v</b>
<b>List of Tables</b>	<b>xi</b>
<b>List of Figures</b>	<b>xii</b>
<b>Foreword</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Wireless Sensor Networks . . . . .	4
1.2 Motivation . . . . .	5
1.3 Thesis Statement . . . . .	6
1.4 Thesis Organization . . . . .	7
<b>2 Related Work</b>	<b>9</b>
2.1 Cross-Layer Optimization . . . . .	9
2.2 Energy Efficient Designs for Short-Range Wireless Networks . . . . .	12

<b>3</b>	<b>Optimizing Physical Layer Parameters for Wireless Networks</b>	<b>32</b>
3.1	Introduction . . . . .	32
3.2	Channel and Physical Layer Model . . . . .	35
3.3	Optimizing Physical Layer Parameters . . . . .	44
3.4	Summary . . . . .	57
<b>4</b>	<b>Joint Optimization of Physical Layer and Link Layer in Narrowband Communication Systems</b>	<b>59</b>
4.1	Introduction . . . . .	60
4.2	System and Signal Model . . . . .	62
4.3	Unconstrained Energy Minimization . . . . .	68
4.4	Energy Minimization with a Fixed Average Power Constraint . . . . .	74
4.5	Energy Minimization with a Fixed Average Rate Constraint . . . . .	78
4.6	Numerical Results . . . . .	79
4.7	Summary . . . . .	94
<b>5</b>	<b>Joint Optimization of Physical Layer and Link Layer in Impulse Radio Ultra-Wideband (IR-UWB) Communication Systems</b>	<b>96</b>
5.1	Introduction . . . . .	97
5.2	System and Channel Models . . . . .	99
5.3	Link Energy Minimization . . . . .	105
5.4	Numerical Results . . . . .	114
5.5	Summary . . . . .	123
<b>6</b>	<b>Maximizing Data Gathering in Clustered Short-Range Wireless Networks</b>	<b>126</b>
6.1	Introduction . . . . .	127

6.2	Transmit Power Optimization . . . . .	129
6.3	Cluster Head Selection . . . . .	137
6.4	To Hop or Not to Hop . . . . .	139
6.5	Results . . . . .	143
6.6	Summary . . . . .	150
<b>7</b>	<b>Maximizing Gathered Samples in Wireless Networks with Slepian-Wolf Coding</b>	<b>151</b>
7.1	Introduction . . . . .	152
7.2	Problem Formulation . . . . .	155
7.3	maximizing collected samples without joint decoder energy constraint .	157
7.4	maximizing collected samples with joint decoder energy constraint . . .	160
7.5	The critical energy of the joint decoder . . . . .	164
7.6	Impact of SW Encoding and Decoding Energy Costs . . . . .	166
7.7	Extensions to Multi-hop Scenarios . . . . .	168
7.8	Numerical Results . . . . .	171
7.9	Conclusions and Discussion . . . . .	178
<b>8</b>	<b>Optimal Rate Allocation for Distributed Source Coding over Gaussian Multiple Access Channels</b>	<b>181</b>
8.1	Introduction . . . . .	182
8.2	Problem Formulation . . . . .	185
8.3	Performance Evaluation for MA Schemes with SW coding . . . . .	188
8.4	Optimal Rate Allocation for MA Schemes with SW coding . . . . .	191
8.5	Optimal MA and Rate Allocation: Two-Source Node Case . . . . .	196
8.6	Numerical Results . . . . .	201
8.7	Summary . . . . .	206

<b>9</b>	<b>Conclusions and Future Directions</b>	<b>207</b>
9.1	Contributions . . . . .	207
9.2	Future Directions . . . . .	212
<b>A</b>	<b>Intersection Points of CDMA and TDMA Rate Regions Boundaries</b>	<b>214</b>
A.1	Finding the points $[r_1^{(1)}, r_2^{(1)}]$ and $[r_1^{(4)}, r_2^{(4)}]$ . . . . .	214
A.2	Finding the points $[r_1^{(2)}, r_2^{(2)}]$ and $[r_1^{(3)}, r_2^{(3)}]$ . . . . .	215
	<b>Bibliography</b>	<b>217</b>

## List of Tables

3.1	Power Consumption Values . . . . .	39
3.2	Table of Symbol Error Formulas . . . . .	42
3.3	Parameters Used in the Models. . . . .	45
3.4	Power Consumption of Motes . . . . .	53
3.5	Percent increase in <i>ESBM</i> by using suboptimal modulation schemes. Data used in figure 3.11(c) . . . . .	56
4.1	Notations . . . . .	63
4.2	Packet Structure Parameters . . . . .	64
4.3	Power Consumption Values . . . . .	66
5.1	Overall Optimum Configurations . . . . .	124
5.2	Overall Optimum Configurations . . . . .	125
7.1	The proposed optimal rate allocation with coding energy cost . . . . .	168
7.2	The average number of iterations $\bar{n}$ w.r.t. $\epsilon$ . . . . .	179
8.1	The optimal MA scheme for $N = 2$ . . . . .	198

## List of Figures

2.1	The seven layers of the OSI model. . . . .	10
2.2	The sensor network protocol stack model [1]. . . . .	11
2.3	A system model with adaptive communication. . . . .	14
2.4	The transmitter and receiver structure in an IR-UWB system. . . . .	21
2.5	An illustration of a multi-hop scheme versus a single hop scheme. . . . .	26
2.6	An illustration of Slepian-Wolf source coding region (2 sources). . . . .	29
3.1	Three examples of a linear wireless network. Network 1 has a short hop distance, Network 2 has a long hop distance, and Network 3 has the optimal hop distance. . . . .	34
3.2	The packet structure used in AWGN channels. . . . .	36
3.3	A typical transmitter structure using linear modulation. . . . .	38
3.4	A typical receiver structure using linear demodulation. . . . .	38
3.5	The energy per successfully received bit ( $ESB$ ) as a function of the transmit energy $E_{s,TX}$ . This plot shows a clear minimum and thus the optimal transmit energy. These results assume a fixed distance $d = 10m$ , BPSK modulation and fixed channel noise. . . . .	41
3.6	The packet structure with header and <i>interleaved</i> training sequence in block Rayleigh fading channels. . . . .	42

3.7	$ESB$ and $E_{s, TX}$ at a fixed distance . . . . .	46
3.8	Optimum distance . . . . .	47
3.9	Figures relating optimum distance . . . . .	48
3.10	Mote Probability of Success . . . . .	49
3.11	Optimum $ESB$ vs. optimum hop distance for various relay distances . .	50
3.12	Packet Size . . . . .	52
3.13	Amplifier Efficiency . . . . .	54
3.14	Gain of finding optimum . . . . .	55
3.15	Optimum $ESB$ vs. optimum hop distance for various relay distances . .	57
4.1	Packet structure. . . . .	64
4.2	A typical transmitter structure using linear modulation. . . . .	65
4.3	A typical receiver structure using linear demodulation. . . . .	65
4.4	The transmission and reception of one packet using $m$ total transmissions. 67	
4.5	The minimum $E_b$ of different modulation schemes vs. distance compared with the theoretical lower bound of $E_b$ . . . . .	80
4.6	Optimized target bit error probability vs. transmission distance ( $L_L = L_L^*$ ). . . . .	81
4.7	Optimized packet length vs. transmission distance ( $P_b = P_b^*$ ). . . . .	82
4.8	The optimized transmit powers vs. transmission distance (theoretical and practical). . . . .	83
4.9	The optimized information rate vs. transmission distance. . . . .	83
4.10	The information rate at distance $d = 1$ m. . . . .	84
4.11	The information rate at distance $d = 40$ m. . . . .	85
4.12	The information rate at distance $d = 70$ m. . . . .	85

4.13	The information rate, $M^*$ , $P_t^*$ , and $\Phi^*$ of $M$ -QAM at different transmission distances. . . . .	86
4.14	The influence of the average power on $(\Phi^*, P_t^*, E_b^*)$ from (4.32). . . . .	87
4.15	The influence of the average power on $R^*$ from (4.32). . . . .	88
4.16	Range of efficient vs. inefficient zones for average power versus distance from (4.32). . . . .	88
4.17	Information rate as a function of average power from (4.32). . . . .	89
4.18	$M^*$ , $P_t^*$ , and $\Phi^*$ for $M$ -QAM at different transmission distances ( $ R' - 20\text{Kbps}  < 100\text{bps}$ ). . . . .	90
4.19	The information rate and $\bar{P}^*$ of $M$ -QAM at different transmission distances (fixed average power compared with fixed average rate). . . . .	91
4.20	The $E_b^*$ of $M$ -QAM at different transmission distances. . . . .	92
4.21	The $L_L^*$ of $M$ -QAM at different transmission distances for the unconstrained, fixed-rate-constrained and fixed-power-constrained cases. . . . .	93
5.1	The transmitter and receiver structure in an IR-UWB system. . . . .	100
5.2	Data packet structure. . . . .	102
5.3	The transmission and reception of one packet using $N$ total transmissions. . . . .	108
5.4	The optimum target BEP versus distance ( $N_p = 3$ ). . . . .	115
5.5	The optimum packet length versus distance ( $N_p = 3$ ). . . . .	116
5.6	The optimum number of RAKE fingers versus distance ( $N_p = 3$ ). . . . .	116
5.7	The minimum energy consumption per information bit versus distance ( $N_p = 3$ ). . . . .	117
5.8	The overall minimum energy consumption per information bit and corresponding modulation/decoding/detection schemes ( $N_p = 3$ ). . . . .	117

5.9	The overall minimum energy consumption per information bit for different repetition coding parameters with optimized $L_L$ , $M$ , and modulation/repetition coding/detection schemes. . . . .	118
5.10	The $M^*$ and $L_L^*$ with optimized modulation/combining/detection schemes for different $N_p$ . . . . .	118
6.1	A typical cluster topology ( $N_d = 5$ ). . . . .	129
6.2	Average $D^*$ as the CH energy increases. . . . .	132
6.3	The transmit power optimization algorithm flowchart. . . . .	137
6.4	A 3-node cluster with a potential multi-hop route. . . . .	140
6.5	The maximum data gathered by clusters operating in the CH-constrained region. . . . .	144
6.6	The maximum data gathered by clusters with different cluster head selection strategies. . . . .	145
6.7	Performance comparison of optimization with/without the possible use of a multi-hop. . . . .	147
6.8	Performance comparison of the optimized and unoptimized cluster. . . . .	149
7.1	A data-gathering WSN with $N$ source nodes where each node employs SW coding for distributed compression and the joint decoder reconstructs all the collected samples losslessly. . . . .	155
7.2	Application of the proposed algorithms in a data-gathering wireless network with multi-hop routes. . . . .	169
7.3	Performance comparison when the network is not joint decoder limited: average $S^*$ versus $h$ for $N = 2$ . . . . .	173
7.4	Performance comparison when the network is joint decoder limited: average $S^*$ versus $h$ for $N = 2$ . . . . .	174

7.5	Performance comparison: average $S^*$ versus $E_{max}$ for $N = 10$ . . . . .	176
7.6	Performance of the proposed iterative algorithm: $S_n^*$ approaches $S_E^*$ as $n$ increases for $N = 10$ . . . . .	178
8.1	The rate regions of TDMA, CDMA, and FDMA for $N = 2$ without (a) and with (b) the consideration of circuit power consumption. . . . .	183
8.2	An illustration of a WSN with SW coding and a Gaussian multiple access channel. . . . .	186
8.3	A visualization of Proposition 1 for $N = 2$ . . . . .	190
8.4	Region selection of TDMA and CDMA for $N = 2$ . . . . .	197
8.5	Rate allocation of CDMA for $N = 2$ . . . . .	198
8.6	Rate allocation of TDMA for $N = 2$ . . . . .	200
8.7	The flowchart of the optimal rate allocation for CDMA and TDMA with $N = 2$ . . . . .	202
8.8	$S^*$ versus $d_1$ for TDMA and $N = 3$ . . . . .	202
8.9	$S^*$ versus $\delta$ for TDMA and $N = 2$ . . . . .	204
8.10	$S^*$ of the “ $R_i^*, \theta_i^*$ ” approach for CDMA and TDMA in terms of $P_1$ and $P_2$ for $N = 3, P_3 = 110$ mW. . . . .	205

# Foreword

The following chapters of this dissertation were jointly produced. My participation and contribution to the research is as follows.

Chapter 3 of my dissertation was co-authored with Matthew Holland, Dr. Bulent Tavli, Dr. Wendi Heinzelman and Dr. Alireza Seyed. My contributions to this chapter are the analysis of the block Rayleigh fading scenario and the detailed analysis of the power models of transceivers. This work has been published in the ACM Transactions on Sensor Networks in 2011.

I conducted all of the research work from Chapter 4 to Chapter 8. I am the primary author of Chapter 4. I collaborated with Dr. Wendi Heinzelman and Dr. Alireza Seyed. This work has been accepted to be published in the Ad Hoc Networks journal in 2012. Part of this work has been published in the proceedings of the IEEE Allerton Conference on Communication, Control, and Computing, 2008.

I am the primary author of Chapter 5. I collaborated with Dr. Wendi Heinzelman and Dr. Alireza Seyed. This work has been published in the IEEE Transactions on Wireless Communications in 2010. Part of this work has been published in the proceedings of the IEEE International Conference on Communications (ICC), 2009.

I am the primary author of Chapter 6. I collaborated with Dr. Wendi Heinzelman and Dr. Alireza Seyed. Part of this work has been published in the proceedings of the IEEE GlobeCom, 2010.

I am the primary author of Chapter 7. I collaborated with Dr. Wendi Heinzelman, Dr. Adadeh Vosoughi, and Dr. Alireza Seyed. This work has been accepted to be

published in the IEEE Transactions on Wireless Communications in 2012. Part of this work has been published in the proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2010.

I am the primary author of Chapter 8. I collaborated with Dr. Wendi Heinzelman, Dr. Adadeh Vosoughi, and Dr. Alireza Seyedi. This work is currently under submission. Part of this work has been published in the proceedings of the IEEE International Conference on Communications (ICC), 2011.

# 1 Introduction

Recently, densely deployed short-range wireless networks (SRWNs), in particular wireless sensor networks (WSNs), have attracted attention in the scientific community. These types of wireless networks have been broadly applied in many environments, providing support for applications ranging from security and surveillance to monitoring and health care.

The popularity of densely deployed wireless networks has been propelled by their unique features. First, the network devices of short-range wireless networks, namely *nodes*, are of very low cost. This makes it possible to densely deploy the network devices to greatly improve the network robustness. Second, due to the self organization feature that is often available in modern short-range wireless networks, the nodes can be deployed in inhospitable environments for the collection of data of interest. For instance, sensors in WSNs can be used to detect survivors in a disaster site where it is too dangerous to send in a search team. There are many other benefits of the application of SRWNs, such as the ability to provide full monitoring coverage of an area. While the value of SRWNs has been fully realized, the research community is still addressing the design challenges of SRWNs to facilitate the design and implementation of new applications for SRWNs.

The network devices used in densely deployed SRWNs are usually small in size

and close in distance. Compared with the devices used in traditional wireless networks, the SRWN nodes are usually much more limited in power, computational ability, and energy. This imposes great challenges on the design and optimization of short-range wireless networks. Due to the nodes' limited energy storage and the difficulties in replacing batteries, the efficient expenditure of energy in the nodes is of particular interest.

One of the major design issues for a SRWN is the consideration of circuit power consumption. Unlike with long range wireless networks where the circuit power consumption is negligible compared with the transmit power, in SRWNs it is crucial to consider the circuit power consumption in the design of the communication protocols. Therefore, the balance between the transmit power and the circuit power must be carefully evaluated. The objective of my dissertation research is to improve the energy efficiency of short-range energy constrained wireless networks through cross-layer optimization techniques that consider these unique characteristics of SRWNs.

## **1.1 Wireless Sensor Networks**

One particular type of SRWN is a wireless sensor network. A wireless sensor network (WSN) consists of spatially distributed autonomous sensors that cooperatively monitor physical or environmental conditions, such as temperature, sound, vibration, pressure, motion or pollutants. In the past ten years, there has been increasing interest in wireless sensor networks. This interest has been fueled, in part, by the availability of small, low cost sensor nodes (motes), enabling the deployment of large-scale networks for a variety of sensing applications [1]. The following are just a few examples of applications that can benefit from wireless sensor networks.

- Agricultural monitoring - evaluation of soil nutrients and moisture.
- Home automation - temperature or movement detection.

- Industrial monitoring - sensing faults in machinery or surveillance of property.
- Wildlife/environmental survey - cataloging animal movements and the status of forested areas.
- Battlefield surveillance - rapidly deployable systems to send situational awareness data to a virtual command center.
- Medical monitoring - monitoring the condition of a patient.

The broad applications of wireless sensor networks have greatly facilitated science exploration. For instance, battery powered sensor networks can be deployed in the Amazon rainforest, the north pole, and even the bottom of the ocean to obtain data that has been inaccessible by traditional means. On the other hand, sensor networks are becoming a part of people's daily lives. Wireless home security systems and temperature control systems make life safe and more comfortable.

Energy efficiency is a paramount design issue for wireless sensor networks. Wireless sensor networks are extremely resource-limited, especially in terms of their energy supply. Sensor nodes are usually powered by batteries, which impose strict constraints on not only the available energy but also the output power. In many wireless sensor networks, the number and location of nodes make recharging or replacing the batteries infeasible at worst and inconvenient at best. For this reason, energy consumption is a universal design issue for wireless sensor networks.

## **1.2 Motivation**

The energy efficiency of short-range energy constrained wireless networks is of paramount importance and has attracted many research efforts [2]. Past research has shown the benefit of cross-layer optimization to reduce the energy consumption of wireless devices [3–7]. In my thesis, I extend this idea of cross-layer optimization to improve the

design and performance of short-range wireless networks. Specifically, in my thesis, I develop several techniques to improve the energy efficiency of short-range wireless networks while taking account of their unique characteristics, i.e., low transmit power, limited energy supply, significant circuit power consumption, and low computational ability.

### 1.3 Thesis Statement

I focus my research on improving the energy efficiency of short-range wireless networks through cross-layer optimization. The optimizations in this thesis are conducted jointly at the physical layer, the data link layer, the multiple access layer, the network layer, and the application layer. The main contributions of this dissertation are highlighted as follows.

- The introduction of a metric, i.e., energy consumption per information bit, to evaluate the energy efficiency of a communication link in short-range wireless networks. This metric incorporates the influences of the circuit power consumption of the nodes, the transmit power consumption, packetization overhead, coding overhead, and possible retransmissions.
- The analysis of the energy-optimal relay distance that minimizes the energy consumption per information bit.
- The analysis of the optimum constellation size, packet length, and duty cycle that minimize the energy consumption per information bit over an additive white Gaussian noise (AWGN) channel.
- The analysis of the optimum packet length, the optimum number of RAKE fingers, the optimum modulation and coding schemes that minimize the energy consumption per information bit in a typical Impulse Radio Ultra Wideband (IR-UWB) system over a frequency selective channel.

- A comprehensive analysis of improving the energy efficiency of a clustered wireless network by choosing the optimal transmit power, selecting the optimal cluster head, and deciding whether or not to use multi-hop routing within a cluster.
- An energy efficient design of wireless networks jointly considering distributed source coding, adaptive transmission, and clustering topology. By applying the philosophy of cross-layer optimization over the physical and application layers, I propose a joint optimization of transmit time durations, transmit powers, and Slepian-Wolf (SW) coding rates of the source nodes. I comprehensively evaluate the impacts of the communication environment and the residual energy of the nodes on the Slepian-Wolf coding rates, and I derive closed form solutions of the optimal transmit time durations, transmit powers, and Slepian-Wolf coding rates that maximize the samples gathered at the fusion center.
- Derivation of a low complexity joint optimal rate allocation of the SW coding rates and the information rates when using a multiple access channel in an average-power constrained short-range wireless network. I examine the optimality of three orthogonal multiple access channel schemes, namely code division multiple access (CDMA), frequency division multiple access (FDMA), and time division multiple access (TDMA), from an information theoretical perspective, and I propose an algorithm to determine how to choose the joint optimal SW coding rates and the information rates that maximize the sample rate of the network.

## 1.4 Thesis Organization

In this thesis, different aspects of the design and optimization of energy efficient short-range wireless networks are discussed. Chapter 2 summarizes related work in the area

of cross-layer optimization of wireless networks and introduces the background of the concepts and techniques used in this dissertation.

Chapter 3 describes an optimization approach to reducing energy dissipation at the physical layer, by finding the optimal transmit (relay) distance and transmit power for a given modulation scheme and a given channel model, in order to maximize network lifetime.

Chapter 4 describes a cross-layer optimization scheme for the link layer and the physical layer for a narrowband communication link in SRWNs. Correspondingly, a cross-layer optimization at the link layer and physical layer for a typical impulse radio ultra wideband (IR-UWB) radio is developed in Chapter 5.

Clustering network topologies are taken into account in the optimization model in Chapter 6. A joint optimization at the physical layer, i.e., adaptive transmit power, and network layer, i.e., cluster head selection and multi-hop selection, is proposed to maximize the data gathering capacity of a clustered wireless network.

Distributed source coding is added into the optimization model in Chapter 7 to improve the information gathering capacity of a short-range wireless network, which is measured in the number of samples gathered at the fusion center. Chapter 8 proposes a joint optimal rate allocation of Slepian-Wolf coding rates over multiple access channels in a SRWN with the consideration of circuit power and average power constraints. Chapter 9 concludes this dissertation.

## 2 Related Work

In the past decades, there has been much research aimed at improving the energy efficiency of wireless networks. Thus, it is essential to have a thorough understanding of current enabling technologies in the energy efficient design of wireless networks, upon which further developments and contribution can be built. In this chapter, I summarize some important concepts and techniques from the literature on the principle of cross-layer optimization and energy efficient design for wireless networks.

### 2.1 Cross-Layer Optimization

Optimization techniques, including linear programming, convex programming, geometric programming, and dynamic programming, have been extensively used in the field of wireless networks to improve the network performance. Before performing any optimization, a clear understanding of a typical communication system is necessary. A traditional 7-layer Open System Interconnection (OSI) protocol stack is shown in Fig. 2.1 [8]. The physical layer (PHY) addresses the issue of establishing a physical link between communication ends, including transmission, reception, modulation and demodulation. The data link layer ensures the reliability of the established physical link and coordinates the resources between different communication links. The network layer is in charge of establishing, maintaining, and terminating end-to-end net-

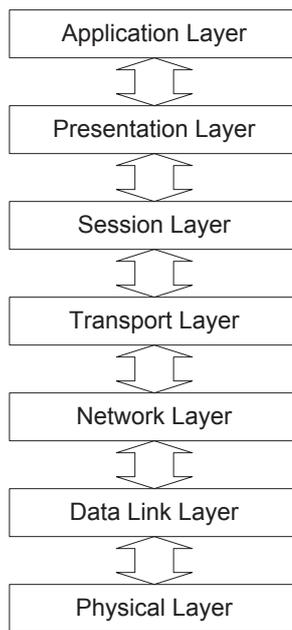


Figure 2.1: The seven layers of the OSI model.

work communication. Its functions include routing and relaying. The above three layers are sometimes referred to as the media layer, and the design of short-range wireless networks is mainly concerned with these three layers.

The transport layer ensures robust end-to-end communication; the session layer establishes system-to-system communication between two hosts; the presentation layer reformats data, including data compression and encryption; and the application layer serves as an interface to the users. These four layers are sometimes called the host layers. In short-range wireless networks, due to the limitations of the computational capacity of the nodes, the functionalities of these four layers may be reduced into one or two layers.

A typical protocol stack for a short-range wireless network such as a sensor network is shown in Fig. 2.2 [1]. Compared with the standard OSI model, the sensor network protocol stack combines the application layer, the presentation layer and the session layer into one application layer. This is because the data processing tasks in sensor

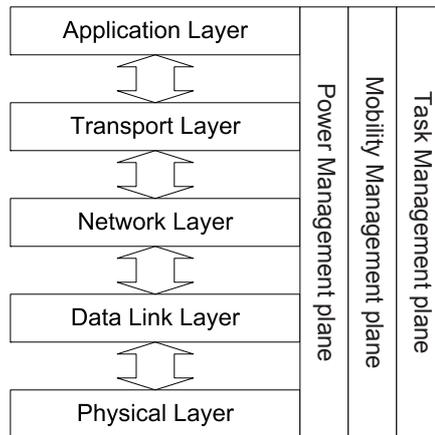


Figure 2.2: The sensor network protocol stack model [1].

networks are usually very simple compared with the tasks in traditional networks, such as the Internet. The remaining four layers remain the same. Note that the transport layer may not be required by sensor networks.

The most significant difference between the sensor network protocol stack and the OSI model is the introduction of the concept of a *plane*. As shown in Fig. 2.2, there are three planes in a sensor network protocol stack, namely the power management plane, the mobility management plane, and the task management plane. These planes are a result of the collaborative nature of wireless sensor networks and the philosophy of cross-layer design. In WSNs, aided by these management planes, cooperation is embodied not only between different stack layers but also among different sensor nodes. For instance, when the sensor battery is low, the power management plane can inform the physical layer to lower its transmit power, the data link layer to sleep longer, and the network layer to avoid participating in a route. As another example, sensors may be deployed in the field to monitor a common phenomenon. The task management plane can schedule some of the nodes to sleep while guaranteeing that the active nodes can still cover the area of interest.

Cross-layer, even cross-node, design principles are of considerable importance in

short-range wireless networks. Cross-layer design is based on the cooperation of different protocol layers. For instance, in [9] the authors propose joint optimization over transmit powers, rates, and link schedules of a wireless sensor network to maximize lifetime. Flow conservation, maximum rate, energy conservation and transmission range are considered as constraints. Convex optimization is used to solve for the optimal rates and powers for a given incidence matrix of the network graph, link gain matrix, and initial energy. However, the computational complexity of the algorithm grows as a double exponential function of the size of the network. Also, the input of the algorithm is difficult to obtain in real-time applications. In the following section, we will introduce some state-of-the-art design and optimization techniques that focus on different layers.

## 2.2 Energy Efficient Designs for Short-Range Wireless Networks

Wireless sensor networks (WSNs) are the most common type of short-range wireless networks, and researchers have studied these networks for decades [10]. Numerous strategies have been investigated to promote the energy efficiency of short-range wireless networks in general and WSNs in particular. These strategies include, but are not limited to, transmit power control, mobile data sink deployment, multiple data sink deployment, non-uniform initial energy assignment, and intelligent node/relay deployment [11, 12]. Moreover, much work has been done to minimize energy dissipation at all levels of system design, from the hardware to the protocols to the algorithms [13–15].

In this section, we highlight some important issues in the design of energy efficient wireless networks. *Adaptive communication* is an optimization technique often used at the physical layer. *impulse radio ultrawideband* and *bursty transmission scheme* are

also introduced. *Multiple access techniques* address the design issue at the data link layer. *Clustered network topology* and *multi-hop topology* can be used to improve the network performance at the network layer. *Distributed source coding* provides lossless compression of the observed random variables and feeds data into the network, which can be viewed as a technique adopted at the application layer.

### 2.2.1 Adaptive Communication

Adaptive communication enables the transmitter to adjust its transmission approach, including transmit power, modulation, and coding scheme, according to the link condition between the transmitter and the receiver. A feedback channel is required between the transmitter and the receiver so that the transmitter can acquire the channel conditions. An illustrative system model using adaptive communication is shown in Fig. 2.3. As shown in this figure, the transmitted signal is demodulated and decoded at the receiver, meanwhile the receiver estimates the channel conditions through the received signal. The estimated channel condition is then sent back to the transmitter via a feedback channel. According to the channel condition, the transmitter will choose an appropriate transmission scheme, including modulation, coding, and transmit power, to achieve a certain quality of service (QoS). For instance, the QoS can be defined as the data rate given a fixed average transmit power, or as outage probability for a fixed data rate [7].

Adaptive communication can be used in a SRWN. However, there are a number of differences between the adaptive communication techniques used in SRWNs and in traditional communication systems. First, traditionally, adaptive communication is used to achieve the maximum data rate given a time varying channel. However, in SRWNs, maximizing the data rate is usually not the priority. Instead, the target of adaptive communication in SRWNs is to minimize the energy consumption. For instance, adaptive communication can be used to find the most energy efficient way to transmit a data packet, given a certain transceiver power model and channel condition. Second, the

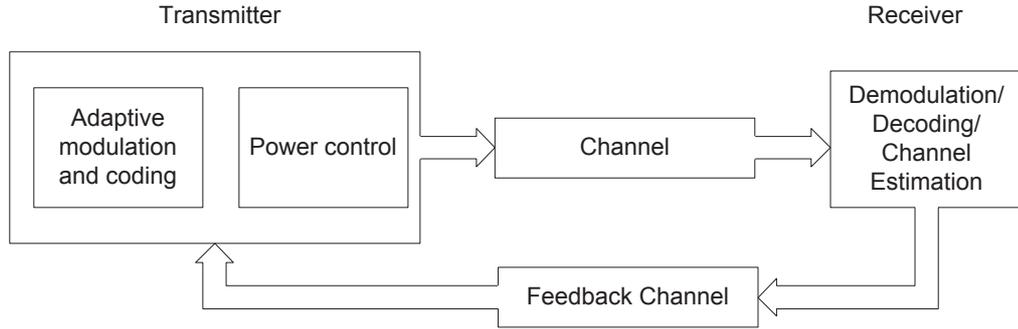


Figure 2.3: A system model with adaptive communication.

adaptive communication scheme in SRWNs is adaptive based on the path loss between a transmitter and different receivers. Traditional adaptive communication is adaptive based on the time varying channel between a transmitter and a receiver. This is caused by the limited computational capacity of SRWN nodes, which are usually small and equipped with simple circuits. Moreover, one node may communicate with different nodes simultaneously. Maintaining a time varying adaptation to multiple links is very demanding for a SRWN node. Thus, time invariant adaptive communication based on path loss are considered in our work.

To facilitate the adoption of the adaptive communication technique based on path loss, the power consumption model of the transceivers needs to be studied, and the energy consumption as a function of transmission distance needs to be established. In [14], the concept of an energy per useful bit metric is proposed. This metric sought to define a way of comparing energy consumption, specifically looking at the impact of the preamble on the effectiveness of the system. The authors define the energy per useful bit (EPUB) metric as:

$$\text{EPUB} = (\text{Preamble Overhead}) \times (\text{Total Energy}) \quad (2.1)$$

$$= \left( \frac{B_D + B_P}{B_D} \right) (P_{TX} + \sigma P_{RX})T \quad (2.2)$$

where  $B_D$  is the average number of data bits and  $B_P$  is the average number of preamble bits. The terms  $P_{TX}$  and  $P_{RX}$  are transmit and receive power, respectively. The parameter  $\sigma$  is determined by the multi-access protocol and represents the proportion of time spent in receive mode compared to the proportion of time spent in transmit mode. Finally,  $T$  is the time to transmit a bit. By looking at this metric, we can see that in finding the minimum EPUB, there is a relationship between the complexity of the MAC (i.e., the size of the preamble) and the reduction in total energy. The authors claim that a more complex MAC can reduce the total energy, but it requires a longer preamble, and the energy consumption of this longer preamble can outweigh the gains of the improved energy from the more complex MAC. The paper compares six physical layers to find the EPUB. The conclusion drawn from the analysis is that simpler non-coherent modulations such as OOK and FSK-NC have the lowest EPUB.

In [3], the authors provide detailed analysis about the power consumptions of the components at both the transmitter and the receiver ends. Moreover, the authors differentiate the power consumptions of different modulation schemes (linear or nonlinear). Both circuit power consumption and transmit power consumption are considered in [3]. A peak-power constrained optimization over the constellation sizes, linear/nonlinear modulations, and coded/uncoded transmission schemes over different transmission distances are provided. The authors concluded that at short transmission distances, bandwidth efficient schemes (uncoded linear modulations with large constellation sizes) are energy efficient; on the other hand, at large transmission distances, energy efficient schemes (coded nonlinear modulations with small constellation sizes) are energy efficient. The authors in [3] assume a fixed target bit error probability and no retransmissions. This assumption may not meet some quality of service (QoS) requirements, such as reliable communications.

In [15], the authors show how startup time, i.e., the time duration a device needs to prepare its circuits for communication, correlates with the energy efficiency of the system. This work is based on the idea that the energy consumed in startup is a sig-

nificant part of the energy consumed in a transmission. For  $M$ -ary modulations, as  $M$  increases the maximum transmit energy must increase for a fixed BER, but the number of transmissions decreases. With higher order modulations the transmitter is on for a shorter time, and therefore even with the higher maximum cost it is shown that higher order modulation schemes are more energy-efficient. However, this result does not hold when there is a large startup time. This work demonstrates the importance of evaluating the startup time of a physical layer, and it shows that for certain startup times, certain modulation schemes are preferable to others.

The idea of finding an energy-efficient optimal hop distance has been evaluated in previous work. In [16], the authors propose a distributed position-based network protocol optimized for minimum energy consumption in wireless networks. In this protocol a node determines the potential relay nodes around it based on the optimum energy dissipation of the combined transmit/receive power of the source and relay nodes. Similarly, in [17] the optimum one-hop transmission distance that will minimize the total system energy is investigated. The main conclusion of this study is that the optimum one-hop transmission distance depends only on the propagation environment and the transceiver characteristics and is independent of other factors (e.g., physical network topology, the number of transmission sources and the total transmission distance). In [18] it is shown that given a route bit error rate (BER) and node spatial density, there exists a global optimal data rate at which the transmit power can be globally minimized. The authors also report that there exists a critical node spatial density at which the optimal transmit power is the minimum possible for a given data rate and a given route BER. In this study the optimal common transmit power is defined as the minimum transmit power used by all nodes necessary to guarantee network connectivity.

The authors in [13] analytically derive the optimal hop distance given a particular radio energy dissipation model. The goal of the derivation is to minimize the total energy consumed by the network to transmit data a distance  $D$ .

$$E_{Total} = \frac{D}{d} E_{Hop} \quad (2.3)$$

where  $D$  is the total distance between the source and the destination,  $d$  is the hop distance and  $E_{Hop}$  is the total energy to transmit the data over one hop.

$$\begin{aligned}
E_{Hop} &= E_{TX} + E_{Hop,Fixed} \\
&= \alpha E_{RX} d^n + E_{TX,Fixed} + E_{RX,Fixed} \\
&\approx \alpha E_{RX} d^n + 2E_{Fixed}
\end{aligned} \tag{2.4}$$

The value  $E_{Hop}$  is made up of 2 components,  $E_{TX}$  and  $E_{Hop,Fixed}$ .  $E_{Hop,Fixed}$  is the fixed energy cost expended during the hop. This energy is based on running the circuits to perform the modulation and any other processing, and it is not dependant on the distance between the nodes or the amount of energy radiated into the channel by the radio.  $E_{Hop,Fixed}$  can be divided into two parts  $E_{TX,Fixed}$  and  $E_{RX,Fixed}$ . These are the fixed energy costs of the transmitter and receiver, respectively. While these two values are not necessarily equal, it is common practice to set them equal and thus the fixed energy is  $2E_{Fixed}$ .

The value  $E_{TX}$  is the energy consumed to appropriately amplify the signal for transmission. It can also be broken into multiple components. As seen in (2.4),  $E_{TX}$  is the product of the received energy,  $E_{RX}$ , the hop distance  $d$  raised to the path loss factor  $n$ , and a scalar  $\alpha$ .  $E_{RX}$  is the energy accumulated at the receiver, or more specifically, the desired received energy. The constant  $\alpha$  is the attenuation of the channel that comes from the wavelength of the signal and antenna gains. This constant also includes the amplifier efficiency.

Combining (2.3) and (2.4) yields the following result.

$$E_{Total} = D(\alpha E_{RX} d^{n-1} + 2E_{Fixed} d^{-1}) \tag{2.5}$$

The optimal hop distance,  $d^*$ , is

$$d^* = \sqrt[n]{\frac{2E_{Fixed}}{\alpha(n-1)E_{RX}}} \tag{2.6}$$

Equation (2.6) is the expression for the energy-efficient optimal hop distance.

In [19] the authors provide an analytical model for determining the transmission range that achieves the most economical use of energy in wireless networks under the assumption of a homogeneous node distribution. Given node locations, the authors propose a transmission strategy to ensure the progress of data packets toward their final destinations. By using the average packet progress for a single common transmission range metric, they determine the transmission range that optimizes this metric.

Optimizing the packet size in wireless networks as an adaptive communication technique has also found considerable attention in the literature [4–6, 20–22]. In [6] techniques for adapting radio parameters (e.g., frame length, error control, processing gain, and equalization) to channel variations is studied to improve link performance while minimizing battery energy consumption. In [20] an algorithm for estimating the channel BER using the acknowledgement history is presented. Estimated channel BER is used to optimize the packet size. It is reported that this algorithm can achieve close to optimal performance using a history of just 10,000 bits.

In [21] the effect of error control on packet size optimization and energy efficiency is examined. It is shown that forward error correction can improve the energy efficiency, while retransmission schemes are found to be energy inefficient. Furthermore, binary BCH codes are found to be more energy efficient than the best performing convolutional codes. In [4] an analytical model characterizing the dependency between packet length and delay characteristics observed at the application layer is presented. It is shown that careful design of packetization schemes in the application layer may significantly improve radio link resource utilization in delay sensitive media streaming under harsh propagation environments.

In [22] link adaption techniques at the MAC layer, which use adaptive frame size, are used to enhance the energy efficiency of wireless sensor nodes. To obtain accurate estimates and to reduce computational complexity, extended Kalman filtering is utilized for predicting the optimal packet size. In [23], the author considers the dynamic sizing of the MAC layer frame to improve wireless link throughput, range and energy

efficiency. The philosophy is to obtain the balance of packet overhead and packet error probability.

In [24], the author utilizes optimum packet size and error control techniques to improve the energy efficiency of wireless sensor networks. However, neither of them considered the energy consumptions of the retransmission procedure, adaptive power control, and power consumption of different components in transceivers.

In [25], the authors proposed the energy-per-useful-bit (EPUB) metric to measure the PHY efficiency of wireless networks. The authors conclude that, to minimize EPUB, high data rates, low carrier frequencies, and simple modulation schemes are preferred. However, the energy minimization procedure in [25] does not consider higher order modulation or retransmissions. For example, the authors assume that the data rate changes only through changing the width of the symbol pulse without considering coding rate and/or high-order modulations. On the other hand, the investigation of the impact of synchronization preambles in [25] is inspiring. The authors point out that the packet header plays an important role in evaluating EPUB and should be fully studied.

The authors investigated the joint optimization of the transmit power and the frame length to improve the energy efficiency of a communication link in wireless sensor networks [26]. The authors concluded that transmit power control is only beneficial within a certain distance range, while at large transmission distances, full power transmission is preferred. The investigation of this paper is experimental and thus specific to a particular device type.

Besides the work mentioned above, other considerable contributions have been made to improve the energy efficiency of wireless networks. For example, Wang *et al.* investigated the energy efficient modulation and MAC for sensor networks with the consideration of the power consumptions of detailed transceiver components as well as the start-up energy consumption [27]. Deng *et al.* studied the optimum transmission range minimizing energy efficiency in Ad Hoc networks based on node density and node coverage area [28]. In [29], the authors derive a simple distributed optimization

scheme, which is an abstracted model without the consideration of detailed channel models, packet structure, and link/MAC layer protocols. Chien *et al.* designed an adaptive radio to minimize energy consumption by adjusting the frame length, error control schemes, processing gain, and equalization based on channel conditions [30]. Cui *et al.* studied the energy per information bit,  $E_b$ , minimization problem considering the dependency of circuit power consumption on modulation and coding schemes and the time duration of a packet containing  $L$  information bits for different coding coding schemes [31]. The fundamental model is

$$E_b = \frac{\text{Power} \times \text{Time to transmit a packet}}{L}. \quad (2.7)$$

Based on this model, the authors conducted an energy minimization considering both transmit power and circuit power.

While this is the first important step towards the analytical modeling and analysis of energy consumption at the PHY layer, there are some relevant factors that are not considered in [31]. First, the authors did not consider retransmissions after a packet loss, which is essential for a reliable link. Second, header and preamble overhead is not considered. Finally, the target bit error probability is assumed to be fixed in [31]. Removal of this restriction allows for further reduction in energy consumption. While all of this previous work has limitations, as described above, it has laid a solid foundation for our study.

### 2.2.2 Impulse Radio Ultra Widedband

Impulse radio ultra wideband (IR-UWB) communications is regarded as an attractive solution to provide high data rate and low radiated power, especially for short-range wireless network applications [32]-[35]. As described previously, WSNs have been used for applications ranging from environmental monitoring and health monitoring to security and surveillance [2][36]. These different applications for WSNs have vastly different data rate requirements. Take, for example, visual sensor networks (VSNs) for

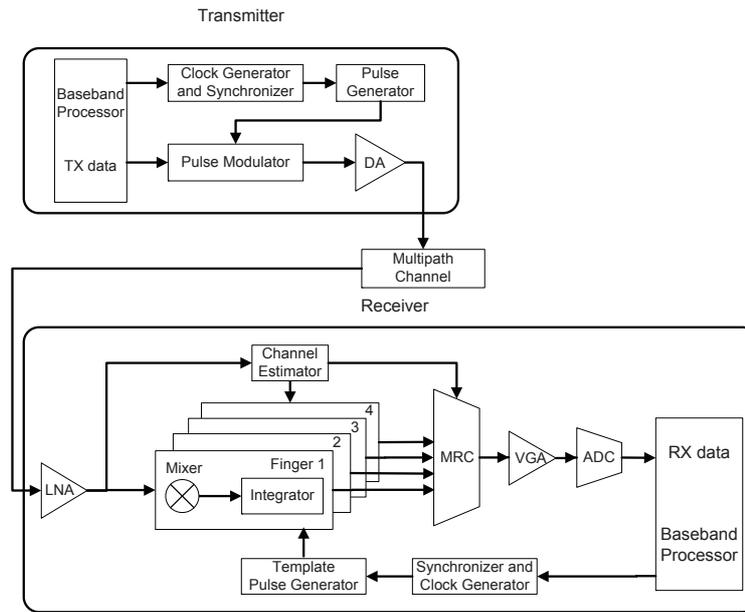


Figure 2.4: The transmitter and receiver structure in an IR-UWB system.

surveillance or health monitoring. These networks require a relatively large data rate to transmit and receive images or video in a timely manner, and a low radiated power to avoid interference with coexisting wireless systems. IR-UWB technology, in this case, has a great potential to facilitate the application of VSNs.

Compared with traditional wideband systems, IR-UWB systems feature low complexity transceiver structures and low emission power. Moreover, the main complexity of an IR-UWB system is in the design of the receiver. Fig. 2.4 shows the structure of a typical IR-UWB transmitter and receiver. As shown in the figure, an IR-UWB transmitter consists of only a pulse generator, a pulse modulator, a digital amplifier (DA), a clock generator, and a synchronizer. However, an IR-UWB receiver is equipped with several correlator branches (mixer and integrator), an analog-to-digital converter (ADC), a low noise amplifier (LNA), a variable gain amplifier (VGA), a pulse generator, a synchronizer, and a channel estimator.

Evidently, the complexity and, correspondingly, the power consumption of an IR-

UWB receiver is more than an IR-UWB transmitter. Moreover, an IR-UWB system is designed as a secondary system that co-exists and shares bandwidth with primary systems. The United States Federal Communications Commission (FCC) has a strict regulation on the transmit power of an IR-UWB system so that it does not cause noticeable degradation to the performance of existing wireless devices [32]. As a result, the power control is not feasible in an IR-UWB transmitter. Therefore, compared with adaptive modulation where the optimization is conducted on the design of the transmitter, the focus of energy efficient IR-UWB system design is on the receiver. In our work, we investigate how to adapt an IR-UWB receiver according to the link conditions.

The performance of IR-UWB has been extensively studied [37]-[39]. Some work on the optimization of IR-UWB systems considering antenna design, synchronization, and channel capacity are also present in the literature [40]-[43]. However, none of these optimizations is aimed at minimizing the energy consumption in IR-UWB systems. The energy capture effect of a RAKE receiver in IR-UWB systems is first studied by Win *et al.* in [44]. The authors analyze the relationship between the diversity level and captured energy. Although no power model is assumed in [44], the authors have concluded that there exists a threshold number of RAKE fingers in IR-UWB systems such that adding more RAKE fingers does not significantly improve performance. Despite the fact that much research has been conducted on IR-UWB systems, a detailed study on link energy minimization in IR-UWB based networks is lacking.

An effective channel model is critical in evaluating the performance of any communication system. Numerous research efforts have been made towards establishing an effective IR-UWB channel model [45]-[47]. In particular, comprehensive IR-UWB channel models for both frequency ranges from 3 – 10 GHz and below 1 GHz are provided in [46]. A single-slope power decay law is adopted to describe the path loss feature of the IR-UWB channel, and Nakagami-distributed amplitude is used to describe the small-scale fading of the IR-UWB channel [46]. This model has been accepted by the IEEE 802.15.4a Task Group as a standard model to evaluate UWB systems, and is

also used in this work to evaluate the energy consumption of different schemes.

### 2.2.3 Bursty Transmission Scheme

Given an average transmit power, it can be shown that maximizing the transmission energy efficiency is equivalent to maximizing the channel capacity. This can be illustrated as follows.

Transmission energy efficiency is measured by cost (energy consumption) per information bit. That is

$$E_{bit} = \frac{P}{C(P)}, \quad (2.8)$$

where  $E_{bit}$  represents the transmission energy consumption per information bit measured in J/bit.  $P$  is the average power consumption (J/second),  $C(P)$  denotes the channel capacity as a function of the average transmit power, measured in bits/second for a given bandwidth. Clearly minimizing  $E_{bit}$  is equivalent to maximizing  $C(P)$  for a given  $P$ .

With the consideration of circuit power consumption, it has been shown that bursty transmission schemes achieve capacity and therefore maximize the transmission energy efficiency [48]. This conclusion is contrary to the common belief that low power, constant transmission, namely lazy scheduling, maximizes the transmission energy efficiency [49]. In fact, both schedules are optimal under different assumptions. Bursty transmission is energy efficient when circuit power is not negligible, while lazy transmission is energy efficient when circuit power is negligible.

In [50, 51], the authors investigate energy-efficient packet transmission scheduling in wireless networks. In [50], it is shown that the transmission energy consumption of a packet is a non-negative, monotonically decreasing function of the transmission duration. That is, the longer it takes to transmit a packet, the less transmission energy will be consumed. However, this is not true when the energy consumed takes into account the circuit power consumption [52] [53]. In SRWNs, it is obvious that circuit

power consumption is an important factor that influences the energy efficiency of nodes. Therefore, bursty transmission schemes should be considered. This coincides with our intuition that a node should sleep as often as possible to save energy. Comprehensive results on improving energy efficiency that can provide simple solutions yet consider circuit power and energy constraints are desired.

#### **2.2.4 Multiple Access Techniques**

Multiple access techniques can be categorized into fixed access and random access techniques. Fixed access techniques include fixed-access time division multiple access (TDMA), code division multiple access (CDMA), frequency division multiple access (FDMA), and space division multiple access (SDMA). Random access protocols for sensor networks include Sensor MAC (S-MAC), SpeckMAC, and WiseMAC [54, 55]. Note that the above random access techniques share time resources among different nodes and thereby can also be viewed as TDMA. The tradeoffs between fixed access and random access are clear. Fixed access requires central control and can only be used in heterogeneous wireless networks. Also the overhead of access scheduling for fixed access techniques is high. However, fixed access can guarantee a collision free multiple access channel for nodes, even under heavy traffic loads. On the contrary, random access techniques do not require central control, have a low scheduling overhead, and incur high collision probability in a heavy traffic scenario. The choice of multi-access technique depends on the features of the SRWN of interest and its applications.

In [56] an optimal variable-length TDMA scheme is obtained in a star-topology wireless network. The authors assumed a specific transceiver power model, finite-length transmitting queues, as well as a fixed deadline for collecting all data from all transmitters. Iterative convex relaxation methods are used to solve the constrained optimization problem. A similar convex modeling method is also used in [31], where the optimal constellation size and modulation method are found. Energy storage con-

straints at the nodes are not considered in both [56] and [31].

### **2.2.5 Clustered Network Topology**

The clustering of nodes consists of grouping nodes together to form a local subnetwork. Using clustering, a flat homogeneous wireless network becomes a hierarchical heterogeneous network. Clustering is an effective method for efficient local scheduling, and it greatly improves the network scalability.

As another option to improve the energy efficiency of SRWNs, clustering protocols have been broadly adopted due to their effectiveness and simplicity. In clustered SRWNs, neighboring nodes are grouped as clusters. One of the nodes in a cluster is selected as the cluster head (CH), and the remaining nodes are the cluster members (CM). The cluster head is usually in charge of certain local coordinations, such as collecting data from the cluster members and communicating with other clusters and the data sink, while cluster members simply transmit data to the cluster head. The cluster head may be selected in a randomized manner, such as in HEED [57] or LEACH [58]. Such a randomized selection of the cluster head, combined with rotating the cluster head position, can effectively avoid the early drain of the energy of a particular node. However, it cannot guarantee the optimality of the selection. On the other hand, the cluster head can be selected by a centralized algorithm through an optimization scheme. This type of cluster head selection scheme requires a powerful control center and does not scale well. The advantage of centralized algorithms is that they can guarantee the optimality of the cluster head selection. More effective distributed algorithms or highly efficient low complexity centralized algorithms for cluster head selection are still desired.

### **2.2.6 Multi-hop Topology**

In SRWNs, nodes are usually densely deployed [1]. Thus, the distances between nodes can be very small. Correspondingly, compared with single hop, the transmit power

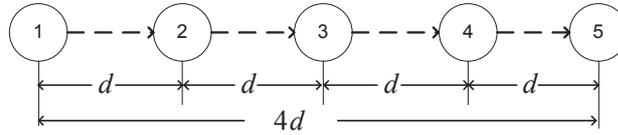


Figure 2.5: An illustration of a multi-hop scheme versus a single hop scheme.

of nodes can be kept low when communicating using multi-hop. This can be shown by the example in Fig. 2.5, where the node distance is  $d$  and is assumed to be the same for all nodes. Suppose node 1 wants to communicate with node 5. Node 1 has two choices : single hop (solid line) and multi-hop (dashed line). Assume the nodes require the received power to be at least  $P_r$  for correct communication, and the path loss exponent is 2. Then if node 1 chooses single hop, the transmit power has to be at least  $P_t = P_r(4d)^2 = 16P_r d^2$ . However, if node 1 chooses a multi-hop route, the transmit power only needs to be  $P_t = P_r d^2$ . The multi-hop scheme significantly lower the transmit power compared with the single hop scheme.

However, in terms of energy efficiency, we cannot guarantee that multi-hop is superior to single hop, especially when the circuit power consumption is taken into account. Still considering the example in Fig. 2.5, if node 1 chooses single hop, the energy consumption of the communication link is

$$E_{\text{link, singlehop}} = (16P_r d^2 + 2P_c)T, \quad (2.9)$$

where  $P_c$  is the circuit power consumption, which is the same for all nodes and is the same for transmission and for reception.  $T$  is the time duration of the transmission. To simplify the analysis, we assume that a node only consumes  $P_c$  when working as a receiver. Similarly, when using multi-hop, the communication consumes

$$E_{\text{link, multi-hop}} = (4P_r d^2 + 5P_c)T. \quad (2.10)$$

Therefore,

$$E_{\text{link, singlehop}} - E_{\text{link, multi-hop}} = (12P_r d^2 - 3P_c)T. \quad (2.11)$$

The above equation is greater than or equal to zero if  $\frac{P_r d^2}{P_c} \geq \frac{1}{4}$ , otherwise it is less than zero. That is, if the nodes are close to each other, or the circuit power consumption is large, the multi-hop scheme can be less energy efficient than using single hop, for it involves more nodes in the route than single hop. Therefore, the multi-hop and single hop schemes must be carefully evaluated in short-range energy constrained wireless networks.

The performance of multi-hop techniques has been broadly studied in the literature. In [59] the performance of multiple multi-hop routing schemes are evaluated in terms of packet delivery ratio and routing overhead. Although the simulation-based performance evaluation in [59] is quite comprehensive, the energy efficiency of different multi-hop routing schemes is not considered. The energy efficiency of a multi-hop link is investigated in [60], where the authors studied the energy expenditure of a path with a large number of short-distance hops and another with a smaller number of large-distance hops. To guarantee reliable communications, two operating models are compared: end-to-end retransmissions and hop-by-hop retransmissions. The optimal routes are found for both models.

The energy efficiency of multi-hop schemes is evaluated from a different perspective in [61], where the optimal load distribution among nodes in a multi-hop scenario has been found. The work is based on the observation that, compared with the nodes that are far away from the sink, those near the sink tend to die faster due to a high local traffic load, since the traffic load accumulates along a multi-hop route to the sink. The authors proposed varying the node's transmission range according to the distance between the node and the sink to evenly distribute the energy consumption over the entire SRWN of interest. There are many more works concerning the performance of multi-hop schemes in the literature. In our work, we seek the answer to the question of whether and how a multi-hop scheme should be used in a cluster to maximize the

amount of data that the cluster head collects from the cluster members.

### 2.2.7 Distributed Source Coding

Distributed source coding (DSC) exploits the spatial correlation of the observed random variables [62, 63]. DSC techniques are especially useful in wireless sensor networks where sensors are densely deployed and the observed random variables are closely correlated among neighboring nodes.

The key feature of DSC is that the sum source coding rate can be reduced to the joint entropy without the nodes communicating with each other. The advantages of using DSC can be better illustrated by an example. In the two source node case, two observed random variables are denoted by  $X_1$  and  $X_2$ . If the nodes do not use distributed source coding, the source coding rates at the nodes are

$$R_1 = H(X_1), R_2 = H(X_2), \quad (2.12)$$

where  $H(\cdot)$  denotes the entropy of a discrete random variable. The two nodes will generate data at a total rate of  $R_1 + R_2 = H(X_1) + H(X_2)$  bits per samples.

If DSC, in particular Slepian Wolf coding [62], is used, the source coding rates at the nodes fall into a rate region shown in Fig. 2.6. The coding rate region can be expressed as

$$\begin{aligned} R_1 &\geq H(X_1|X_2), \\ R_2 &\geq H(X_2|X_1), \\ R_1 + R_2 &\geq H(X_1, X_2). \end{aligned} \quad (2.13)$$

In this case, the two nodes can generate data at a total rate of  $R_1 + R_2 = H(X_1, X_2)$  bits per samples. We have

$$H(X_1, X_2) = H(X_1|X_2) + H(X_2) \leq H(X_1) + H(X_2), \quad (2.14)$$

with equality only if  $X_1$  and  $X_2$  are independent. Therefore, as long as the observed random variables are correlated,  $H(X_1|X_2) < H(X_1)$ , and DSC generates fewer total

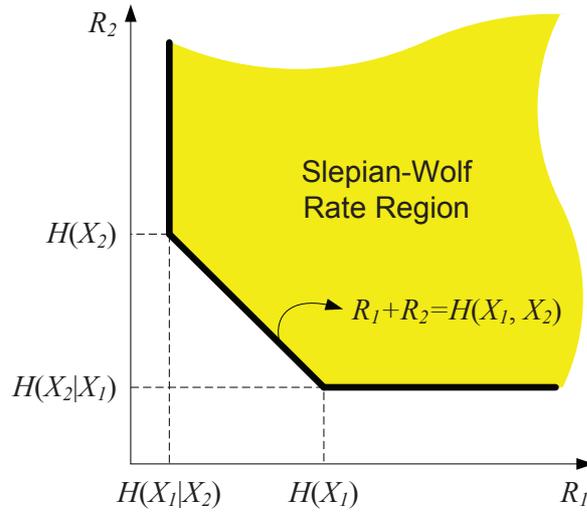


Figure 2.6: An illustration of Slepian-Wolf source coding region (2 sources).

bits per sample than individual compression. The more correlated the random variables are (smaller  $H(X_1|X_2)$ ), the more data bits can be saved by using DSC techniques. Fewer data bits will reduce the communication burden of the nodes and thereby improve the energy efficiency of the wireless network.

DSC deals with the compression of several correlated data sources, sensed by remotely located nodes, without communication among the nodes (hence distributed) such that a destination that knows the spatial correlation can recover all the data [63][64]. Consider two discrete correlated sources  $X_1$  and  $X_2$ . Slepian and Wolf showed that coding at a combined rate of  $R = H(X_1, X_2)$  is sufficient even for distributed encoding of correlated sources [62]. Specifically, the Slepian-Wolf theorem states that the achievable rate region of lossless DSC for discrete sources  $X_1$  and  $X_2$  is given by (2.13). The achievability of Slepian-Wolf coding is generalized to an arbitrary number of discrete correlated sources [63]. For lossless distributed compression of  $N$  correlated discrete sources  $X_1, X_2, \dots, X_N$  the combined rate  $H(X_1, X_2, \dots, X_N)$  is sufficient for perfect reconstruction of all the sources. In other words, there is no rate penalty due to lack of explicit side information at the encoders.

By relying on the duality between source coding and channel coding, distributed source codes have been constructed that achieve different points on the boundaries of the Slepian-Wolf rate-region [65][66]. Powerful channel codes such as low-density parity-check (LDPC) codes [67], Trellis codes [68], and convolutional and Turbo codes [69] are employed to construct lossless DSC codes for two correlated binary sources. Some of these codes are generalized to include two  $q$ -ary or  $N > 2$  binary correlated sources [70].

Theoretical lossless DSC results have been recently utilized to address energy efficient gathering of correlated data in wireless networks with mathematical optimization techniques [71]-[79]. These works address the problem of constructing correlated data gathering tree on a graph, which is different from classical network flow theory. Since the data are correlated, standard solutions (e.g., shortest path spanning tree, minimum cost flow) are not optimal, leading to new original rate allocation problems and original tree building problems, depending on the source coding model. The prominent approach is to formulate a flow-based linear programming problem to minimize the energy consumption, taking into account the capacity constraints associated with the wireless shared media, flow conservation constraints, and rate constraints enforced by the lossless DSC rate region. In these works, the authors have jointly optimized the data gathering tree and the rate allocation across the source nodes. To obtain these solutions, the wireless media is abstracted as a graph with fixed cost per information bit, which is often not the case in practice. Rather, different communication parameters such as node distances and available energy in each node eschew the optimal allocation of the communication burden.

An important question is how to assign the coding rates among nodes. In the above two source case, we want to select  $R_1$  and  $R_2$  so that  $R_1$  and  $R_2$  are within the coding region shown in Fig. 2.6 and the nodes can send the highest number of samples during their limited lifetime. The optimal coding rates should be related with the condition of the node (e.g., how much energy it has, its circuit power consumption), the

network topology (e.g., the distance from the source node to the data sink), and the observed random variables (e.g., how much entropy they have). Such detailed analysis and evaluations are conducted in our research.

## 3 Optimizing Physical Layer Parameters for Wireless Networks

In this chapter, we investigate the problem of energy-efficient transmission of data over a noisy channel, focusing on the setting of physical layer parameters. We derive a metric called the energy per successfully received bit, which specifies the expected energy required to transmit a bit successfully over a particular distance given a channel noise model. By minimizing this metric, we can find, for different modulation schemes, the energy-optimal relay distance and the optimal transmit energy as a function of channel noise level and path loss exponent. These results enable network designers to select the hop distance, transmit power and/or modulation scheme that maximize network lifetime.

### 3.1 Introduction

In short-range wireless networks (SRWNs), the energy of the nodes is usually very limited. To make the best use of the limited energy available to the nodes, and hence to the network, it is important to appropriately set parameters of the protocols in the network stack. Here, we specifically look at the physical layer, where the parameters available for optimization include: modulation scheme, transmit power and hop distance. The optimal values of these parameters will depend on the channel model. In this chapter,

we consider both an additive white Gaussian noise (AWGN) channel model as well as a block Rayleigh fading channel model. Moreover, we examine the relationship among these physical layer parameters as the channel model parameters are varied.

When a wireless transmission is received, it can be decoded with a certain probability of error, based on the ratio of the signal power to the noise power of the channel (i.e., the SNR). As the energy used in transmission increases, the probability of error goes down, and thus the number of retransmissions goes down. Thus, there exists an optimal tradeoff between the expected number of retransmissions and the transmit power to minimize the total energy dissipated to receive the data.

At the physical layer, there are two main components that contribute to energy consumption in a wireless transmission, the energy consumption for retransmission caused by channel loss and the fixed energy cost to run the transmission and reception circuitry [80]. The loss in the channel increases as a power of the hop distance, while the fixed circuitry energy cost increases linearly with the number of hops. This implies that there is an optimal hop distance where the minimum amount of energy is expended to send a packet across a multi-hop network. Similarly, there is a tradeoff between the transmit power and the probability of error. In this tradeoff, there are two parameters that a network designer can change to optimize the energy consumed: transmit power and hop distance. The third option for physical layer parameter selection is much broader than the other two. The coding/modulation of the system determines the probability of success of the transmission. Changes in the probability of a successful transmission lead to changes in the optimal values for the other physical layer parameters [15]. Here we look at the case where the probability of error is a function of the basic modulation scheme in an AWGN channel and a block Rayleigh fading channel, and it depends on the noise level of the channel and the received energy of the signal (i.e., it depends on the SNR). However, this work can be extended to incorporate any packet error or symbol error model (e.g., models that incorporate channel coding).

To illustrate these physical layer tradeoffs, consider the linear network shown in

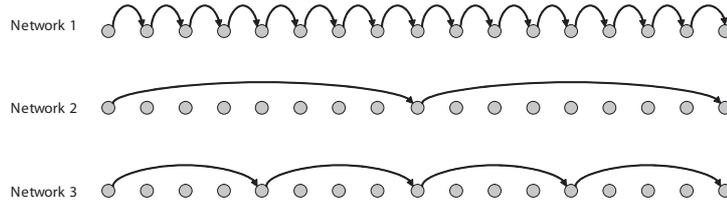


Figure 3.1: Three examples of a linear wireless network. Network 1 has a short hop distance, Network 2 has a long hop distance, and Network 3 has the optimal hop distance.

Fig. 3.1. In this network, a node must send data back to the base station. The first physical layer consideration is hop distance. In the first case (Network 1), the hop distance is very small, which translates to low per-hop energy dissipation. Because the transmit energy must be proportional to  $d^n$  where  $n \geq 2$  and  $d$  is the distance between the transmitter and receiver, the total transmit energy to get the data to the base station will be much less using the multi-hop approach than a direct transmission [80]. However, in this network, the main factor in the energy dissipation of the transmission is the large number of hops. The fixed energy cost to route through each intermediate hop will cause the total energy dissipation to be high.

In the second case (Network 2), the hop distance is very large. With so few hops there is little drain of energy on the network due to the fixed energy cost. However, there is a large energy drain on the nodes due to the high energy cost to transmit data over the long individual hop distances. With a large path loss factor, the total energy in this case will far exceed the total energy in the case of short hops. Thus it is clear that a balance must be struck, as shown in Network 3, so that the total energy consumed in the network is at a minimum.

The contribution of this chapter is a method of finding the optimum physical layer parameters to minimize energy dissipation in a multi-hop wireless network. To achieve

this goal, first we define a metric that specifies the energy per successfully received bit ( $ESB$ ). This metric is a function of three physical layer parameters: hop distance,  $d$ , transmit energy,  $E_{s, TX}$ , and the modulation scheme. In addition,  $ESB$  depends on the channel model. Given a specific channel model and a constraint on any two of the three physical layer parameters, this formula allows a network designer to determine the remaining physical layer parameter that will minimize energy dissipation and hence optimize the performance of the network.

This chapter is organized as follows. In 3.2, we explain the channel and physical layer models that are used in this work, and we describe the analytical framework used to optimize the physical layer parameters. In 3.3, we show the results of experiments to analyze the relationship between the three physical layer parameters as a function of different channel models. Section 3.4 provides analysis and discussion of the experiments as well as thoughts on future work that can be done in this area.

## **3.2 Channel and Physical Layer Model**

In this section, we derive the model for the energy per successfully received bit ( $ESB$ ) for a given transmitter/receiver structure and packet structures. The  $ESB$  model is established for AWGN channels and for block Rayleigh fading channels.

### **3.2.1 ESB Over AWGN Channels**

#### **3.2.1.1 Packet structure**

In communications systems, packets must be sent with a training sequence in order to estimate the channel conditions and facilitate the synchronization of the transmitter and receiver. The length of the training sequence depends on the estimation algorithm, synchronization algorithm, RF technology, oversampling rate, and the required system performance [81]. Usually, the longer the training sequence is, the more accurate the

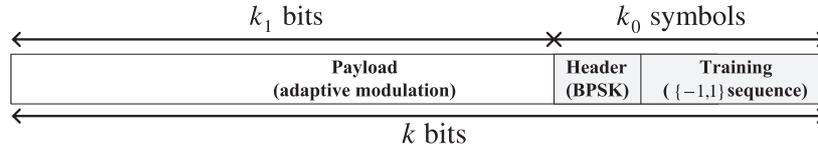


Figure 3.2: The packet structure used in AWGN channels.

channel estimate and synchronization are. Also, using more robust modulation schemes and operating at high SNRs will shorten the required training sequence length [82]. In [83], the authors state that in a slowly changing Rayleigh fading channel, a training sequence of 50 symbols can completely remove any phase offset. Thus, we assume a training sequence length of 50 symbols for our work.

Additionally, in adaptive communications systems, a header must be included to inform the receiver of the modulation scheme used for the information bits (packet payload). We assume a header length of 14 symbols. The training sequence and header must be transmitted using a predetermined modulation scheme, which will be fixed regardless of the modulation scheme used for the information bits. The modulation used for the training sequence/header should be robust even though it may be bandwidth inefficient. In this work, we assume that the training sequence consists of a binary signal ( $\{1, -1\}$ ), and the header is always modulated using BPSK, regardless of the modulation scheme used in the packet body.

We assume that a packet of length  $k$  contains  $k_1$  information-bearing bits and  $k_0$  bits of training sequence and header. Further, we assume that the training sequence and header bits are always error-free. The packet structure used for AWGN channels is shown in Fig. 3.2.

### 3.2.1.2 Energy for a single packet transmission.

We use the model from [13] for the total energy for a single packet transmission:

$$E_{Consumed} = \alpha E_{RX} \left( \frac{d}{d_0} \right)^n + E_{Fixed}, \quad (3.1)$$

where  $d_0 = 1$  meter is the reference distance. Some fixed energy is required both in the transmitter and in the receiver to run the circuitry.  $E_{Fixed}$  represents the total fixed energy in both the transmitter and receiver to transmit/receive one packet, and  $E_{RX}$  is the received energy per packet.

The relationship between the transmit and circuit power consumption and energy consumption per symbol can also be determined. Assume each symbol contains  $b$  bits and the signal bandwidth is  $B$  Hz, then the time duration to transmit a packet of  $k$  bits (with  $k_1$  information bits and  $k_0$  overhead bits) is

$$T_k = \frac{k_1}{bB} + \frac{k_0}{B}. \quad (3.2)$$

Also, we assume that the transmit power at the transmitter is  $P_t$  and the total circuit power of the transmitter and receiver is  $P_c$ . Thus, the energy to transmit and receive a packet of  $k$  bits is

$$\begin{aligned} E_{Consumed} &= (P_t + P_c)T_k, \\ &= (P_t + P_c) \left( \frac{k_1}{bB} + \frac{k_0}{B} \right). \end{aligned} \quad (3.3)$$

Since each packet contains  $k_1/b + k_0$  symbols, then the energy consumption per symbol is

$$E_s = \frac{E_{Consumed}}{k_1/b + k_0} = \frac{P_t + P_c}{B} = E_{s,TX} + E_{s,Fixed}, \quad (3.4)$$

where  $E_{s,TX} = P_t/B$  is the transmitted energy per symbol and  $E_{s,Fixed} = P_c/B$  is the fixed energy consumption per symbol. Therefore, for a fixed bandwidth,  $E_{s,TX}$  can be adjusted by changing the transmit power  $P_t$ .  $E_{s,Fixed}$  is determined by the circuitry power consumption  $P_c$ . The circuitry power consumption can be found according to the transceiver structure, modulation schemes, coding techniques, etc. In this work, we

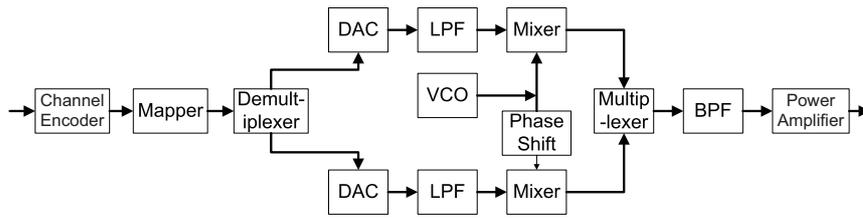


Figure 3.3: A typical transmitter structure using linear modulation.

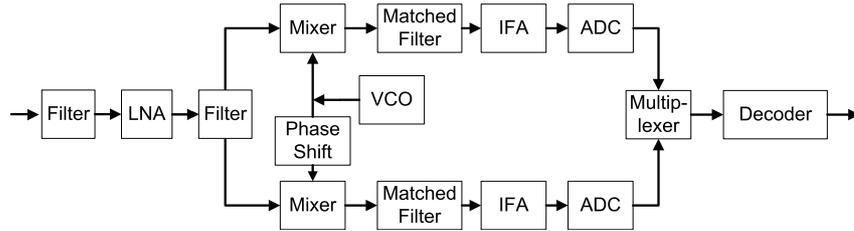


Figure 3.4: A typical receiver structure using linear demodulation.

only consider linear modulation schemes (e.g.,  $M$ -QAM), which have typical transmitter and receiver structures as shown in Figs. 3.3 and 3.4.

As shown in Fig. 3.3, the major energy consuming components at the transmitter are the digital-to-analog converter (DAC), the low pass filter (LPF), the bandpass filter (BPF), the mixer, the frequency synthesizer and the power amplifier (PA). In this work, the power consumption of the LPF, BPF, mixer, and frequency synthesizer are viewed as constants, while the power consumption of the DAC follows the model in [3]. Also, the power amplifier does not have perfect efficiency (see Section 3.3.7). The circuit power consumption here excludes the power consumed by the power amplifier. The energy consumption from the power amplifier is considered as a part of  $E_{s,TX}$ .

Fig. 3.4 shows the major energy consuming components at the receiver, which are the analog-to-digital converter (ADC), the low pass filter (LPF), the low noise amplifier (LNA), the mixer, the frequency synthesizer, and decoder. In this work, the power consumption of the LPF, LNA, mixer, and frequency synthesizer are viewed as constants. The power consumptions of the ADC and the Viterbi decoder follow the models in [3].

Table 3.1: Power Consumption Values

	$P_{filter}$	$P_{mixer}$	$P_{LNA}$	$P_{syn}$
Transmitter $P_{ct}$	2.5 mW	30.3 mW	-	50mW
Receiver $P_{cr}$	2.5 mW	30.3 mW	20 mW	50mW

The power consumption of the circuit components of the transmitter (excluding the power amplifier) and the receiver is defined as

$$P_c = 2P_{mixer} + 2P_{syn} + P_{filter} + P_{DAC} + P_{LNA} + P_{ADC} + P_v,$$

where  $P_{mixer}$ ,  $P_{syn}$ ,  $P_{filter}$  and  $P_{LNA}$  are the power consumptions of the mixers, frequency synthesizers, filters, and LNA, respectively. The above power consumptions are assumed to be constant. The values for these parameters are chosen based on typical implementations, as shown in Table 4.3 [3].  $P_{DAC}$  and  $P_{ADC}$  represent the power consumption of the DAC and the ADC, respectively.  $P_v$  is the power consumption of the Viterbi decoder.  $P_v = 0$  when uncoded modulation schemes are used. These power consumptions can be determined using the formulas in [3]. From the value of  $P_c$  and the signal bandwidth  $B$ , we can calculate  $E_{s,Fixed}$ . For example, when  $P_c = 286$  mW and  $B = 100$  kHz,  $E_{s,Fixed} = \frac{P_c}{B} = 2.86 \mu\text{J}$ .

### 3.2.1.3 ESB model

We model the probability of error in data reception using an AWGN channel with noise variance  $N_0$  to find the energy required to successfully receive a data packet. We assume that an error in the reception of the packet implies that the packet needs to be retransmitted. Thus there is a tradeoff that can be balanced to reduce energy dissipation through appropriate selection of physical layer parameters.

First, we need to find the relationship between the energy per received symbol

$E_{s,RX}$  and the transmitted energy  $E_{s,TX}$ .

$$E_{s,RX} = \frac{E_{s,TX}}{\alpha d^n} \quad (3.5)$$

The parameter  $\alpha$  is the reciprocal of the product of the amplifier efficiency ( $L$ ) and the loss in the channel. For instance, in the free space model:

$$\alpha = \frac{(4\pi)^2}{G_T G_R \lambda^2 L} \quad (3.6)$$

where in general  $L$  is a constant. Section 3.3.7 investigates the case where  $L$  is a function of  $E_{s,TX}$ . The term  $E_{s,RX}$  is used to determine the SNR of the received signal, which is important for determining the probability of error.

The probability of a successful packet transmission is as follows:

$$P_{s,p} = (1 - P_{e,s})^{\frac{k_1}{b}} \quad (3.7)$$

where  $P_{e,s}$ , the probability of a symbol error, is dependent on the SNR of the signal. Note that the above calculation of the probability assumes that the  $k_0$ -bit training sequence bits are error free. The formulas for  $P_{e,s}$  are given in Table 3.2 for a selection of modulation techniques. The value  $k_1$  is the number of information bits per packet, and  $b = \log_2 M$  is the number of bits per symbol. Thus the value  $\frac{k_1}{b}$  is the number of symbols needed for a  $k$ -bit packet containing  $k_1$  information bits.

The product of the probability of packet success and the number of data bits per packet gives the expected amount of data received per packet.

$$T = k_1 P_{s,p} \quad (3.8)$$

The ratio of the total energy to send a packet and the expected amount of data per packet gives the metric *energy per successfully received bit* ( $ESB$ ). This is the value that should be minimized by appropriate setting of the physical layer parameters.

$$\begin{aligned} ESB &= \frac{(\frac{k_1}{b} + k_0)(E_{s,TX} + E_{s,Fixed})}{T} \\ &= \frac{(\frac{k_1}{b} + k_0)(E_{s,TX} + E_{s,Fixed})}{k_1(1 - P_{e,s})^{\frac{k_1}{b}}} \end{aligned} \quad (3.9)$$

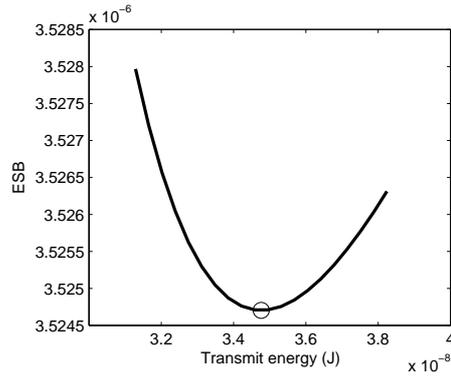


Figure 3.5: The energy per successfully received bit ( $ESB$ ) as a function of the transmit energy  $E_{s, TX}$ . This plot shows a clear minimum and thus the optimal transmit energy. These results assume a fixed distance  $d = 10m$ , BPSK modulation and fixed channel noise.

So, for BPSK modulation, the equation for  $ESB$  (see Table 3.2 for  $P_{e,s, BPSK}$ ) is:

$$ESB_{BPSK} = \frac{k(E_{s, TX} + E_{s, Fixed})}{k_1 \left[ 1 - Q \left( \sqrt{\frac{2E_{s, TX}}{\alpha d^n N_o}} \right) \right]^{k_1}} \quad (3.10)$$

Equation (3.9), the energy per successfully received bit, is the primary metric for determining the energy efficiency values. As shown in Fig. 3.5,  $ESB$  has a minimum with respect to the transmit energy  $E_{s, TX}$ .

To find the minimum of  $ESB$ , we can take the derivative with respect to  $E_{s, TX}$  and set it equal to zero. However, the equation  $\frac{d}{dE_{s, TX}} ESB = 0$  has no closed-form solution and thus the values that minimize  $ESB$  must be calculated numerically.

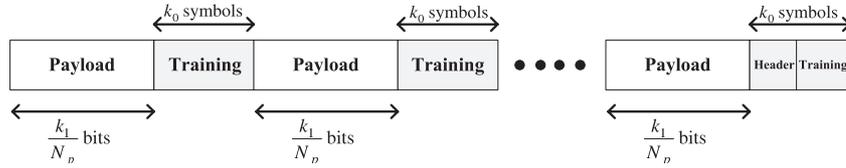
## 3.2.2 ESB Over Block Fading Channels

### 3.2.2.1 Packet structure

In narrowband communication networks, the transmitted signal most often encounters block fading. In block fading environments, the training sequence at the beginning

Table 3.2: Table of symbol error formulas from [84].

Modulation	$P_{e,s}$	
BPSK	$Q\left(\sqrt{\frac{2E_{s,RX}}{N_o}}\right)$	
QPSK	$2Q\left(\sqrt{\frac{E_{s,RX}}{N_o}}\right)$	$1 - \frac{1}{2}Q\left(\sqrt{\frac{E_{s,RX}}{N_o}}\right)$
M-PSK	$2Q\left(\sqrt{\frac{4E_{s,RX}}{N_o}} \sin \frac{\pi}{M}\right)$	
M-QAM	$1 - \left[1 - 2\left(1 - \frac{1}{\sqrt{M}}\right)Q\left(\sqrt{\frac{3}{(M-1)} \frac{E_{s,RX}}{N_o}}\right)\right]^2$	

Figure 3.6: The packet structure with header and *interleaved* training sequence in block Rayleigh fading channels.

of a packet cannot provide an effective estimation of the channel, especially when the packet length is large. Therefore, *interleaved* training sequences can be used to update the channel estimation periodically according to the coherence time of the block fading channel. The packet structure for this case of block fading is shown in Fig. 3.6.

Assume that there are  $N_p$  inserted training sequences, each of length  $k_0$ , and the coherence time of the Rayleigh fading channel is  $\tau_c$ . To have the maximum efficiency and maintain estimation accuracy, we should have

$$\frac{k_1}{bB} + \frac{N_p k_0}{B} \approx N_p \tau_c, \quad (3.11)$$

where  $k_1$  is the total number of information bits in a packet. Thus, the total number of bits in a packet is  $k = k_1 + N_p k_0$ .

The number of required training sequences is therefore

$$N_p = \frac{k_1}{b(B\tau_c - k_0)}. \quad (3.12)$$

### 3.2.2.2 Energy for a single packet transmission

For the sake of conciseness, we assume the same energy model for the transmitter and receiver in block Rayleigh fading channels as for AWGN channels. Although there are additional components in the transceiver when considering block fading channels, such as automatic gain controller (AGC) to fight Rayleigh fading, their power consumptions are constant and can be viewed as a small amount of increment over the circuit power,  $P_c$ , in AWGN channels. For example, the AGC will increase  $P_c$  by about 7 mW [85].

### 3.2.2.3 ESB model

From equation 3.3, we have

$$\begin{aligned} E_{Consumed} &= (P_t + P_c) \left( \frac{k_1}{bB} + N_p \frac{k_0}{B} \right) \\ &= (E_{s,TX} + E_{s,Fixed}) \left( \frac{k_1}{b} + N_p k_0 \right) \\ &= (E_{s,TX} + E_{s,Fixed}) \frac{k_1 B \tau_c}{b(B\tau_c - k_0)}. \end{aligned} \quad (3.13)$$

Thus, the ESB is now

$$\begin{aligned} ESB &= \frac{E_{Consumed}}{T} \\ &= (E_{s,TX} + E_{s,Fixed}) \frac{k_1 B \tau_c}{b(B\tau_c - k_0)} \frac{1}{k_1(1 - P_{e,s}) \frac{k_1}{b}} \\ &= (E_{s,TX} + E_{s,Fixed}) \frac{B \tau_c}{b(B\tau_c - k_0)(1 - P_{e,s}) \frac{k_1}{b}}. \end{aligned} \quad (3.14)$$

### 3.2.2.4 ESB model with average system outage probabilities

In fading channels, the system outage probabilities must be considered in system design. Assume that the SNR threshold is  $\gamma_T$ , then the system outage probability can be defined as

$$Pr(\gamma < \gamma_T) = \int_0^{\gamma_T} \frac{1}{\bar{\gamma}} e^{-\gamma/\bar{\gamma}} d\gamma, \quad (3.15)$$

where  $\bar{\gamma} = \frac{E_{s,RX}}{N_0}$  is the average SNR at a given distance, which is determined by path loss. Then, the ESB considering average system outage probabilities becomes

$$ESB = (E_{s,TX} + E_{s,Fixed}) \frac{B\tau_c}{b(B\tau_c - k_0)(1 - P_{e,s})^{\frac{k_1}{b}} (1 - Pr(\gamma < \gamma_T))}. \quad (3.16)$$

The selection of SNR threshold  $\gamma_T$  is very important, especially considering multi-hop transmission, since  $\gamma_T$  reflects the configuration of the transmission range of a node. A high  $\gamma_T$  will increase the outage-probability-scaled ESB in equation 3.16 and require the designer to choose more nodes to cover a given distance. On the other hand, a low  $\gamma_T$  will decrease the outage-probability-scaled ESB and make it possible to use fewer nodes to cover a given distance. However, in this work, we do not focus on the selection of SNR threshold. Instead, we view  $\gamma_T$  as a predetermined system-level parameter.

### 3.3 Optimizing Physical Layer Parameters

We performed several numerical calculations to minimize  $ESB$ , the energy per successfully received bit, and hence find the optimum transmit energy and the energy-optimal hop distances for different modulation schemes. There are considerable similarities in the analysis for AWGN and block Rayleigh fading channels. Therefore, for the sake of brevity, we focus on the analysis in AWGN channels (Sections 3.3.2 - 3.3.8), with Section 4.9 providing an illustration of the performance in block Rayleigh fading channels.

#### 3.3.1 Numerical Calculations

All numerical optimizations are performed in MATLAB. The primary optimization metric is  $ESB$ , the energy per successfully received bit. The goal is to minimize this value to reduce the energy required to transmit data successfully in the presence of

Table 3.3: Parameters Used in the Models.

Description	Parameter	Value
Fixed radio cost	$E_{s,Fixed}$	2.86 $\mu$ J/symbol
Packet size	$k$	360 bits
Overhead bits per packet	$k_0$	64 bits
Path loss exponent	$n$	3.5
Amplifier efficiency	$L$	0.02
Carrier frequency	$f$	2.4 GHz
Signal bandwidth	$B$	100 kHz
Channel coherence time	$\tau_c$	1 ms
Outage threshold	$\gamma_T$	0.1 (-10 dB)

channel noise. Because there is no closed-form solution, MATLAB is used to numerically solve the optimization of  $ESB$  with respect to transmit energy. All that is needed to find the minimum transmit energy at an arbitrary distance is to search  $ESB$  for a minima through different  $E_{s,TX}$  values. Finding optimum distances is more difficult and is described in Section 3.3.3.

As a basis, the reference noise value  $N_0$  is chosen such that the bit error rate (BER) of a BPSK symbol is  $10^{-5}$  for an energy per received bit  $E_{b,RX} = 50$  nJ. In simulations where a range of noise values are considered, the values are logarithmically spaced from  $N_0$  to  $128N_0$ . Unless otherwise specified, we used the parameters shown in Table 3.3 for determining  $ESB$ .

### 3.3.2 Optimum Transmit Energy in AWGN Channels

In this section we evaluate the case where hop distance is fixed. Finding the optimum transmit energy is a simple matter of finding the minimum of the  $ESB$  function with

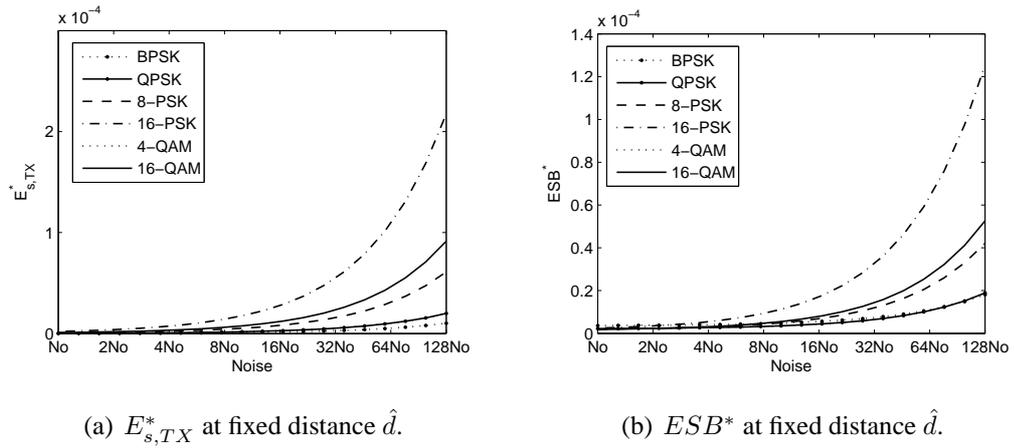


Figure 3.7:  $E_{s,TX}^*$  and  $ESB^*$  for a fixed distance,  $\hat{d} = 15m$  and a range of noise values for different modulations.

respect to energy  $E_{s,TX}$  for a particular channel  $N_0$ , and  $n$  and at a particular hop distance,  $d$ , and modulation. It was shown in Fig. 3.5 that  $ESB$  has a minimum with respect to  $E_{s,TX}$ . This value cannot be solved analytically because of the multiple Q-functions in the derivative of the  $ESB$  formula. However, the optimal  $E_{s,TX}$  can be solved numerically. Fig. 3.7 shows the optimum values of  $E_{s,TX}$  and  $ESB$  over a range of channel noise values and at different modulations. The figures were created by fixing the hop distance  $d$  to 15 m and iteratively changing the noise value  $N_0$  and modulation. For each iteration, the value of  $E_{s,TX}$  that minimizes  $ESB$  is found. The optimal  $ESB$  ( $ESB^*$ ) and the optimal  $E_{s,TX}$  ( $E_{s,TX}^*$ ) values were stored and plotted against the noise value in Fig. 3.7.

Fig. 3.7(a) shows that  $E_{s,TX}^*$  increases with channel noise. This result is expected to maintain the optimal  $ESB$ , as increased channel noise must be offset with increased transmission power to maintain a certain SNR. Fig. 3.7(b) shows that as the noise increases, the optimal  $ESB$  also increases.

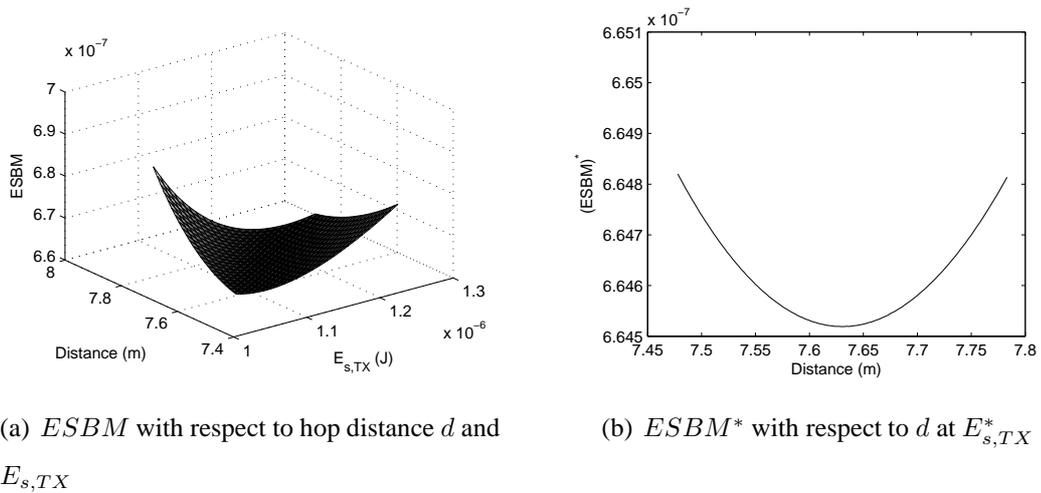


Figure 3.8: Determining optimal hop distance.

### 3.3.3 Optimum Distance in AWGN Channels

In addition to finding the optimum transmit energy, we also want to find the optimal hop distance. In this section we evaluate the case where transmit energy and modulation are fixed, and we want to find the optimum relay distance. The optimum energy-efficient hop distance  $d^*$  can be found by minimizing the  $ESB$  divided by the hop distance  $d$  (e.g.,  $ESB/d$ ). This gives the value of energy per successfully received bit per meter,  $ESBM$ . This metric is important, because if a packet needs to travel a route of distance  $D$ , then  $ESBM \times D$  gives the  $ESB$  of the entire route. Thus, by minimizing  $ESBM$ , then  $ESB$  is minimized for the entire route.

The optimal distance can be seen by looking at a plot of  $ESBM$  versus transmit energy and hop distance, shown in Fig. 3.8(a). The line of minimum values occur at each distances' optimum transmit energy value. It may appear that  $ESBM$  has a range of values that are minimum, but as seen in Fig. 3.8(b), a plot of the values along the trench,  $ESBM$  has a clear minimum value and, thus, an optimum hop distance.

Fig. 3.9 shows the optimal distance  $d^*$  and  $ESBM^*$ . Both plots were generated with  $E_{s, TX} = 5$  nJ. Fig. 3.9(a) shows that the optimum distance decreases with in-

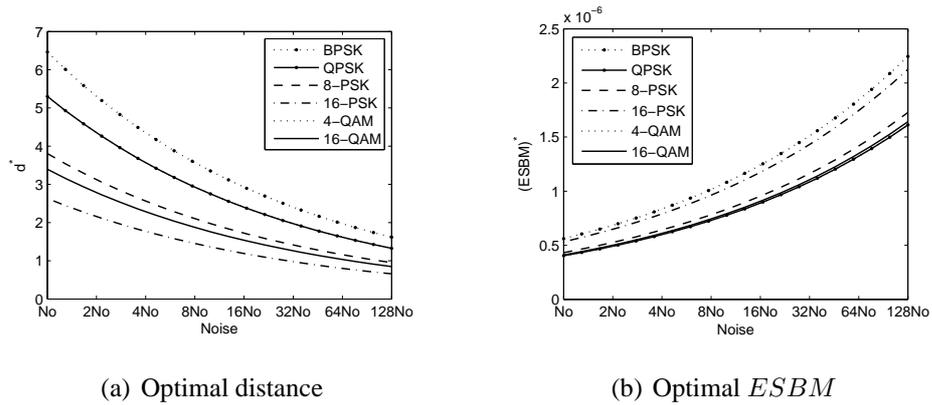


Figure 3.9: Energy optimal hop distance as a function of noise.  $E_{s,TX} = 5$  nJ.

creasing channel noise. Similarly, Fig. 3.9(b) shows that as the channel noise increases,  $ESBM^*$  increases. This is as expected, since as the channel gets worse, more energy on average to transmit the data is needed due to the increased probability of retransmission.

### 3.3.4 $ESB$ at the Optimum Distance and Transmit Energy in AWGN Channels

In Sections 3.3.2 and 3.3.3, the metric  $ESB$  was evaluated with one degree of freedom, namely,  $E_{s,TX}$  or  $d$ , respectively. In this section we look at the case where  $E_{s,TX}$  and  $d$  are both allowed to be set to their optimum values. For the analysis in this section, all the desired modulations and channel noise values were iteratively evaluated. In each iteration, the optimum hop distance was found, but instead of using one transmit power, the optimal transmit power (as described in Section 3.3.2) was found for each hop distance considered.

Fig. 3.10 shows the results when both parameters are set to their optimal values. Fig. 3.10(a) shows the optimal hop distance. As expected the optimal hop distance decreases with an increase in channel noise. Unexpectedly, Figs. 3.10(b) and 3.10(c) show that the optimal  $ESB$  and  $E_{s,TX}$  are independent of channel noise. This means

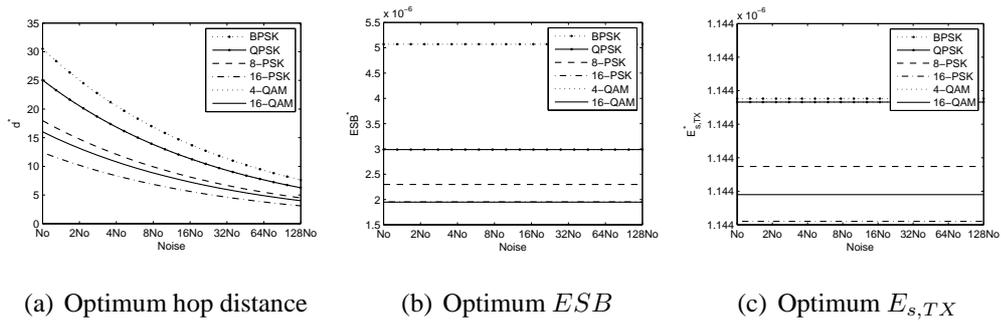


Figure 3.10: Parameters calculated using  $E_{s,TX}^*$  and  $d^*$  at each point considered.

that nodes can be set with the predetermined optimal transmit power, and that the optimal energy-efficient solution can be obtained by simply changing the hop distance as channel noise varies. This can be seen by rewriting equation (3.9) as follows:

$$ESB = \frac{\left(\frac{k_1}{b} + k_0\right)(E_{s,TX} + E_{s,Fixed})}{k_1 \left(1 - P_{e,s}\left(\frac{E_{s,TX}}{\alpha d^n N_0}\right)\right)^{\frac{k_1}{b}}}.$$

In this equation we can see that the only places that the hop distance and the noise term appear are as a product of one another. Thus the two can be regarded as one term. Once the desired  $ESB$  is found, any change in the environment that causes  $N_o \rightarrow \xi N_o$ , then the same minimum  $ESB$  can be achieved by scaling the hop distance  $d \rightarrow \frac{1}{\sqrt[n]{\xi}}d$ .

### 3.3.5 Selecting the Optimal Modulation Scheme

In Section 3.3.2 we showed how to find, for different modulation schemes, the optimal transmit energy for a given hop distance, and in Section 3.3.3 we showed how to find the energy optimal hop distance. If these two parameters of hop distance and transmit energy were the constraints on the network and it was up to the network designer to decide what type of modulation and coding to use, then it may seem that the proper solution is to find which modulation scheme has its optimal distance and transmit energy parameters nearest to the desired values provided by the network designer. However, this will not provide the best (minimum total energy) solution. As can be seen in Fig.

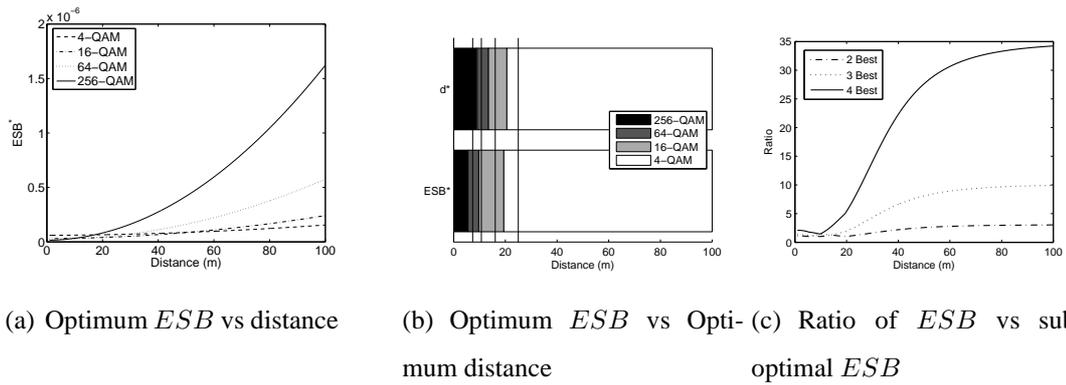


Figure 3.11: Selection of the modulation scheme for each (noise, distance) value based on (a) which modulation scheme's optimal is closest to the point and (b) which modulation scheme obtains the optimum  $ESB$  at that point. Subfigure (c) shows the ratio of  $ESB$  using the  $n$ th best modulation and the best modulation scheme.

3.11(a), for each hop distance, there is an optimal modulation scheme that minimizes energy dissipation.

Fig. 3.11(b) shows that using a particular modulation's optimum hop distance does not guarantee that it is the most efficient means of modulation. The vertical lines show where the optimal relay distances are for each modulation. The top bar shows which modulation is closest to its optimal for each distance. The lower bar shows which modulation scheme has a minimum  $ESB$  for each relay distance. We can see that these two bars are not the same, and thus we need to select the modulation scheme based on which scheme has a minimum  $ESB$  for the particular hop distance in order to minimize energy.

Fig. 3.11(c) is an evaluation of the effects of using a suboptimal modulation scheme. In this figure, the ratio between the best and the  $n$ th best modulation scheme are compared. This figure shows that the penalty for using a modulation that is only one off from the optimal scheme does not have a great impact on  $ESB$ , but using a modulation that is much different from the optimal one will perform quite poorly. Thus it is important to use either the optimal or the next-optimal modulation scheme to save

energy.

### 3.3.6 Effect of Packet Size

Packet size has a significant effect on the efficiency of the system. The model we are using gives the probability of packet success as the product of all symbol successes, as shown in (3.7). Then, for a given modulation scheme, the probability of a successfully received packet decreases as the packet size increases. Thus there is an increase in energy efficiency with small packets. However, this is only true if we do not consider the per-packet overhead. Equation (3.8) shows that the throughput of the system approaches zero as the bits per packet,  $k$ , approaches the number of overhead bits,  $k_0$ . Thus there is some optimal packet size to obtain the highest energy efficiency.

This tradeoff in packet size can be seen in Fig. 3.12, which shows the optimal energy per successfully received bit,  $ESB$ , as packet size is varied for different amounts of per-packet overhead. The case where packets have zero overhead shows the minimal energy tending to zero. However, when packet overhead is considered, there is a non-zero minimum energy packet size. As expected, as the size of the overhead increases the optimal packet size also increases.

### 3.3.7 Amplifier Efficiency

In our model, parameter  $\alpha$  that is used to encapsulate both the loss in the channel and the amplification efficiency. In all the previous experiments, this term was constant. The amplification efficiency term is due to the loss in energy from the loss in amplification of the signal before it is sent to the antenna. In a traditional model for a radio, there is some fixed cost for operating the radio. That is, for every 1 mW put into the amplifier, there will be  $\delta$  mW radiated out of the antenna, where  $\delta < 1$ .

However, this is not the most important term in the analysis of this work, as this term has only a relational impact on the equations. Rewriting (3.9) to be in terms of

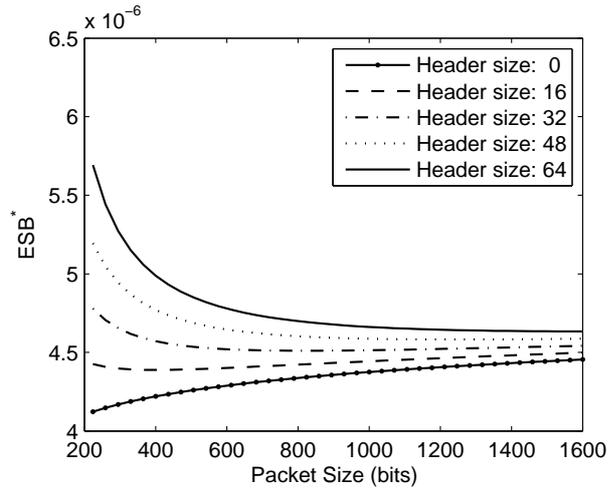


Figure 3.12: Effect of packet size on the  $ESB$ .

transmitted energy shows that the only impact of  $\alpha$  is as a scalar to the noise,  $N_0$ . As described in Section 3.3.1, the reference noise level was defined for a BPSK system to have a BER of  $10^{-5}$  and an  $E_{b,RX} = 50$  nJ. This means that using an  $\alpha$  that depends on the amplifier efficiency is equivalent to scaling the noise term, as shown in this equation:

$$ESB = \frac{\left(\frac{k_1}{b} + k_0\right)(E_{s,TX} + E_{s,Fixed})}{k_1 \left[1 - P_{e,s} \left(\frac{E_{s,TX}}{\alpha d^n N_0}\right)\right]^{\frac{k_1}{b}}}. \quad (3.17)$$

Using a constant  $\alpha$  is not the most accurate model, because in actual hardware the amplifier is more efficient at higher power levels. For example, the Tmote Sky motes developed by Sentilla Corporation (formerly MoteIV Corporation [86]) have a table that specifies the current draw of the system, which provides us with the energy values shown in Table 3.4.

Fig. 3.13 shows the optimal  $ESB$  at different noise levels, for various values of  $\alpha$ . This plot shows how the optimal  $ESB$  changes when  $\alpha$  changes. The solid line shows an example of how a non-constant  $\alpha$  changes the optimal  $ESB$ . This figure shows a

Table 3.4: Table of power consumed based on transmit power for the MoteIV Tmote Sky. Based on information from [86].

Transmit Power (mW)	Consumed Power (mW)
1.00	52
0.79	49
0.50	45
0.31	41
0.20	37
0.10	33
0.03	29
0.003	25

slight change in the shape of the curve as the value of  $\alpha$  changes. The exact shape and degree of the distortion depend on the range and degree of the nonlinearity in amplifier efficiency as a function of transmit power. As seen in this example, the distortion is not very severe and does not significantly affect the results obtained in the previous sections.

### 3.3.8 Gain Achieved By Optimizing Physical Layer Parameters in AWGN Channels

In actual wireless networks it would not be possible to place all nodes in such a way as to guarantee that nodes could always use the optimal hop distance, nor would it be possible to set transmit powers to the exact optimum level. In both cases, the physical constraints of the system in terms of topology of the SRWN and the limitations on the hardware's precision will prevent the system from achieving this theoretical optimum

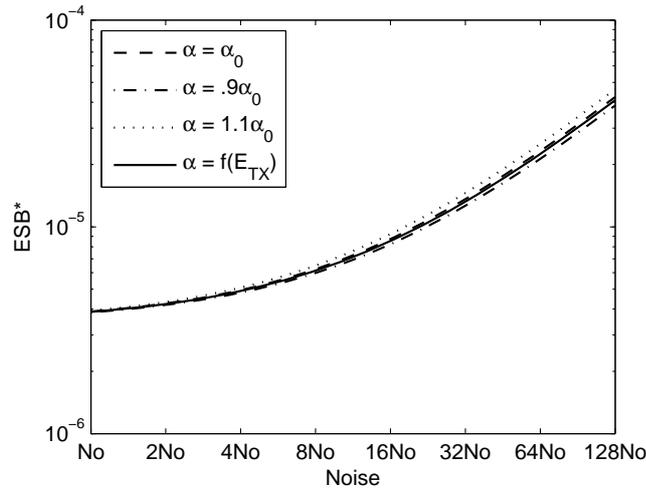


Figure 3.13:  $ESB^*$  as a function of channel noise  $N_0$  for different amplifier efficiency values.

behavior. Thus, the overall benefit of finding an optimum must be considered.

The two ways that a node could be used sub-optimally are in its hop distance and in its transmit energy precision. If the nodes' transmit energy is calibrated to transmit a particular distance, and the actual distance covered is different from this calibrated distance, then there will be a waste of energy. If the distance is smaller, the transmitter could have used less power to send the message with a similar probability of success. If the distance is longer, the probability of error will dominate and the number of re-transmissions will negatively affect the efficiency. Similarly, if the transmit power is non-optimal, there will be energy waste.

Figs. 3.14(a) and 3.14(b) show the impact of deviation from the optimum transmit energy and hop distance values, respectively. Fig. 3.14(a) shows how error in  $E_{s,TX}$  affects the performance of the system. The figure shows the ratio of  $ESB^*$  at an arbitrary distance and  $ESB$  with different  $E_{s,TX}$  used for that same arbitrary distance of 20 m. The range of  $E_{s,TX}$  used are shown in percent of  $E_{s,TX}^*$ . The figure shows that underestimating  $E_{s,TX}$  requires more energy overall than overestimating this parame-

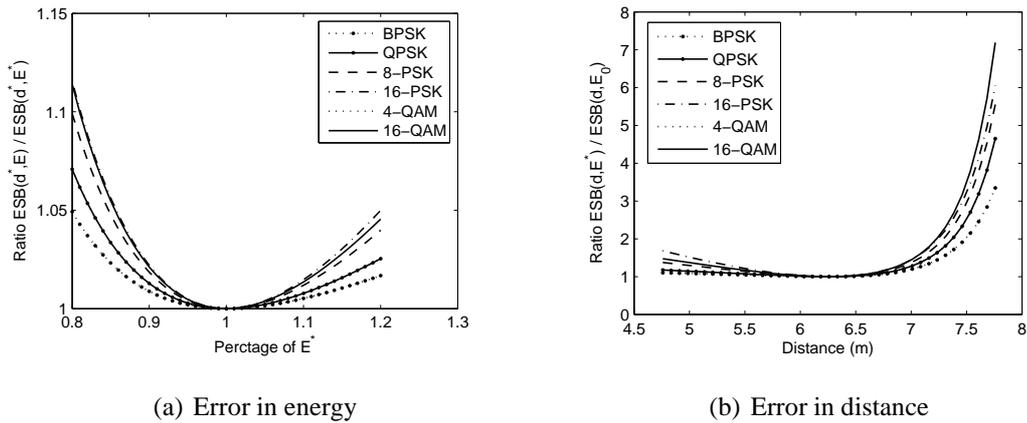


Figure 3.14: Gain of finding optimal transmit energy and optimal distance.

ter.

Fig. 3.14(b) shows the effect of using hop distances other than the one used to find the optimal transmit power. In this figure, the optimal transmit power was found for a distance of 20 m. The  $ESB$  was then found for that transmit power over the given range of distances. This was divided by the value of  $ESB$  if the optimal transmit power had been recalculated for each distance. This shows that hop distances that are greater than expected will cost much more energy than distances less than expected. Distances greater than expected would be equivalent to underestimating the transmit power, so both figures in Fig. 3.14 show that it is better to use more energy in transmission when there is uncertainty or an inability to get exact values of  $E_{s, TX}$  and  $d$ .

Table 3.5 shows the effects on  $ESBM$  of using suboptimal modulation schemes. This data tells us that the penalty for using a suboptimal modulation scheme can be quite high, and thus it is important to match the modulation scheme with the expected hop distance and channel model to reduce energy to send data in SRWNs.

Table 3.5: Percent increase in *ESBM* by using suboptimal modulation schemes. Data used in figure 3.11(c).

Maximum difference		Optimum Modulation			
		4-QAM	16-QAM	56-QAM	256-QAM
Modulation	4-QAM	0%	43%	77%	110%
	16-QAM	203%	0%	17%	37%
	56-QAM	893%	82%	0%	12%
	256-QAM	3323%	393%	41%	0%
Average difference		4-QAM	16-QAM	56-QAM	256-QAM
Modulation	4-QAM	0%	24%	63%	100%
	16-QAM	150%	0%	10%	31%
	56-QAM	683%	37%	0%	8%
	256-QAM	2566%	201%	19%	0%

### 3.3.9 The Performance in Block Rayleigh Fading Channels with Outage Probability

The performance of different modulations is also evaluated in block Rayleigh fading channels. The ESB model in this case is from (3.16). By observing (3.9) and (3.16), we find that the ESB models in AWGN channels and block Rayleigh fading channels are similar. Compared with the ESB model in AWGN channels, the ESB in block Rayleigh fading channels is scaled by the outage probability and multiple sequences of training symbols. Some illustrative results for block Rayleigh fading channels are shown in Fig. 3.15. Fig. 3.15(a) shows that, for each hop distance, there is an optimal modulation scheme that minimizes energy dissipation in block Rayleigh fading channels. Fig. 3.15(b) shows that using a particular modulation's optimum hop distance does not guarantee the most energy efficiency. Fig. 3.15(c) shows the importance of using either

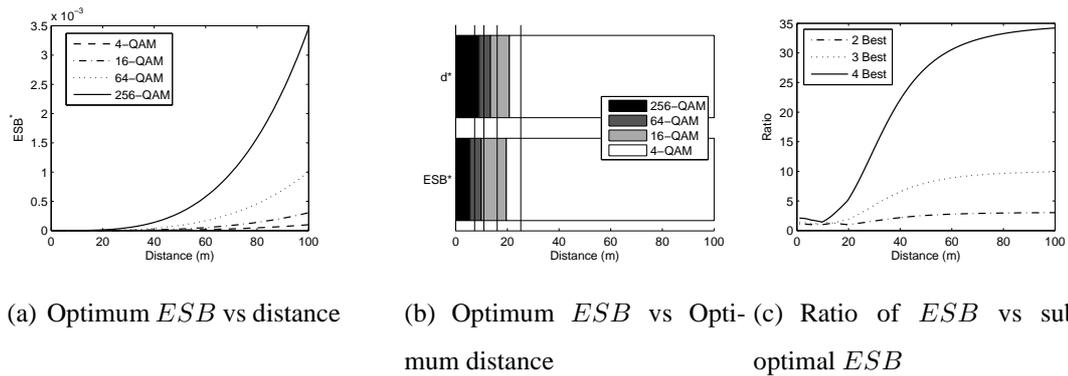


Figure 3.15: Selection of the modulation scheme for each (noise, distance) value based on (a) which modulation scheme's optimal is closest to the point and (b) which modulation scheme obtains the optimum ESB at that point. Subfigure (c) shows the ratio of ESB using the  $n$ th best modulation and the best modulation scheme.

the optimal or the next-optimal modulation scheme to save energy. These results are similar to the results for AWGN channels, and similar conclusions about optimal selection of transmit power, hop distance and modulation scheme can be made. The most significant differences in the results using AWGN and block Rayleigh fading channels are due to the increased energy consumption caused by the outage probability and the multiple sequences of training symbols. For example, the optimized ESB for 4-QAM is about  $10^{-7}$  J at  $d = 50$  m in AWGN channels; while the optimized ESB for 4-QAM increases to  $1.1 \times 10^{-5}$  J at  $d = 50$  m in block Rayleigh fading channels.

### 3.4 Summary

In this chapter we investigated the impact of physical layer parameter selection on the energy efficiency of short-range wireless networks. The analysis is conducted mostly in AWGN channels, while we show that a similar procedure can be readily adopted for the analysis in block Rayleigh fading channels. The results presented in this chapter can be of great help to adaptive network designs. For example, as the simulation results

show, once the channel and modulation scheme are known, one can easily find the optimum distance that the node should hop to get its data to the destination, as well as the optimum transmit energy. The contributions of this study are itemized as follows:

- The main conclusion of this study is that using optimal transmit energy and optimal relay distance are crucial in achieving energy efficiency for a SRWN.
- Optimizing only the transmit energy without optimizing the relay distance is not enough to achieve the best possible ESB.
- Over-estimating the transmit energy is preferable over under-estimating the transmit energy.
- If the system is operating at the optimum distance, then the transmit energy and ESB become independent of channel noise. This means that to maintain the same ESB, as the noise floor of the channel increases, the hop distance can be scaled without requiring a change in the transmit energy.
- It is important to match the modulation scheme with the expected hop distance and channel noise model in order to efficiently use the limited node energy. Average increases in ESBM from using a suboptimal modulation scheme range from 8% up to greater than 2500%.
- The results presented for AWGN channels can be extended to block Rayleigh fading channels.
- As all networks will not be operating under the same conditions, it is important for future short-range wireless network standards to allow for adaptation in order to achieve long network lifetimes.

## **4 Joint Optimization of Physical Layer and Link Layer in Narrowband Communication Systems**

In the previous chapter, we proposed a metric called energy per successfully received bit to evaluate the energy efficiency of a communication link. In this chapter, we look at cross-layer optimization of the physical and link layer, further detailing the energy per successfully received bit with the consideration of circuit power consumption, packetization and retransmission overhead, bit and packet error probability, and the duty cycle of the transceiver.

We formalize the problem of minimizing the energy dissipated to successfully transmit a single information bit over a link. In our model, we optimize the packet length and transmit power as a function of distance between the transmitter and the receiver for different modulation schemes. We propose a general unconstrained energy consumption model that provides a lower bound on the energy dissipated per information bit. A practical unconstrained physical layer optimization scheme is also provided to illustrate the utilization of the model. Furthermore, minimized energy consumptions of different modulation schemes are compared over an additive white Gaussian noise

(AWGN) channel.

We extend this general energy consumption minimization problem by considering two particular constraints: fixed average power and fixed average rate. The impacts of the average power and the information rate constraints on energy consumption are explored. We determine the optimum constellation size, packet length, and duty cycle. While only numerical optimization is used in the previous chapter, here we provide analytical expression for the optimal packet length.

## 4.1 Introduction

In the design of communication systems, often the goal is to minimize the transmit power [87, 88]. In recent years, with the advent of battery operated wireless communication nodes operating over small distances, much more attention is being paid to the overall energy consumption. In this work, we investigate the problem of joint optimization of PHY-layer and data link layer parameters to maximize the energy efficiency of a communication link.

Although an effective energy efficiency metric and a solid physical layer optimization scheme are proposed in the previous chapter, there are noticeable limitations to this work. First, in the reliable transmission scheme, the retransmission cost, both of time and energy, is not considered. Second, the optimal transmit energy per symbol and the optimal hop distance are derived for a given noise power. However, the noise power is usually fixed for a given bandwidth and noise power spectral density while the transmission distances are often unknown and variable. Therefore, to better facilitate the wireless network optimization, the optimal configurations should be provided with respect to a given transmission distance instead of a given noise power. Third, although the potential impacts of a variable packet length on the energy minimization are briefly described in the previous chapter, the packet length is not considered as an optimization parameter. Furthermore, the work in the previous chapter relies only on

numerical calculations and did not exploit the possibility of analytical solutions. Last but not least, only unconstrained optimizations are considered in the previous chapter, while in practice communication links are often bounded by many constraints, such as transmission rate requirements and average power consumption limitations.

In this chapter, we extend our work from the previous chapter to address these limitations. First, we derive a detailed model of an automatic repeat request (ARQ) based retransmission scheme by including the different energy costs from different stages of retransmissions.

Second, we thoroughly investigate the joint influence of transmit power, packet length, modulation and coding schemes on the energy performance of a wireless link, and we provide the optimal configurations with respect to transmission distances. This allows a direct adoption of the optimal configuration in the deployment of a practical wireless network. To further facilitate the real time calculation of the optimal communication scheme, we derive closed form solutions for both the optimal packet length and the optimal target bit error probabilities for a given transmission distance and  $M$ -QAM modulation scheme.

Third, we derive an analytical solution for the lower bound on the energy consumption per information bit and the optimal transmit power from an information-theoretical point of view, with the consideration of circuit power consumption and retransmission overhead.

Fourth, we study the link energy minimization problem given average power and average rate constraints by further including the optimization over duty cycles. The fixed average power constraint is particularly important in wireless networks that require a predictable lifetime, while the fixed average rate constraint is useful in wireless networks that must provide a certain quality of service (QoS), such as guaranteeing the continuity of a video stream. For both constrained optimization problems, we provide in-depth analyses of the impact of the constraints on the energy cost. Moreover, we derive analytical solutions of the optimal transmit power and duty cycle from an

information theoretical perspective in the fixed average power case.

In summary, the contribution of this work is three fold: first, we derive a comprehensive link-level energy consumption model that includes transmit power, circuit power, retransmission overhead, packetization, and duty cycle. Second, the energy minimization procedure is highly parameterized. That is, it is straightforward to adapt the proposed energy consumption model and the energy minimization procedure to transceivers with different circuitries and different channel models, such as Rayleigh and Ricean fading channels. Third, some analytical solutions are derived to allow real time calculations of the optimal configurations.

## 4.2 System and Signal Model

This section introduces the packet structure, transmitter/receiver structure, and automatic repeat request (ARQ) scheme that lay the foundation for analyzing the impacts of packetization, circuit power, and retransmissions on energy consumption. The notation used in this chapter is summarized in Table 4.1.

### 4.2.1 Packet Structure

The packet structure considered in this work is shown in Fig. 4.1. It consists of four components: payload, upper layer header, PHY/MAC-header, and preamble. We assume that there are  $L_L$  bits in the payload of each packet. The upper layer header contains the control information added by the upper layers, such as routing information, packet ID, etc. We assume there are  $L_{UH}$  bits in the upper layer header. From the view of the PHY and MAC layers, the payload and the upper layer header are indistinguishable. Therefore, the payload and the upper layer header are modulated and coded similarly.

Table 4.1: Notations

Notation	Meaning
$B$	Signal bandwidth
$d$	Transmission distance
$P_b$	Bit error probability
$N_0$	Noise power spectral density
$\eta$	Bandwidth efficiency
$\gamma$	Signal-to-noise ratio
$G$	Path loss
$G_c$	Coding gain
$R_c$	Coding rate
$P_r$	Received signal power
$P_t$	Transmit signal power
$P_c$	Circuit power (including both transmitter and receiver)
$\Lambda$	Retransmission overhead
$\Phi$	Duty cycle
$L_L$	Number of information bits per packet
$E_b$	Energy consumption per information bit

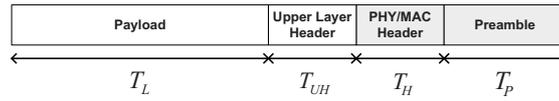


Figure 4.1: Packet structure.

Table 4.2: Packet Structure Parameters

Component	Length (bits)	Duration (s)	Modulation
Payload	$L_L$	$T_L$	Adaptive
Upper layer header	$L_{UH}$	$T_{UH}$	Adaptive
PHY/MAC header	$L_H$	$T_H$	BPSK/coded BPSK
Preamble	-	$T_P$	-

Conversely, PHY and MAC headers are modulated using a predefined modulation scheme, such as BPSK for an uncoded system and coded BPSK for a coded system. This is because the PHY and MAC headers carry important control information, such as information regarding modulation and coding for the payload and the upper layer header. Therefore, the modulation scheme of the PHY/MAC-header has to be robust and known to the receiver *a priori*, so that the receiver can always demodulate the received PHY/MAC-header, no matter what modulation scheme the payload and upper layer header use. Finally, the preamble is a predefined sequence that serves the purpose of synchronization, automatic gain control (AGC), etc. Moreover, we assume that the transmit power is constant during the entire packet. A summary of the length and duration parameters for these components are provided in Table 4.2.

### 4.2.2 Transceiver Model

In a node, energy is consumed for sensing, data processing and communications [36][89]. In this work, only the energy consumption involved in the communications is considered, since the energy consumption of sensing and data processing does not affect our

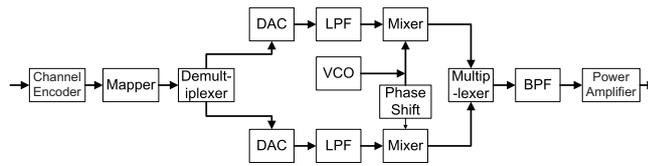


Figure 4.2: A typical transmitter structure using linear modulation.

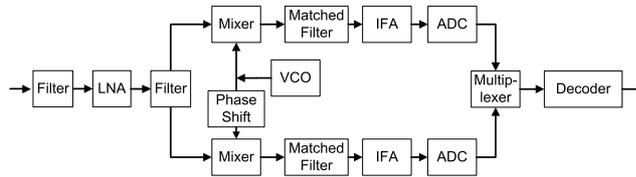


Figure 4.3: A typical receiver structure using linear demodulation.

optimization scheme. At the transmitter, energy consumption consists of the transmitted energy and the energy consumed in the circuits. At the receiver, the only energy consumption is that of the circuitry. To facilitate the analysis of the energy consumption, we assume generic transmitter and receiver models as shown in Figs. 4.2 and 4.3.

#### 4.2.2.1 Transmitter

As shown in Fig. 4.2, the major energy consuming components at the transmitter are the digital-to-analog converter (DAC), low pass filter (LPF), bandpass filter (BPF), mixer, frequency synthesizer and power amplifier (PA). In this work, the power consumption of the LPF, BPF, mixer, and frequency synthesizer are viewed as constants, while the power consumption of the DAC follows the model in [31]. The power consumption of the power amplifier can be expressed as

$$P_{amp} = \beta P_t, \quad (4.1)$$

where  $P_t$  is the transmission power and  $\beta = \frac{\varepsilon}{\rho} - 1$ ,  $\varepsilon$  is the peak-to-average ratio, and  $\rho$  is the drain efficiency of the power amplifier. Note that  $\varepsilon$  and  $\rho$  are both determined by the modulation scheme.

Table 4.3: Power Consumption Values

	$P_{filter}$	$P_{mixer}$	$P_{amp}$	$P_{LNA}$	$P_{syn}$
Transmitter $P_{ct}$	2.5 mW	30.3 mW	$\beta P_t$	-	50mW
Receiver $P_{cr}$	2.5 mW	30.3 mW	-	20 mW	50mW

#### 4.2.2.2 Receiver

As shown in Fig. 4.3, the major energy consuming components at the receiver are the analog-to-digital converter (ADC), low pass filter (LPF), low noise amplifier (LNA), mixer, frequency synthesizer, and decoder. In this work, the power consumption of the LPF, LNA, mixer, and frequency synthesizer are viewed as constants. The power consumptions of the ADC and Viterbi decoder follow the models in [31].

The power consumption of the circuit components of the transmitter (excluding the power amplifier) and the receiver is defined as

$$P_c = 2P_{mixer} + 2P_{syn} + P_{filter} + P_{DAC} + P_{LNA} + P_{ADC} + P_v,$$

where  $P_{mixer}$ ,  $P_{syn}$ ,  $P_{filter}$  and  $P_{LNA}$  are the power consumptions of the mixers, frequency synthesizer, filters, and LNA, respectively. The above power consumptions are assumed to be constant. The values for these parameters are chosen based on typical implementations, as shown in Table 4.3 [31].  $P_{DAC}$  and  $P_{ADC}$  represent the power consumption of the DAC and the ADC, respectively.  $P_v$  is the power consumption of the Viterbi decoder. These power consumptions can be determined using the formulas in [31].

#### 4.2.3 Automatic Repeat Request Sessions

In this work, ARQ is used as the link-layer protocol, that is retransmissions are required when any bit error is detected. Considering retransmission, the procedure for

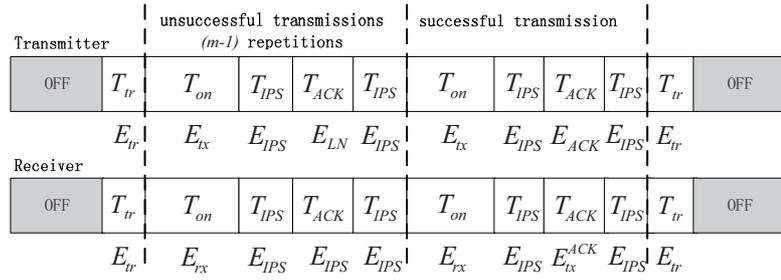


Figure 4.4: The transmission and reception of one packet using  $m$  total transmissions.

successfully transmitting/receiving one packet is shown in Fig. 4.4. We assume that before transmission or reception of a packet, the transmitter and receiver will spend  $T_{tr}$  seconds to go from the off (sleep) state to an on (active) state. Also, for a given implementation, the time period to start up the frequency synthesizer,  $T_{tr}$ , is assumed to be fixed.  $T_{IPS}$  denotes the inter packet space (IPS).  $T_{on}$  is the time duration for the transmission of one packet.  $T_{on} = (T_L + T_{UH} + T_H)/R_c + T_p$ , where  $R_c$  is the channel code rate and is set to 1 for the uncoded case.  $T_{ACK}$  is the time period when the transmitter listens for an acknowledgement. We set  $T_{ACK} = \frac{L_H}{BR_c} + T_p$ .

Assume that to successfully deliver one packet, the total number of transmissions is  $m$ . In the first  $m - 1$  transmissions, the energy consumption during the  $T_{ACK}$  period at the transmitter is denoted by  $E_{LN}$ . In the last delivery, the energy consumption during the  $T_{ACK}$  period at the transmitter is  $E_{ACK}$ .  $E_{tx}^{ACK}$  is the energy consumption of transmitting the acknowledgement after receiving the  $m^{th}$  packet. We assume that during the inter-frame space,  $T_{IPS}$ , only the frequency synthesizer contributes to the energy consumption denoted by  $E_{IPS}$ . Also, we assume that in the first  $(m - 1) T_{ACK}$  periods, the energy consumption at the receiver is  $E_{IPS}$  as well, since no ACK is transmitted.  $E_{tr}$  is the energy consumption during the transient mode,  $E_{tx}$  is the energy consumption at the transmitter to transmit one packet, and  $E_{rx}$  is the energy consumption at the receiver to receive one packet. The detailed expressions of the energy consumptions above will be defined in the following section.

## 4.3 Unconstrained Energy Minimization

### 4.3.1 Lower Bound on Energy Consumption per Information Bit

A classic model of the energy consumption per information bit when communicating at rate  $R$  is [90]

$$E_b = \frac{P_t}{R}, \quad (4.2)$$

where  $P_t$  is the transmit power and  $R$  is the information rate. However, to obtain a more refined model, we need to at least consider the following factors:

- the circuit power consumption;
- the reduction in information rate by packetization, duty cycle mode, and ARQ;
- the increase in power consumption caused by overhead.

We now provide a revised energy consumption per information bit model considering the above factors.

Firstly, considering the impact of circuit power consumption, the lower bound on energy consumption per information bit can be modeled as

$$E_b \geq \frac{P_t(1+\beta)+P_c}{B \log\left(1+\frac{P_t}{2GBN_0}\right)}, \quad (4.3)$$

where  $P_t/G$  is the received signal power after path loss  $G$ , and  $2BN_0$  is the total noise power within bandwidth  $B$ . Since  $B \log(1 + P_t/2GBN_0)$  is the channel capacity, it represents the maximum possible information rate and hence (4.3) provides a lower bound on the energy consumption per information bit.

If we assume that  $P_t/P_c = \alpha$ , a minimum value of (4.3) can be found for any given distance with respect to  $\alpha$ . The corresponding optimum value of  $\alpha$  is denoted as

$$\alpha^* = \operatorname{argmin} \frac{(1+\alpha+\alpha\beta)P_c}{B \log\left(1+\frac{P_c\alpha}{2GBN_0}\right)}. \quad (4.4)$$

The function  $\frac{(1+\alpha+\alpha\beta)P_c}{B \log(1+P_c\alpha/2GBN_0)}$  is a strict convex function of  $\alpha$  in  $\alpha > 0$ , which implies  $\alpha^* = \arg\{\frac{\partial E_b}{\partial \alpha} = 0\}$ . Thus, a closed-form expression of  $\alpha^*$  can be found:

$$\alpha^* = \left[ \frac{\left(\frac{\kappa}{1+\beta} - 1\right) \log_2 10}{W\left(\left[\frac{\kappa}{1+\beta} - 1\right] \frac{\log_2 10}{10^{\ln 2}}\right)} - 1 \right] \frac{1}{\kappa}, \quad (4.5)$$

where  $\kappa = P_c/(2BN_0G)$ , and  $W(\cdot)$  is the Lambert W function [91]. Equation (4.5) shows that  $\alpha^*$  is also a monotonically decreasing function of  $P_c$  for any given path loss, which is a function of distance between the transmitter and the receiver.

The corresponding optimum information rate and transmit power will be

$$\begin{aligned} P_t^* &= P_c \alpha^* \\ R^* &= B \log \left( 1 + \frac{P_c \alpha^*}{G2BN_0} \right). \end{aligned} \quad (4.6)$$

Based on the given parameters  $G$ ,  $B$ ,  $N_0$ , the optimum parameters  $\alpha^*$ ,  $P_t^*$ , and  $R^*$  can be calculated for any given path loss.

Furthermore we modify the initial model in (4.3) according to the reduction in information rate that comes from overhead in a real system. For example, packetization overhead, retransmission overhead, and duty cycle must be included. To incorporate the ARQ scheme and packet structure considered in this work, the lower bound on  $E_b$  in (4.3) must be modified as

$$E_b \geq \frac{(1 + \alpha + \alpha\beta)P_c}{\Lambda B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right)}, \quad (4.7)$$

where  $\Lambda$  represents the overhead induced by the link-layer protocol and the frame structure. For instance, inheriting previous assumptions,  $\Lambda$  can be defined as

$$\Lambda = \frac{L_L/B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right)}{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right)}}. \quad (4.8)$$

It is straightforward to extend the model in (4.7) to include the duty cycle of wireless transceivers, since the duty cycling can be viewed as a direct reduction in the informa-

tion rate. If the transmitter/receiver only works at  $\Phi$  fraction of the total time (we refer to  $\Phi$  as the duty cycle), the information rate then becomes

$$R = \Phi \Lambda B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right). \quad (4.9)$$

The duty cycle does not affect  $E_b$  in the unconstrained case we consider here, since

$$E_b \geq \frac{\Phi((1 + \beta)P_t + P_c)}{\Phi \Lambda B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right)} = \frac{(1 + \alpha + \alpha\beta)P_c}{\Lambda B \log \left( 1 + \frac{\alpha P_c}{2GBN_0} \right)}, \quad (4.10)$$

which is the same as the model expressed in equation (4.7). However, the duty cycle mode will have a major influence in some constrained situations, such as fixed average power transmission, as discussed in Section 4.4.

### 4.3.2 Minimization of Energy Consumption with Practical Modulation and Coding Schemes

The information rate in practice is much lower than the bound provided by the capacity  $B \log(1 + P_t/2GBN_0)$ , since the information rate is reduced by the imperfections of coding, packetization overheads, etc. Therefore, to use the general  $E_b$  model in (4.3) in practice, we need to obtain the realistic information rate and use this as the denominator in (4.3). Therefore, we modify the model of  $E_b$  from (4.3) to adopt practical information rates as follows

$$E_b = \frac{(1 + \beta)P_t + P_c}{R} = N \frac{[(1 + \beta)P_t + P_c]T_{on}}{L_L}, \quad (4.11)$$

where  $T_{on}$  is the time duration to transmit one packet,  $L_L$  is the number of information bits in one packet, and  $N$  is the total number of retransmissions needed to successfully deliver one packet. That is, the energy consumption per information bit can be equivalently expressed as the energy consumption per transmitting/receiving one packet multiplied by the average number of retransmissions required to successfully deliver the packet divided by the number of information bits contained in the packet.

### 4.3.2.1 Total number of retransmissions

We assume that there are no errors in the PHY/MAC-header. This assumption is reasonable for two reasons. The robust modulation schemes used by the PHY/MAC-header ensure that errors rarely occur in the PHY/MAC-header. Also, we assume that whenever there is a bit error in the received packet, a retransmission is required. For a packet containing  $L_L$  information bits, the probability of a packet error is

$$P_{pe} = 1 - (1 - P_b)^{L_L + L_{UH}}. \quad (4.12)$$

The expected total number of retransmissions to successfully deliver one packet is

$$N = \frac{1}{(1 - P_b)^{L_L + L_{UH}}}. \quad (4.13)$$

### 4.3.2.2 Average energy consumption per packet

Since the circuit power  $P_c$  is fixed, we only need to find the transmit power  $P_t$ . The transmit power can be determined from the SNR  $\gamma$  at the receiver and the desired bit error probability  $P_b$ . The SNR per symbol is defined as  $\gamma = P_r / (2BN_0)$ , where  $P_r$  is the received power,  $B$  is the signal bandwidth, and  $N_0$  is the spectral power density of the white Gaussian noise. The  $\gamma$ - $P_b$  function of  $M$ -QAM modulations over AWGN channel is well defined as [87]

$$P_b \approx \frac{2}{\log_2 M} e^{-\frac{3}{2(M-1)}\gamma}. \quad (4.14)$$

Then, the SNR-BER relation in  $M$ -QAM modulation is

$$\gamma = f(\gamma) = \frac{2}{3}(2^b - 1) \ln \frac{2}{bP_b}, \quad (4.15)$$

where  $b = \log_2 M$ .

Also, based on the signal propagation model, we have  $P_t = GP_r$ , where  $G$  represents the path loss, whose decibel value is determined by

$$G_{(\text{dB})} = G_{1(\text{dB})} + 10k \log_{10} d + L_{M(\text{dB})}, \quad (4.16)$$

where  $G_1 = 30$  dB is the reference path loss at 1 m,  $k = 3.5$  is the path loss constant, and  $L_M = 40$  dB is the link margin [31].

Therefore, the transmit power can be eventually denoted as

$$P_t = 2BN_0G\gamma/G_c = 2BN_0Gf(P_b)/G_c, \quad (4.17)$$

where  $G_c$  denotes the coding gain, and  $G_c = 1$  for uncoded modulation.

The ARQ procedure has been shown in Fig. 4.4. The energy consumption during each session is summarized as follows:

$$\begin{aligned} E_{tr} &= P_{syn}T_{tr}, \\ E_{IPS} &= P_{syn}T_{IPS}, \\ E_{LN} &= (P_{cr} - P_v)T_{ACK}, \\ E_{ACK} &= P_{cr}T_{ACK}, \\ E_{tx} &= [2(1 + \beta)BN_0G\gamma/G_c + P_{ct}]T_{on}, \\ E_{tx}^{ACK} &= [2(1 + \beta)BN_0G\gamma/G_c + P_{ct}]T_{ACK}, \\ E_{rx} &= P_{cr}T_{on}. \end{aligned}$$

In the above equations,  $P_{ct}$  and  $P_{cr}$  represent the power consumption of the circuits components of the transmitter and the receiver, respectively.  $P_v$  is the power consumption of the Viterbi decoder, the value of which can be calculated from [92]. Moreover, we have

$$\begin{aligned} T_{on} &= (T_L + T_{UH} + T_H)/R_c + T_p, \\ &= (L_L + L_{UH})/(R_c\eta B) + L_H/R_cB + T_p. \end{aligned} \quad (4.18)$$

That is,  $T_{on}$  is a function of packet length.

The total transmit and receive energy consumptions of  $m$  deliveries are

$$\begin{aligned} E_t(m) &= (2E_{IPS} + E_{tx} + E_{LN})(m - 1) + 2E_{tr} \\ &\quad + 2E_{IPS} + E_{tx} + E_{ACK}, \\ E_r(m) &= (3E_{IPS} + E_{rx})m + 2E_{tr} + E_{tx}^{ACK}. \end{aligned}$$

Consequently, to successfully deliver a packet, the average energy consumption is

$$\bar{E} = \sum_{i=1}^{\infty} [E_t(i) + E_r(i)] Pr\{m = i\}, \quad (4.19)$$

where  $m$  is the number of transmissions and  $Pr\{m = i\}$  denotes the probability that the number of transmissions equals  $i$ , which is given by  $Pr\{m = i\} = P_{pe}^{i-1}(1 - P_{pe})$ . After simplification, we have

$$\begin{aligned} \bar{E} &= (2E_{IPS} + E_{tx} + E_{LN})N + 2E_{tr} + P_v T_{ACK} \\ &+ (3E_{IPS} + E_{rx})N + 2E_{tr} + E_{tx}^{ACK}. \end{aligned} \quad (4.20)$$

From previous analysis,  $\bar{E}$  is a function of target bit error probability  $P_b$  and packet length  $L_L$ . Thus, the minimization of  $E_b$  can be conducted over  $L_L$  and  $P_b$ .

#### 4.3.2.3 Minimization of energy consumption per information bit

Each packet contains  $L_L$  information bits. Therefore, the average energy consumption per information bit is

$$\bar{E}_{bit} = \frac{\bar{E}}{L_L} \quad (4.21)$$

To minimize  $\bar{E}_{bit}$  with respect to  $L_L$ , we set  $\frac{\partial \bar{E}_{bit}}{\partial L_L} = 0$ , which gives us

$$A_1 L_L^2 + B_1 L_L + C_1 = 0, \quad (4.22)$$

where

$$\begin{aligned} A_1 &= \frac{P_{on} P_b}{B\eta}, \\ B_1 &= P_b \left( 5E_{IPS} + E_{LN} + P_{on} T_p + \frac{P_{on} L_H}{BR_c} + \frac{P_{on} L_{UH}}{B\eta R_c} \right), \\ C_1 &= - \left( 5E_{IPS} + E_{LN} + 4E_{tr} + E_{tx}^{ACK} + P_v T_{ACK} \right. \\ &\quad \left. + P_{on} T_p + \frac{P_{on} L_H}{BR_c} + \frac{P_{on} L_{UH}}{B\eta R_c} \right), \\ P_{on} &= 2(1 + \beta)BN_0G\gamma/G_c + P_c. \end{aligned}$$

Solving (4.22) yields the optimum number of information bits per packet,  $L_L$

$$L_L^* = \frac{-B_1 + \sqrt{B_1^2 - 4A_1C_1}}{2A_1}. \quad (4.23)$$

Correspondingly, the optimum target  $P_b$  can be found by solving  $\frac{\partial \bar{E}_{bit}}{\partial P_b} \Big|_{L_L^*} = 0$ . When the  $M$ -QAM family is used, the corresponding closed-form solution of the optimum target  $P_b$  can be found approximately as

$$P_b^* \approx \frac{1}{1+(L_L+L_{UH}) \left[ \ln\left(\frac{2}{b}\right) + 10 + \frac{P_c T_{on} + 5E_{IPS} + E_{LN}}{\frac{2}{3}(2^b-1)A_2} \right]}, \quad (4.24)$$

where  $A_2 = \frac{2(1+\beta)BN_0GT_{on}}{G_c}$ .

When transmission distance  $d$  is large,  $A_2 \rightarrow \infty$ , and equation (4.24) becomes

$$P_b^* \approx \frac{1}{1 + (L_L + L_{UH}) \left[ \ln\left(\frac{2}{b}\right) + 10 \right]}. \quad (4.25)$$

Therefore, the target bit error probability will eventually converge to a value solely determined by the packet length and the modulation scheme. The optimum target bit error probabilities of other modulation schemes and their corresponding convergence values can be obtained similarly. Furthermore, equation (4.23) reveals a one-to-one relation between  $P_b^*$  and  $L_L^*$  at any given distance. Thus, as  $P_b^*$  converges,  $L_L^*$  will also converge for higher transmission distances. The analysis and calculation results for this model will be shown and discussed in detail in Section 4.6.

So far, we have discussed the minimization of the energy consumption per information bit in an unconstrained framework. However, in practice, different constraints may apply, among which the most common ones are an average power constraint and an average rate constraint. In the following sections, we will study how these two constraints affect the minimization of energy consumption.

## 4.4 Energy Minimization with a Fixed Average Power Constraint

In this section, we consider bursty transmissions with a strict average power constraint. A bursty transmission in this work means that the transceiver only transmits/receives

for a fraction of time (duty cycle). Assume that the average power constraint is

$$\Phi[(1 + \beta)P_t + P_c] = \bar{P}. \quad (4.26)$$

Then,  $E_b$  in (4.7) can be expressed as

$$E_b \geq \frac{(1 + \beta)P_t + P_c}{\Lambda B \log\left(1 + \frac{P_t}{2GBN_0}\right)} = \frac{\bar{P}}{\Phi \Lambda B \log\left(1 + \frac{\bar{P}/\Phi - P_c}{2(1+\beta)GBN_0}\right)}. \quad (4.27)$$

For a given  $\bar{P}$ , the minimization of  $E_b$  is equivalent to the maximization of  $R = \Phi \Lambda B \log\left(1 + \frac{\bar{P}/\Phi - P_c}{2(1+\beta)GBN_0}\right)$ . Using the expression of  $\Lambda$  given in (4.8), we have the following optimization model

$$\max R = \frac{\Phi L_L}{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B \log\left(1 + \frac{\bar{P}/\Phi - P_c}{2(1+\beta)GBN_0}\right)}}, \quad (4.28)$$

$$\text{subject to } \Phi[(1 + \beta)P_t + P_c] = \bar{P},$$

$$0 \leq \Phi \leq 1.$$

where  $L_L$  represents the number of information bits contained in a packet. It is obvious that to achieve the maximum information rate  $R$ , we should have  $L_L \rightarrow \infty$ . However, in practice, the packet length is always finite and overheads are inevitable. Equation (4.28) implies that overheads of both packetization and ARQ decrease the maximum possible information rate from the ideal information rate.

For a given modulation scheme with bandwidth efficiency  $\eta$ , the channel capacity of a wireless channel with bandwidth  $B$  is limited to  $B\eta$ . Moreover, if we consider the influence of finite packet length and channel distortion, the information rate can be further specified as

$$\begin{aligned} \max R &= g(\Phi, L_L, \eta) P_{pc} \\ &= \frac{\Phi L_L}{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B\eta}} (1 - P_b)^{L_L + L_{UH}}, \end{aligned} \quad (4.29)$$

$$\text{subject to } \Phi[(1 + \beta)P_t + P_c] = \bar{P},$$

$$0 \leq \Phi \leq 1,$$

where  $g(\Phi, L_L, \eta) = \frac{\Phi L_L}{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B\eta}}$  which is an increasing function of both  $L_L$  and  $\Phi$  for a given  $\eta$ , and  $P_{pc}$  is the packet-correctly-received probability. For a given distance,  $\bar{P}$  and  $P_c$ ,  $P_b$  is a monotonically increasing function of  $\Phi$ , since large  $\Phi$  implies small  $P_t$ . This makes  $P_{pc} = (1 - P_b)^{L_L + L_{UH}}$  a decreasing function of  $\Phi$ . In addition,  $P_{pc}$  is also a decreasing function of  $L_L$ , since the larger the packet is, the greater the possibility of an error in the packet. Therefore, there exists an optimum combination of  $(L_L^*, \Phi^*)$  that balances  $g(\Phi, L_L, \eta)$  and  $P_{pc}$  and thereby maximizes  $R$ . This idea can be further explained by the following example.

Take  $M$ -QAM using fixed average power as an example. The expression of  $P_b$  in this case is

$$P_b \approx \frac{2}{\log_2 M} e^{-\frac{3(\bar{P}/\Phi - P_c)/2(1+\beta)GBN_0}{2(M-1)}}. \quad (4.30)$$

Also, we have bandwidth efficiency  $\eta = \log_2 M$  in this case. Consequently, (4.29) becomes

$$\max R = \frac{\Phi L_L}{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B \log_2 M}} (1 - P_b)^{L_L + L_{UH}} \quad (4.31)$$

subject to  $0 \leq \Phi \leq 1$ ,

$$P_b = \frac{2}{\log_2 M} e^{-\frac{3(\bar{P}/\Phi - P_c)/2(1+\beta)GBN_0}{2(M-1)}}.$$

The above minimization problem is readily solvable through numerical methods. The results are presented in Section 4.6 and explained in detail.

#### 4.4.1 The Influence of Average Power

The imposed average power constraint plays an important role on the resulting minimum energy consumption per bit. That is, the given average power constraint determines how close the resulting  $E_b^*$  is to the unconstrained global minimum value of  $E_b$ . In this subsection, we investigate the influence of the average power constraint and find the condition under which the fixed average power constrained model could give the same global minimum  $E_b$  as in the unconstrained case.

The minimization model can be rewritten as

$$\begin{aligned} \min E_b &= \frac{\bar{P}}{\Phi} \frac{2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B \log(1 + (P/\Phi - P_c)/2(1+\beta)GBN_0)}}{L_L}, \\ &= \frac{(P_t(1+\beta) + P_c)(2T_{IPS} + T_{ACK} + T_P + \frac{L_L + L_{UH}}{B \log(1 + P_t/2GBN_0)})}{L_L}, \end{aligned} \quad (4.32)$$

$$\begin{aligned} \text{subject to } \Phi[P_t(1 + \beta) + P_c] &= \bar{P}, \\ 0 &\leq \Phi \leq 1. \end{aligned}$$

This model can be used to describe any modulation scheme. The minimum possible  $E_b$  should be achieved over all possible  $P_t > 0$ . However, the constraints in (4.32) indicate that

$$\begin{aligned} \Phi[P_t(1 + \beta) + P_c] = \bar{P} &\Rightarrow P_t = (\frac{\bar{P}}{\Phi} - P_c)/(1 + \beta), \\ 0 \leq \Phi \leq 1 &\Rightarrow (\bar{P} - P_c)/(1 + \beta) \leq P_t < \infty. \end{aligned} \quad (4.33)$$

Thus,  $P_t \in [(\bar{P} - P_c)/(1 + \beta), \infty)$ . To ensure that  $P_t$  is a nonnegative value, we have that  $\bar{P} - P_c \leq 0$ . That is, when  $\bar{P} \leq P_c$ , the fixed-power transmission can achieve the same minimum  $E_b$  as that of the non-constrained transmission presented in Section 4.3. This is because when  $\bar{P} \leq P_c$ ,  $P_t$  can be any nonnegative value between  $[0, \infty)$ .

#### 4.4.2 The Equivalence of $\Phi$ , $\alpha$ , and $P_t$

In this subsection, we assume that  $\bar{P}$  is properly set so that (4.32) could achieve a global minimum value. That is  $\bar{P} < P_c$ .

From the following relations

$$\begin{aligned} \Phi[P_t(1 + \beta) + P_c] &= \bar{P}, \\ P_t &= \alpha P_c, \end{aligned} \quad (4.34)$$

we have  $\alpha = (\frac{\bar{P}}{\Phi P_c} - 1)/(1 + \beta)$ . That is, for a given pair of  $(\bar{P}, P_c)$ , if we can freely adjust  $\Phi$  and  $\alpha$ , we can achieve an optimum pair  $(\Phi^*, \alpha^*)$ , where  $\alpha^*$  is achieved by solely solving model (4.3) and  $\Phi^*$  is achieved by solely solving model (4.28).

On the other hand, if we are given a fixed  $\Phi$  and a relaxed average power consumption constraint as

$$\Phi[P_t(1 + \beta) + P_c] \leq \bar{P}, \quad (4.35)$$

then we have  $\alpha \leq \left(\frac{\bar{P}}{\Phi P_c} - 1\right)/(1 + \beta)$ . For a given  $\Phi$ , the resultant optimized  $\alpha$  may not be overall optimum. In fact, we have

$$\alpha^* = \begin{cases} \left[ \frac{\left(\frac{\kappa}{1+\beta} - 1\right) \log_2 10}{W\left(\left(\frac{\kappa}{1+\beta} - 1\right) \frac{\log_2 10}{10^{ln2}}\right)} - 1 \right] \frac{1}{\kappa}, & \text{if } \Phi \leq \Phi^{th} \\ \left(\frac{\bar{P}}{\Phi P_c} - 1\right)/(1 + \beta), & \text{if } \Phi > \Phi^{th} \end{cases} \quad (4.36)$$

where  $\kappa = P_c/(2BN_0G)$ , and  $W(\cdot)$  is the Lambert W function, and

$$\Phi^{th} = \frac{\bar{P}}{P_c} \left( \left[ \frac{\left(\frac{\kappa}{1+\beta} - 1\right) \log_2 10}{W\left(\left(\frac{\kappa}{1+\beta} - 1\right) \frac{\log_2 10}{10^{ln2}}\right)} - 1 \right] \frac{1 + \beta}{\kappa} + 1 \right)^{-1}. \quad (4.37)$$

That is, when the duty cycle is larger than a critical value  $\Phi^{th}$ , the resulting  $\alpha^*$  is the overall optimal. Otherwise, the resulting  $\alpha^*$  will be suboptimal. In summary, for minimizing energy consumption, the assigned duty cycle should be sufficiently large ( $\Phi > \Phi^{th}$ ).

We have discussed the energy minimization problem with fixed average power constraint. Additionally, we have investigated the impacts of the average power constraint and duty cycle on the energy performance.

## 4.5 Energy Minimization with a Fixed Average Rate Constraint

In some cases, such as a sustainable video stream, a stable node-to-node throughput is desired so that a certain quality-of-service (QoS) can be guaranteed. The above analysis can be readily adapted to fixed average rate situations. The minimization of

energy consumption per information bit under the fixed average rate constraint for a given modulation scheme can be modeled as

$$\min E_b = \frac{[P_t(1+\beta)+P_c]\Phi}{R} \quad (4.38)$$

$$\text{subject to } |R - R_{const}| \leq \Delta_R,$$

where  $R = \frac{\Phi L_L}{2T_{IPS}+T_{ACK}+T_P+\frac{L_L+L_{UH}}{B\eta}}(1-P_b)^{L_L+L_{UH}}$ , as defined previously,  $R_{const}$  is the desired information rate and  $\Delta_R$  is the allowed information rate deviation, since the deviation is unavoidable in practice due to unpredictable circumstances (*e.g.*, processing delays from upper layers).

Moreover, since an average power constraint no longer exists, the one-to-one relationship between  $\Phi$ ,  $\alpha$  and  $P_t$  disappears. This minimization problem thus must be conducted over all possible  $L_L$ ,  $\Phi$ ,  $M$ , and  $P_t$ . The resulting average power and minimum  $E_b$  will be

$$\begin{aligned} E_b^* &= \frac{[P_t^*(1+\beta)+P_c]\Phi^*}{R_{const}}, \\ \bar{P}^* &= [P_t^*(1+\beta) + P_c]\Phi^*. \end{aligned} \quad (4.39)$$

As with the fixed average power constraint, the minimization of energy consumption per information bit with the fixed average rate constraint is not the overall minimum  $E_b^*$  unless the target information rate is the overall optimal rate that minimizes energy consumption in the unconstrained case.

## 4.6 Numerical Results

In this section, we present numerical results that verify the previous analysis and provide insight into the performance of the different optimization frameworks. We assume a bandwidth of  $B = 10\text{KHz}$ ,  $L_{UH} = 160$  bits,  $L_H = 32$  bits,  $T_P = 20$  ms, coding rate  $R_c = 1/2$ , and coding gain  $G_c = 6.47$ . The transient period for the transceiver is set to  $T_{tr} = 5 \mu\text{s}$ . The inter packet space  $T_{IPS} = 5$  ms. The power consumption values

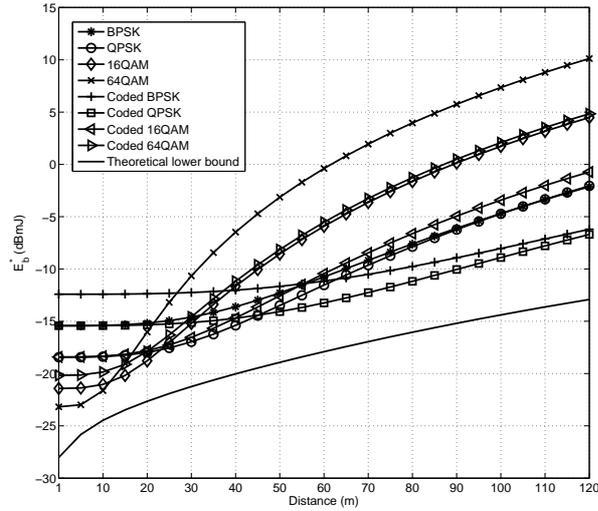


Figure 4.5: The minimum  $E_b$  of different modulation schemes vs. distance compared with the theoretical lower bound of  $E_b$ .

are shown in Table 4.3, which gives us the total circuit power consumption at the transmitter and receiver as  $P_c = P_{ct} + P_{cr} = 0.2884$  W. The power amplifier coefficient  $\beta = 0.35$ . The optimization is implemented using a simple search algorithm. This approach is sufficient in practice, as the optimization can be performed and the optimum configurations determined off-line before network deployment. Then a simple look-up table can provide the optimal parameters depending on the existing conditions.

#### 4.6.1 Unconstrained Energy Minimization

In this subsection, we compare the unconstrained energy consumption per information bit lower bound in (4.3) with the practically minimized energy consumption per information bit (4.21). In the practical scheme, the modulations considered are confined to coded and uncoded BPSK, QPSK, 16-QAM and 64-QAM.

Fig. 4.5 shows a comparison of the practical minimum  $E_b$  of different modulation schemes and the theoretical lower bound of  $E_b$ , which is given by (4.3). As transmission

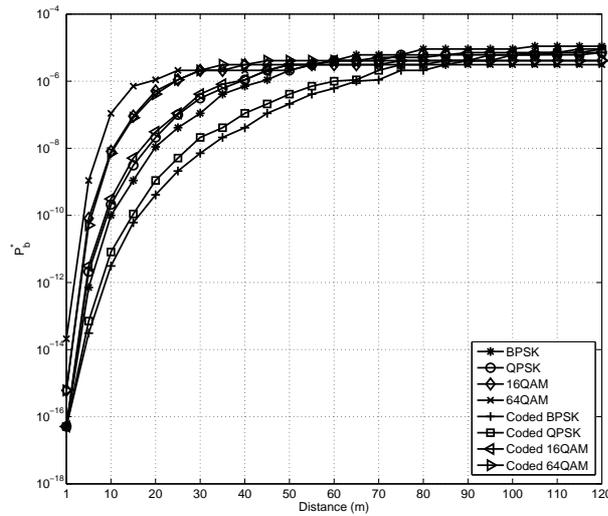


Figure 4.6: Optimized target bit error probability vs. transmission distance ( $L_L = L_L^*$ ).

distance increases, the total energy consumption per information bit increases. This is mainly caused by the increasing transmitted energy. As shown in Fig. 4.5, uncoded 64-QAM, uncoded 16-QAM, uncoded QPSK, and coded QPSK are preferred for ultra short, short, medium and long distances, respectively. This observation is justified by noting the fact that at short distances, the energy consumption is dominated by that of the circuitry. Consequently, bandwidth efficient modulation schemes that lead to shorter on time will have an advantage. On the other hand, at longer distances, the energy consumption is dominated by the transmitted energy. Hence, modulation and coding schemes that require lower SNR will have an advantage. The curve labeled ‘Theoretical lower bound’ in Fig. 4.5 is obtained by directly inserting  $\alpha^*$  in (4.4) into  $E_b$  from (4.3).

Fig. 4.6 presents  $P_b^*$  at different transmission distances. As the transmission distance increases,  $P_b^*$  will increase as well. This is because, as transmission distance increases, a higher target  $P_b$  is preferred lest the transmission energy increase dramatically to mitigate the path loss. Moreover, as transmission distance increases, a flattening of  $P_b^*$  can be observed, which is consistent with (4.25).

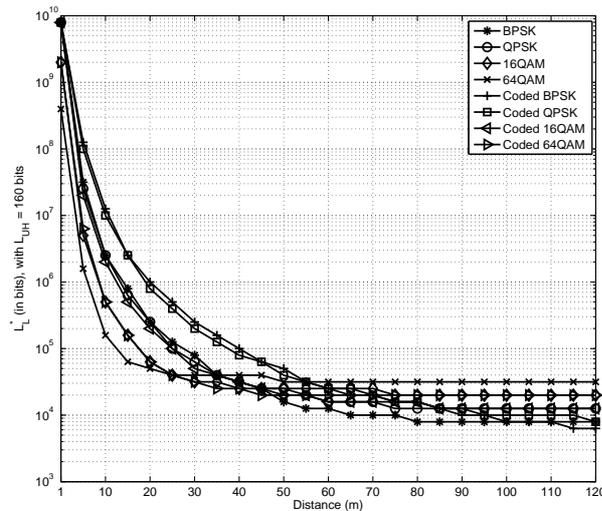


Figure 4.7: Optimized packet length vs. transmission distance ( $P_b = P_b^*$ ).

Fig. 4.7 depicts  $L_L^*$  at different transmission distances.  $L_L^*$  decreases as transmission distance increases and converges to a certain value at large transmission distances. Recall that  $P_b^*$  increases as  $d$  increases, which gives rise to a higher retransmission probability. Therefore, to reduce the retransmission cost, a shorter packet length is preferred. Also, the convergence of  $L_L^*$  occurs at large transmission distances as  $P_b^*$  flattens.

Fig. 4.8 shows the optimum transmit power,  $P_t^*$ , that minimizes  $E_b$ . The theoretical and practical  $P_t^*$ s have the same trend. However, the theoretical  $P_t^*$  curve is smooth, while the practical  $P_t^*$  curve exhibits a saw-toothed shape. This irregularity is caused by the limitations of using a discrete modulation and the packetization parameters used in the calculations for the practical model.

Fig. 4.9 compares the optimum rate,  $R^*$ , obtained through (4.31) and the optimum information rate from (4.9). We can tell that the framework from (4.31) provides an upper bound for the optimum information rate. Also, the staircase type curve of the calculated optimum information rate is caused by the discrete nature of modulation used in the numerical calculations.

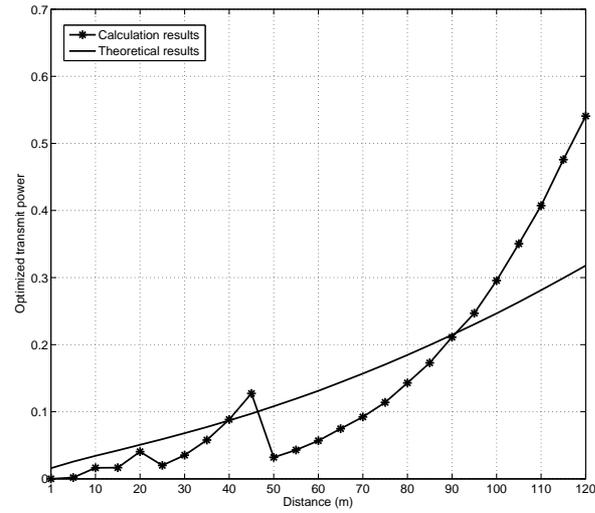


Figure 4.8: The optimized transmit powers vs. transmission distance (theoretical and practical).

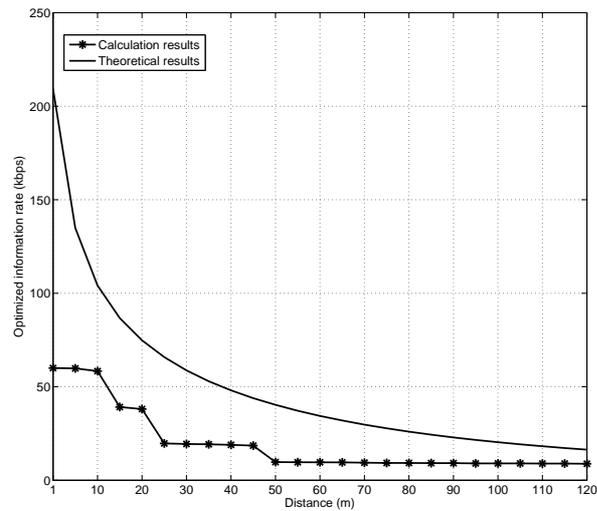


Figure 4.9: The optimized information rate vs. transmission distance.

#### 4.6.2 Energy Minimization with Fixed Average Power Constraint

This subsection presents the optimization results of the energy minimization with fixed average power constraint. The average power constraint  $\bar{P} = 0.2894$  W, and the circuit

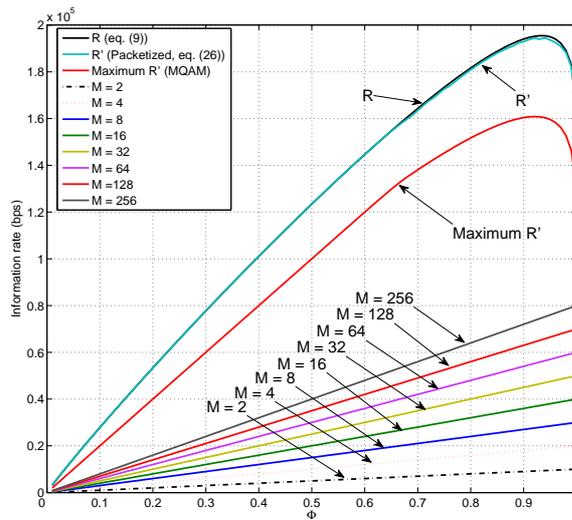


Figure 4.10: The information rate at distance  $d = 1$  m.

power consumption  $P_c = 0.2884$  W.

Figs. 4.10 to 4.12 show the achievable information rates at distances 1 m, 40 m and 70 m, respectively. Note that the packet length  $L_L$  and target bit error rate  $P_b$  have been optimized for any specific  $\Phi$ . The bound  $R$  is achieved using (4.9). The packetized/ARQ information bound is obtained using (4.28), and the information rates are achieved using (4.31). The maximum  $R$  is achieved through (4.31) by searching the constellation size  $M$  up to  $10^{12}$ . Here we have allowed non-integer  $M$ .

As shown in Fig. 4.10, at very short distances, a large  $M$  and a large  $\Phi$  will maximize the information rate and thereby minimize the energy per bit under fixed average power constraint. Moreover, Figs. 4.11 and 4.12 show that, as distance increases, the parameters  $(M^*, \Phi^*, L_L^*)$  will decrease. For example, the set  $(M^*, \Phi^*, L_L^*)$  is  $(8, 0.58, 3 \times 10^4)$  at  $d = 40$  m, while the set  $(M^*, \Phi^*, L_L^*)$  is  $(4, 0.32, 1 \times 10^4)$  at  $d = 70$  m. This trend is caused by the fact that the reliable (energy efficient) modulation and coding schemes gradually outweighs the high-speed (bandwidth efficient) modulation and coding schemes as the communication environment deteriorates. The cost we pay to save energy is the information rate. As shown in Fig. 4.13, to achieve energy

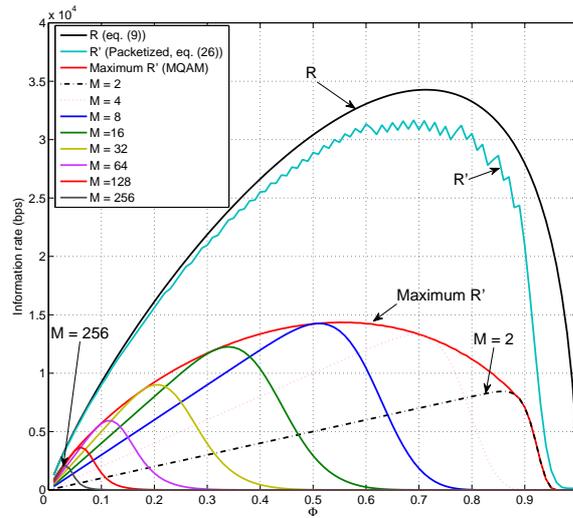


Figure 4.11: The information rate at distance  $d = 40$  m.

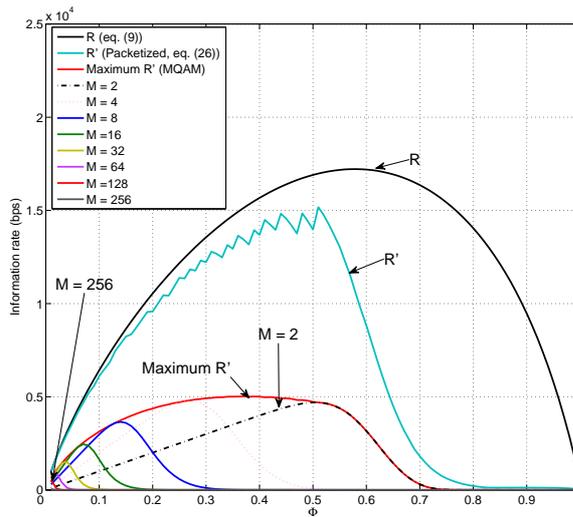


Figure 4.12: The information rate at distance  $d = 70$  m.

efficiency, the information rate will drop rapidly. At  $d = 120$  m, the  $R^*$  is only about 1 Kbps over a bandwidth of  $B = 10$  KHz.

For a given distance and a given family of modulation schemes, it is straightforward to find the optimum  $\Phi^*$  with respect to a fixed constellation size  $M$ , or the optimum

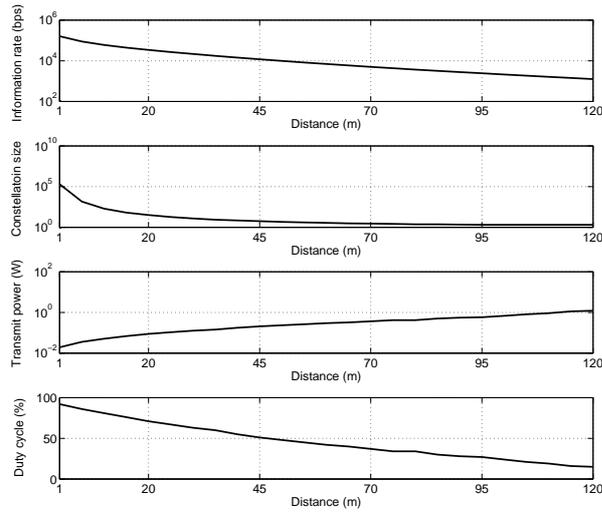


Figure 4.13: The information rate,  $M^*$ ,  $P_t^*$ , and  $\Phi^*$  of  $M$ -QAM at different transmission distances.

constellation size  $M^*$  with respect to a fixed  $\Phi$  by taking partial derivatives of (4.31) and setting them to zero. For instance, the  $\Phi^*$  and  $M^*$  of  $M$ -QAM at different distances are shown in Fig. 4.13. From Fig. 4.13, we can see that, under a strict average power constraint, the transceiver has to sacrifice its maximum possible information rate by slashing its duty cycle to satisfy this constraint. For example, as shown in Fig. 4.13, at  $d = 100$  m, the duty cycle should be as low as 24.6% to provide an appropriate amount of transmit power  $P_t$  to guarantee a reasonably low  $P_b$  even when the constellation size  $M = 2$ . This limits the maximum possible information rate  $R' \leq B \log_2 M\Phi = 2.46$  Kbps.

#### 4.6.2.1 The influence of average power

Fig. 4.14 illustrates the impact of average power on the choice of duty cycle, minimum energy consumption per information bit, and transmit power, where the circuit power consumption is  $P_c = 0.288W$ . The average power configurations are

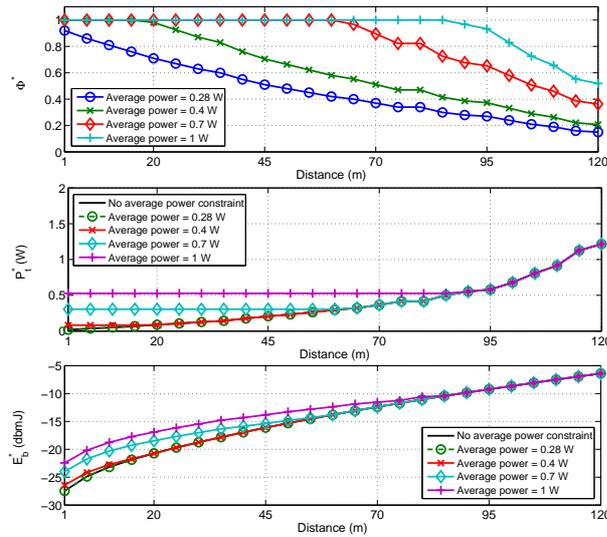


Figure 4.14: The influence of the average power on  $(\Phi^*, P_t^*, E_b^*)$  from (4.32).

$\bar{P} = 0.28, 0.4, 0.7$ , and  $1$  W. Fig. 4.14 shows that, for a given distance,  $\Phi^*$  is larger for a higher  $\bar{P}$ . That is, when  $P_t$  is fixed as  $P_t^*(d)$  (the unconstrained optimal transmit power), the transmitter has to be on for a longer period of time to meet a higher  $\bar{P}$ . However, when  $\bar{P}$  increases beyond a point when  $\Phi = 1$ , the transmit power  $P_t$  has to start increasing to maintain the average power consumption requirement  $\bar{P}$ . In this case,  $P_t > P_t^*(d)$  and the resulting minimum  $E_b$  becomes suboptimal. The sufficient average power constraint at any given distance is  $\bar{P}^*(d) = P_t^*(d)(1 + \beta) + P_c$ . That is, for a given transmission distance  $d$ , any  $\bar{P}$  above  $\bar{P}^*$  is unnecessary and not energy efficient. In fact, Fig. 4.14 shows that, under an average power constraint where  $\Phi[P_t(1 + \beta) + P_c] = \bar{P}$ , the transmit power sometimes has to be higher than the unconstrained optimum value just to maintain an unnecessarily high average power constraint.

Although a low average power constraint  $\bar{P} < P_t^*(1 + \beta) + P_c$  benefits energy efficiency, it lowers the information rate correspondingly. An example is shown in Fig. 4.15 where a higher average power obviously provides higher information rate. This is because a low  $\bar{P}$  requires the transmitter to sleep for a larger portion of the

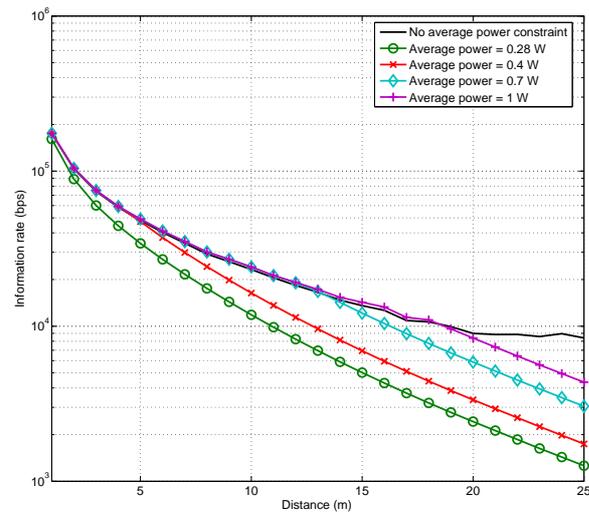


Figure 4.15: The influence of the average power on  $R^*$  from (4.32).

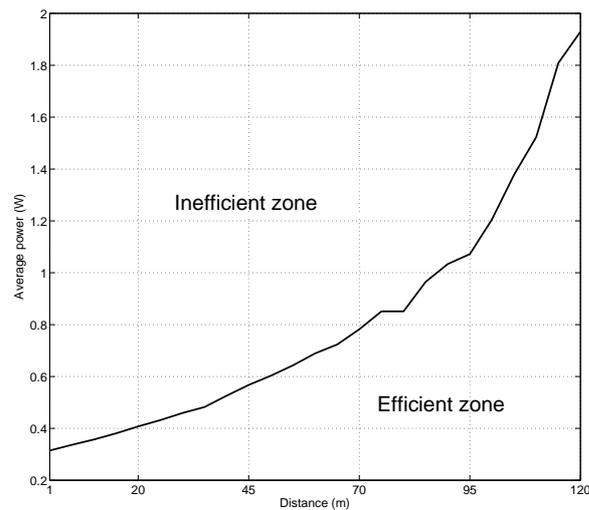


Figure 4.16: Range of efficient vs. inefficient zones for average power versus distance from (4.32).

duty cycle (small  $\Phi$ ). On the other hand, the information rate  $R$  is proportional to  $\Phi$ . Therefore, the information rate will drop with decreasing  $\bar{P}$ . Thus, the average power  $\bar{P}$  reflects the trade-off between the energy-efficiency (low  $\bar{P}$ ) and the information rate

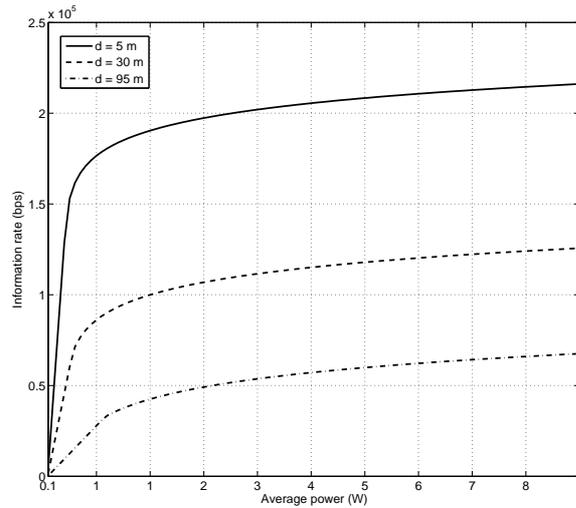


Figure 4.17: Information rate as a function of average power from (4.32).

(high  $\bar{P}$ ). However, the tradeoff of energy-efficiency and the information rate caused by  $\bar{P}$  should always be evaluated in the range  $(0, P_t^*(1 + \beta) + P_c]$ , since beyond this range the transmitter will operate in a classic inefficient trade-off between  $E_b$  (linear increase with respect to  $\bar{P}$ ) and  $R$  (logarithmic increase with respect to  $\bar{P}$ ). The efficient versus inefficient zone of  $\bar{P}$  is shown in Fig. 4.16 where the boundary is  $P_t^*(1 + \beta) + P_c$ . Fig. 4.17 gives some examples to show the trend of information rate versus average power. There clearly exists a point of  $\bar{P}$  where the slope of the curve changes. For instance, at  $d = 95$  m, this point is about  $\bar{P} = 1$  W, below which the  $R-\bar{P}$  curve increases linearly (operating in efficient zone) and above which the  $R-\bar{P}$  curve increases logarithmically (operating in inefficient zone).

### 4.6.3 Energy Minimization with Fixed Average Rate Constraint

In the case of energy minimization with fixed average rate constraint, some illustrative results are presented in Figs. 4.18 through 4.21, where the information rate constraint  $R_{const} = 20$  Kbps with the rate deviation  $\Delta_R = 100$  bps.

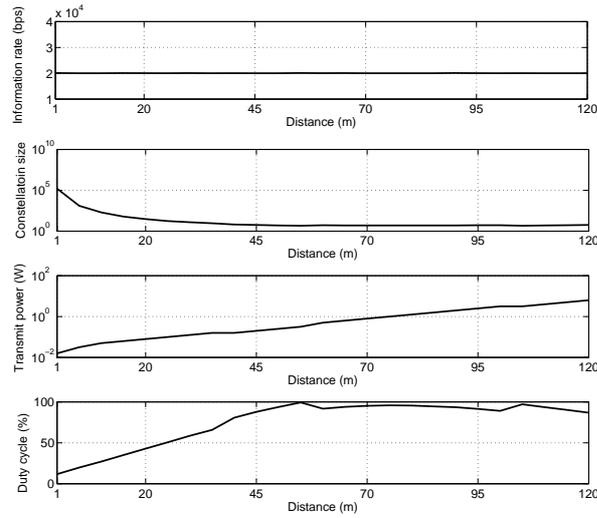


Figure 4.18:  $M^*$ ,  $P_t^*$ , and  $\Phi^*$  for  $M$ -QAM at different transmission distances ( $|R' - 20\text{Kbps}| < 100\text{bps}$ ).

Fig. 4.18 shows the optimum  $(M^*, P_t^*, \Phi^*)$  of  $M$ -QAM at different transmission distances under this average rate constraint. As the transmission distance increases,  $M^*$  decreases due to the fact that a robust modulation scheme becomes energy efficient as the communication environment worsens. In contrast to the fixed average power case,  $\Phi^*$  also increases in the fixed average rate scenario. This is because,  $\Phi^*$  is only used to control the information rate in the fixed average rate transmission. Therefore, at short distances, when large  $M^*$  is energy efficient, a small  $\Phi^*$  needs to be chosen to meet the information rate requirement. As  $M^*$  decreases,  $\Phi^*$  has to increase to maintain information rate requirement. Therefore, for fixed average rate transmissions, at short distances where  $P_c$  is comparable with  $P_t$ , a time sharing use of the channel is energy efficient; at large distances where  $P_t \gg P_c$ , a continuous use of the channel is energy efficient.

Fig. 4.19 shows the information rate and  $\bar{P}^*$  of  $M$ -QAM at different transmission distances, under average power constrained and average rate constrained cases, respectively. It is easy to see why  $\bar{P}^*$  does not change with distance in the average-

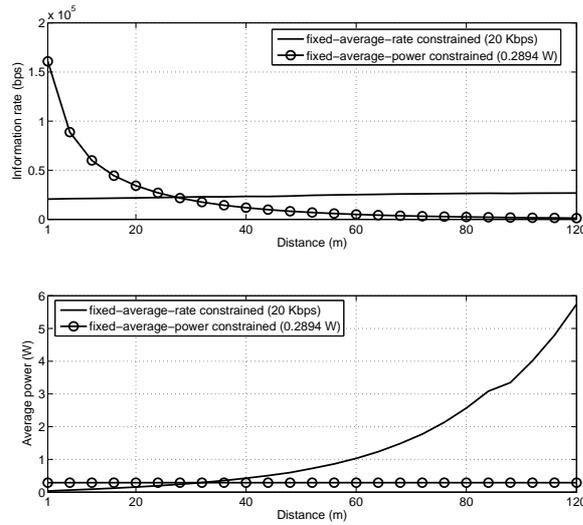


Figure 4.19: The information rate and  $\bar{P}^*$  of  $M$ -QAM at different transmission distances (fixed average power compared with fixed average rate).

power-constrained transmission case (here  $\bar{P}^* = \bar{P} = 0.2894$  W). In the fixed average rate transmission case,  $\bar{P}^*$  will increase with distance. This is directly caused by the increase in the duty cycle and the transmit power as shown in Fig. 4.18. In this particular example, average rate constrained transmission achieves a lower  $\bar{P}^*$  than average power constrained transmission when  $d \leq 35$  m. That is, for fixed average rate transmission, a high average power budget is not necessarily beneficial in terms of energy efficiency when transmission distance is below a certain value  $d_{th}$ , where  $d_{th}$  is a relative value determined by  $R_{const}$  and  $\bar{P}$ . For example, when  $R_{const} \rightarrow 0$ , the fixed average rate transmission always achieves a lower  $\bar{P}^*$  than the fixed average power transmission with  $\bar{P} = 0.2894$  W for any  $d > d_{th} = 0$ . On the other hand, although the fixed average power constrained transmission provides a low power consumption at longer distances, the provided information rate drops dramatically.

Fig. 4.20 shows the  $E_b^*$  of  $M$ -QAM at different transmission distances for the unconstrained, fixed average rate transmission, and fixed average power transmission, respectively. The  $E_b^*$  of fixed average rate constrained transmissions is always higher

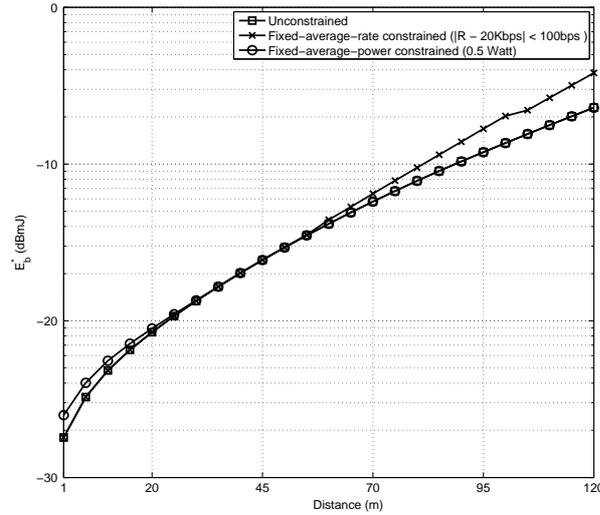


Figure 4.20: The  $E_b^*$  of  $M$ -QAM at different transmission distances.

than or equal to the unconstrained optimal  $E_b^*$ . From Fig. 4.20,  $R_{const} = 20$  Kbps approaches the global energy efficient information rate in the transmission range  $d \leq 55$  m. Compare this with Fig. 4.18, this is the range where  $\Phi^*$  increases linearly. In this distance range, the transceiver can achieve the desired information rate yet obtain the overall minimum energy consumption by increasing  $\Phi^*$ . Beyond this range, since  $\Phi^*$  cannot be further increased, extra energy has to be wasted to maintain this information rate. Therefore, a reasonable target information rate should be set according to the transmission distance, or in general, the communication environment. In this case, the target information rate  $R_{const} = 20$  Kbps is suitable in terms of energy efficiency for a wireless network with average distance  $\bar{d} \leq 55$  meters. When  $\bar{d}$  is larger than this range, the target information rate should be reduced to save energy. On the other hand, for a given target information rate, this algorithm gives us a target node density that is energy efficient.

A fixed average power constrained case is also shown in Fig. 4.20 with  $\bar{P} = 0.5$  W. At short distances, the minimized  $E_b$  is suboptimal due to the superfluous average power. However, as  $d$  increases, the minimized  $E_b$  becomes the global minimum value

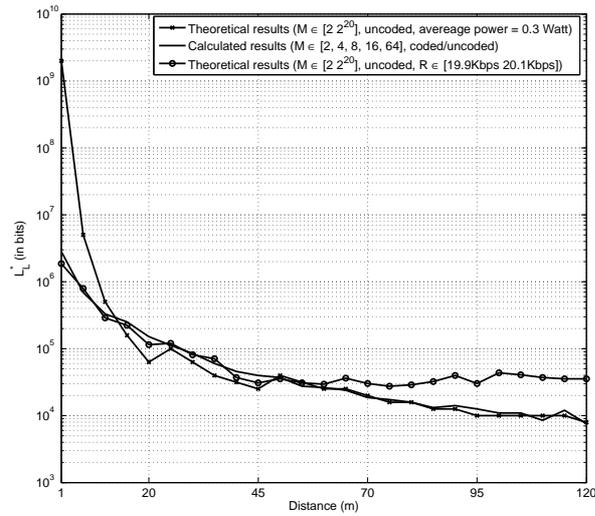


Figure 4.21: The  $L_L^*$  of  $M$ -QAM at different transmission distances for the unconstrained, fixed-rate-constrained and fixed-power-constrained cases.

since there is no more wasted  $\bar{P}$ . Fig. 4.20 indicates that, in general, fixed average rate transmissions and duty cycle are energy efficient at short transmission distances, while fixed average power transmissions and duty cycle are energy efficient at large transmission distances.

Fig. 4.21 presents the optimum  $L_L^*$  of  $M$ -QAM at different transmission distances for unconstrained and fixed average rate constrained transmissions, respectively. The calculated results of the optimal packet length in Section 4.3 (unconstrained) are also provided as a comparison. Note that the calculated results are obtained when the modulation schemes are restricted within coded/uncoded BPSK, QPSK, 16-QAM, and 64-QAM.

We can see that the theoretical  $L_L^*$  in the unconstrained transmissions is almost the same as the optimum packet lengths given by the calculations. The differences between the calculated and theoretical results are mainly caused by the limited resolutions of constellation sizes in the calculation. However, in the case of the fixed average rate

( $R_{const} = 20$  Kbps) transmission, the optimum packet length (denoted by  $\hat{L}_L^*$ ) is smaller than the overall optimum packet length  $L_L^*$  at short distances, while larger than  $L_L^*$  at large distances. This is caused by the requirement of maintaining throughput. At short distances ( $d < 15$  m in this case), nodes should use larger constellation sizes to save energy. However, a larger constellation size will increase the throughput beyond the acceptable range. Thus, other parameters must be adjusted to shrink the information rate. These parameters include: adopting lower duty cycle and using shorter packet. Therefore, the transmitter adopts a low duty cycle at very short distances as shown in Fig. 4.18 and a packet length  $\hat{L}_L^* < L_L^*$ . On the other hand, as distance increases, the effect of increasing  $\Phi^*$  to maintain the information rate starts to fade ( $\Phi^* \approx 1$ ), the  $\hat{L}_L^*$  then needs to be larger than the  $L_L^*$  to increase the effective information rate by amortizing the overhead over a larger number of information bits.

## 4.7 Summary

In this chapter, we investigated the energy consumption minimization problem for a single link in a wireless network. Specifically, we proposed a generic model for energy consumption per information bit, considering circuit power, packetization, overhead and duty cycle. We have considered the unconstrained, fixed average power, and fixed average rate cases.

For the unconstrained case, the results reveal that when transmission distance is short, a system adopting large packet length, small target bit error probability, and high bandwidth-efficient modulation schemes (e.g., high-order uncoded QAM) is more energy efficient. On the other hand, when transmission distance is large, a system using small packet length, large target bit error probability, and high energy efficient modulation schemes (e.g., coded BPSK) is energy efficient. Moreover, as transmission distance increases, a flattening of the optimum values of packet length and target bit error probability is observed.

In the fixed average power case, we conclude that the minimization of energy consumption per information bit is equivalent to the maximization of information rate. At short distances, large constellation sizes and large duty cycle are energy efficient, while the optimum constellation size and duty cycle both decrease with distance. This indicates that, within the limits of average power constraint, bandwidth efficient modulations and continuous use of channels are energy efficient at short distances, while robust modulations and duty cycling are energy efficient at large distances. The cost associated with maintaining a fixed average power is the decrease of information rate with distance.

In the fixed average rate case, at short distances, large constellation sizes and small duty cycle are energy efficient. As transmission distance increases, the optimum constellation size decreases and optimum duty cycle increases to get data through while minimizing the energy consumption. That is, under a strict average rate constraint, bandwidth efficient modulations and duty cycling are energy efficient at short distances, while robust modulations and continuous use of the channels are energy efficient at long distances.

## **5 Joint Optimization of Physical Layer and Link Layer in Impulse Radio Ultra-Wideband (IR-UWB) Communication Systems**

In this chapter, we extend the work of the previous chapter by considering an Impulse Radio Ultra Wideband (IR-UWB) based wireless network. Using detailed models of typical IR-UWB transmitter and receiver structures, we model the energy consumption per information bit in a single link of an IR-UWB system, considering packet overhead, retransmissions, and a Nakagami- $m$  fading channel.

Using this model, we minimize the energy consumption per information bit by finding the optimum packet length and the optimum number of RAKE fingers at the receiver for different transmission distances, using Differential Phase-Shift Keying (DBPSK), Differential Pulse-Position Modulation (DPPM) and On-Off Keying (OOK), with coherent and non-coherent detection. Symbol repetition schemes with hard decision (HD) combining and soft decision (SD) combining are also compared in this chapter.

Our results show that at very short distances, it is optimum to use large packets, OOK with non-coherent detection, and HD combining, while at longer distances, it is optimum to use small packets, DBPSK with coherent detection, and SD combining.

The optimum number of RAKE fingers are also found for given transmission schemes.

## 5.1 Introduction

Energy consumption is a very important design consideration in any IR-UWB based system. Unlike in traditional communications systems, where transmit power can be flexibly adjusted to minimize the energy consumption [31][52], there is a strict limit on the effective isotropic radiated power (EIRP) in IR-UWB systems due to their overlay nature. Regulations mandate that the spectrum of the signal be limited to  $-41.25$  dBm/MHz [32]. Since the IR-UWB system needs to operate at or near this limit to achieve a reasonable range, the traditional optimization techniques, which mainly operate by adjusting the transmit power, cannot be used for IR-UWB systems. However, there are other parameters of the IR-UWB system that can be adjusted, such as the number of RAKE fingers, the packet length, the modulation scheme, the detection scheme, and the coding or spreading scheme.

In IR-UWB communications, the channel delays are often resolvable due to the narrow width of the IR-UWB pulse. Therefore, a RAKE receiver structure can achieve considerable diversity gain [93][94]. Another important utility of the RAKE receiver structure is that it can increase the collection of the transmitted power through multiple paths. The diversity gain and collected power will be increased by adding more RAKE fingers (correlator taps), which in turn will increase the power consumption of the receiver. Therefore, the tradeoff between the diversity gain as well as the power collection and the power consumption at the receiver must be evaluated.

Packet length is another important factor that influences the energy consumption of a communication link. A long packet will increase the packet error probability; thereby increasing the average number of transmissions in an automatic-repeat-request (ARQ) system. On the other hand, a short packet will lower the system efficiency due to the packet overhead. Thus, an optimum packet length should be found to minimize the

energy consumption.

Binary modulation schemes, such as DBPSK, OOK, and DPPM, are usually used in IR-UWB systems due to their simplicity and good performance. Among the three modulation schemes considered in this chapter (DBPSK, OOK, and DPPM), DBPSK is the most robust, but also the most energy consuming. Compared with DBPSK, OOK requires less energy to transmit each bit, but has a lower performance. The performance of DPPM is between DBPSK and OOK. The comparison and evaluation of these modulation schemes are important for the design of energy-efficient IR-UWB systems.

In our previous work, we proposed an energy consumption model of an IR-UWB based communication link, and we compared the energy consumption features of DBPSK and OOK modulations with coherent/noncoherent detections schemes [53]. Although this work builds a solid background for the work in this chapter, the studies in [53] need to be improved. First, in the channel model in [53], Rayleigh distribution is used to depict the small-scale fading feature of an IR-UWB channel. This channel model, however, is outdated. Second, comprehensive analysis on the energy consumption characteristics of an IR-UWB communication link is missing. Third, an extensive comparison of practical schemes is still desired. In this chapter we address the above issues in detail.

We provide detailed power consumption models of a typical IR-UWB transmitter and both coherent and noncoherent receivers. The optimization model considers these detailed power consumption models as well as the packet structure and the ARQ procedure. Using this model we optimize packet length and the number of RAKE fingers at different transmission distances for DBPSK, OOK, and DPPM, with both coherent (CO) and noncoherent (NC) detection. Moreover, in IR-UWB systems, to increase the effective energy per bit, repetition coding schemes are commonly used. At the receiver, hard decision (HD) based combining or soft decision (SD) based combining may be used. HD combining provides relatively low performance but it can be operated using a low-power, one bit analog-to-digital converter (ADC). On the other hand, SD

combining provides good performance while demanding a high-resolution ADC and memory units. The tradeoffs between these combining methods are also evaluated in this chapter.

The remainder of this chapter is organized as follows. Section 5.2 introduces the packet structure, transceiver power model, and channel model used in this work. In Section 5.3, after deriving a lower bound on the energy consumption per information bit in IR-UWB systems, we minimize the energy consumption per information bit over packet length and number of RAKE fingers. Numerical results are presented in Section 5.4. Section 5.5 concludes this chapter.

## 5.2 System and Channel Models

We consider an IR-UWB system with a symbol repetition scheme. The coding rate  $R_c = 1/N_p$ , where  $N_p$ , which is an odd number, is the coding parameter. Moreover, in order to avoid inter symbol interference (ISI), the maximum pulse rate is limited. Also, perfect knowledge of the channel is assumed at the receiver.

### 5.2.1 IR-UWB Transceiver Power Consumption Model

A typical IR-UWB transmitter and a typical IR-UWB receiver with four RAKE fingers and maximal ratio combining (MRC) are shown in Fig. 5.1. When DBPSK, DPPM, and OOK are used at the transmitter, the power consumption of the transmitter can be modeled as

$$P_t = P_{\text{SYN}} + E_p R_p, \quad (5.1)$$

where  $E_p$  is the fixed energy per pulse and  $R_p$  is the pulse rate. The pulse rate  $R_p = \rho_t R_b$ , where  $\rho_t = 1$  for DBPSK and DPPM,  $\rho_t = 0.5$  for OOK, and  $R_b$  is the bit rate. We have assumed that an information bit may be 0 or 1 with equal probability. Further-

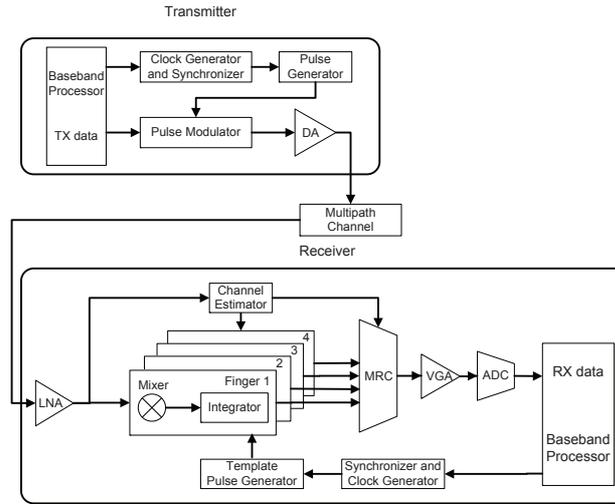


Figure 5.1: The transmitter and receiver structure in an IR-UWB system.

more,  $P_{\text{SYN}}$  represents the power consumption of the transmitter components that are independent from the data transmission, namely the clock generator and synchronizer.

In our model, the power consumption of an IR-UWB transmitter, as described by equation (5.1), is grouped into two parts: the power consumption from the circuit components that are not related to pulse generating ( $P_{\text{SYN}}$ ), and the power consumption from the ones that are related to pulse generating ( $E_p R_p$ ). That is,  $E_p R_p$  includes the power consumptions of the pulse generator, pulse modulator and digital amplifier (DA), while  $P_{\text{SYN}}$  is simply the power consumption of the clock generator and synchronizer.

As shown in Fig. 5.1, the power consumption of an IR-UWB receiver can be modeled as

$$P_r = M P_{\text{COR}} + \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}} + \rho_r (P_{\text{GEN}} + P_{\text{SYN}} + P_{\text{EST}}), \quad (5.2)$$

where  $P_{\text{COR}}$ ,  $P_{\text{ADC}}$ ,  $P_{\text{LNA}}$ ,  $P_{\text{VGA}}$ ,  $P_{\text{GEN}}$ ,  $P_{\text{SYN}}$ , and  $P_{\text{EST}}$  respectively represent: the power consumptions of one correlator branch (mixer and integrator), the analog-to-digital converter (ADC), the low noise amplifier (LNA), the variable gain amplifier (VGA), the pulse generator, the synchronizer, and the channel estimator.  $M$  represents the number of RAKE fingers at the receiver.  $\rho_r$  is determined by the receiver structure. That is,

$\rho_r = 1$  for coherent demodulation and  $\rho_r = 0$  for noncoherent demodulation. This is because for a noncoherent UWB receiver, the pulse generator, clock generator, synchronizer, and channel estimator are not necessary. Moreover,  $\rho_c = 1$  for SD combining and  $\rho_c = 0$  for HD combining. For SD combining, a 5-bit ADC is assumed [95], while for HD combining, the power consumption of the ADC (one-bit ADC) is assumed to be negligible.

At the receiver, we consider an IR-UWB receiver that is able to choose the coherent or noncoherent demodulation after the signal passed through RAKE fingers and MRC. When the IR-UWB receiver uses the coherent detection, the received signal will pass through a matched filter and a template pulse needs to be generated to configure the matched filter. When the IR-UWB receiver adopts the noncoherent detection, neither a matched filter nor a template pulse is needed. The received signal will be either correlated with the previously received signal, or simply be multiplied by itself (envelop detection). A differential modulation scheme can cooperate with either the coherent detection or noncoherent detection. For example, a DBPSK modulated signal can be noncoherently detected by a correlation with the previously received signal so that only the difference between the two signals will be at output of the ADC, or each DBPSK modulated signal can be coherently detected individually through a matched filter and the difference between two adjacent bits can be measured in the digital domain. In general, the coherent detection provides a better performance than the noncoherent detection in terms of bit error probability. However, the coherent detection requires more circuit components (template pulse generator, synchronizer, and etc.) and thereby consumes more power than the noncoherent detection.

As with the transmitter, we also group the power consumption of an IR-UWB receiver into two parts: the power consumption of the circuit components that are not related to the detection schemes,  $MP_{\text{COR}} + \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}}$ , and the power consumption of the circuit components that are related to the detection schemes,  $\rho_r (P_{\text{GEN}} + P_{\text{SYN}} + P_{\text{EST}})$ . That is,  $MP_{\text{COR}} + \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}}$  represents the power consump-

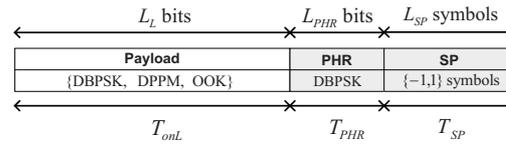


Figure 5.2: Data packet structure.

tion of  $M$  correlator branches, the ADC, the LNA, and the VGA. These components need to be active whether coherent or noncoherent detection is used at the receiver. However, the pulse generator, the synchronizer, and the channel estimator are only active during coherent detection, where a template pulse has to be generated to correlate with the received pulse and the channel information is required. During noncoherent detection, the pulse generator, the synchronizer and the channel estimator are not necessary because no template pulse is needed and the received signal pulse is only correlated with the previously received pulse. The power consumption of the MRC is not considered, since a MRC is simply an adder.

### 5.2.2 Packet Structure

The packet structure consists of three components: synchronization preamble (SP), PHY-header (PHR), and payload. We assume that there are  $L_L$  bits in the payload,  $L_{PHR}$  bits in the PHR, and  $L_{SP}$  symbols in the SP. Correspondingly, the time durations to deliver the payload, the PHR, and the SP are denoted by  $T_{onL}$ ,  $T_{PHR}$ , and  $T_{SP}$ , respectively. The energy consumption to transmit a packet once is the summation of two parts:  $E_O$ , the energy consumed on delivering the SP and PHR, and  $E_L$ , the energy consumed on the payload.

We assume that the synchronization preamble has values  $\{-1, 1\}$ . Moreover, the PHR is modulated using DBPSK and always received coherently. This is to ensure that the PHR is transmitted using the modulation and detection schemes with the highest performance, since the PHR carries important information. Also for the sake of sim-

plicity, we assume that the PHR is coded in the same manner as the payload. Therefore, the overhead energy consumption is

$$\begin{aligned} E_O &= E_O^{(\text{TX})} + E_O^{(\text{RX})} \\ &= (L_{\text{SP}} + L_{\text{PHR}}/R_c)E_p + P_{\text{SYN}}T_O + P_rT_O, \end{aligned} \quad (5.3)$$

where  $L_{\text{SP}}$  is the number of SP symbols,  $L_{\text{PHR}}$  is the number of PHR bits, and  $T_O = T_{\text{SP}} + T_{\text{PHR}} = (L_{\text{SP}} + \frac{L_{\text{PHR}}}{R_c})/R_{\text{base}}$ , where  $R_{\text{base}}$  is the fixed base data rate. In this work, a frequency selective slow fading channel is assumed. Therefore, the channel estimation block consumes  $P_{\text{EST}}$  amount of power only during the reception of the overhead.

The energy consumption for the payload can be modeled as

$$E_L = E_L^{(\text{TX})} + E_L^{(\text{RX})}, \quad (5.4)$$

where  $E_L^{(\text{TX})}$  and  $E_L^{(\text{RX})}$  represent the energy consumption to transmit/receive the payload containing  $L_L$  information bits, respectively. For  $E_L^{(\text{TX})}$ , we have

$$E_L^{(\text{TX})} = \rho_t E_p L_L / R_c + P_{\text{SYN}} T_{\text{onL}}, \quad (5.5)$$

where  $T_{\text{onL}} = L_L / R_b R_c$ , is the time duration to transmit the payload containing  $L_L$  bits, and  $R_c$  is the coding rate.

The energy consumption to receive  $L_L$  information bits is given by

$$\begin{aligned} E_L^{(\text{RX})} &= \rho_t (M P_{\text{COR}} + \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}}) T_{\text{onL}} \\ &\quad + \rho_r (P_{\text{GEN}} + P_{\text{SYN}}) T_{\text{onL}}. \end{aligned} \quad (5.6)$$

The receiver does not consume power on channel estimation during the reception of information bits when using either coherent or noncoherent detection, since the channel information has been estimated during the reception of the overhead bits.

### 5.2.3 Channel Model

The channel model consists of a path loss model and a frequency selective fading model. In this work, we focus our research on the frequency range from 3-10 GHz.

### 5.2.3.1 Path Loss Model

The UWB path loss model is both distance and frequency dependent and can be modeled as [46]

$$G_d = G_0 - 20(\kappa + 1) \log_{10} \left( \frac{f}{f_c} \right) - 10n \log_{10} d - 3, \quad (5.7)$$

where  $d$  is the transmission distance,  $G_0$  is the path gain at the reference distance ( $d = 1$  m),  $n$  is the path loss exponent,  $f$  is the UWB transmission center frequency,  $f_c$  is the reference frequency, and  $\kappa$  is the frequency dependency decaying factor. Both  $G_d$  and  $G_0$  are expressed in dB.

### 5.2.3.2 Frequency Selective Fading

In an IR-UWB system, the transmitted signal inevitably encounters frequency selective fading. The baseband channel impulse response of a frequency selective fading channel in UWB systems consists of clusters and rays and can be represented as [46]

$$c(t) = \sum_{l=0}^L \sum_{k=0}^K \alpha_{k,l} e^{-\theta_{k,l}} \delta[t - T_l - \tau_{k,l}], \quad (5.8)$$

where  $\theta_{k,l}$  follows a uniform distribution (over  $[0, 2\pi]$ ) and  $\alpha_{k,l}$  is the amplitude gain of the  $k$ th ray in the  $l$ th cluster.  $L$  and  $K$  represent the number of clusters and rays, respectively.  $T_l$  is the arrival time of the  $l$ th cluster, and  $\tau_{k,l}$  is the arrival time of the  $k$ th ray in the  $l$ th cluster.  $T_l$  and  $\tau_{k,l}$  follow the following independent interarrival exponential probability density functions, the details of which can be found in [46].

The average power gain of the  $k$ th ray in the  $l$ th cluster is modeled as

$$E[\alpha_{k,l}^2] = E[\alpha_{0,0}^2] e^{-T_l/\Gamma} e^{-\tau_{k,l}/\gamma}, \quad (5.9)$$

where  $\Gamma$  and  $\gamma$  are power-delay time constraints for the clusters and rays, respectively.  $E[\alpha_{0,0}^2]$  is the average power gain of the first ray of the first cluster, which is expressed as

$$E[\alpha_{0,0}^2] = \frac{G_d}{\gamma\lambda}. \quad (5.10)$$

The parameter  $\alpha_{k,l}$  follows a Nakagami distribution described as

$$p(\alpha_{k,l}) = \frac{2}{\Gamma(m)} \left( \frac{m}{E[\alpha_{k,l}^2]} \right)^m \alpha_{k,l}^{2m-1} e^{-\frac{m\alpha_{k,l}^2}{E[\alpha_{k,l}^2]}}, \quad (5.11)$$

where  $m > 0.5$  is the Nakagami m-factor and  $\Gamma(m)$  is the gamma function.

## 5.3 Link Energy Minimization

### 5.3.1 Lower Bound on Average Energy Consumption per Information Bit

#### 5.3.1.1 Lower bound based on channel capacity

The transmit power is strictly limited in IR-UWB systems to avoid interfering with pre-existing communication systems. In the following analysis, we assume the transmit power is a constant denoted by  $P_{\text{tx}}$ . Considering the power consumption at the transmitter and receiver and the channel capacity, the lower bound of energy consumption per information bit is modeled as

$$E_b \geq \frac{P_t + P_r}{B \log \left( 1 + \frac{\sum_{i=1}^M |\alpha_i|^2 P_{\text{tx}}}{G_d \sigma^2} \right)} = \frac{P_{\text{SYN}} + E_p R_p + M P_{\text{COR}} + P_{\text{CNST}}}{B \log \left( 1 + \frac{\sum_{i=1}^M |\alpha_i|^2 P_{\text{tx}}}{G_d \sigma^2} \right)}, \quad (5.12)$$

where  $P_{\text{CNST}} = \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}} + \rho_r (P_{\text{GEN}} + P_{\text{SYN}})$  represents the receiver power consumption that is independent of the number of RAKE fingers, and  $|\alpha_i|^2$  represents the average power gain of the  $i$ th selected path. Note that the power consumption of the channel estimator is not considered, since the channel estimator is not involved in the actual data communication. We assume that the positioning of the RAKE fingers is ideal and therefore  $|\alpha_i|^2$  are the largest  $M$  values of  $E[|\alpha_{k,l}|^2]$  from (5.9).  $M$  is the number of RAKE fingers, and  $B$  is the signal bandwidth. Correspondingly,  $B \log(1 + \sum_{i=1}^M |\alpha_i|^2 P_{\text{tx}} / G_d \sigma^2)$  represents the channel capacity [96].

Equation (5.12) can be minimized through proper selection of the number of RAKE fingers,  $M$ . As transmission distance increases, the optimum number of RAKE fingers increases and eventually converges to a certain value. In other words, when received SNR approaches zero, there exists an optimum number of RAKE fingers that minimizes the energy consumption in IR-UWB systems. This optimum number is only a function of the power delay profile of the channel and the power consumption values of the components of the transceiver.

Removing the integer constraint on  $M$ , we can determine this convergence by finding the discrete derivative of the right hand side of (5.12) with respect to  $M$  and setting the resulting equation to zero, i.e.,

$$(\log_2 e) \frac{\sum_{i=1}^{M^*} |\alpha_i|^2}{|\alpha_{M^*}|^2} \left( 1 + \frac{\sum_{i=1}^{M^*} |\alpha_i|^2 P_{\text{tx}}}{G_d \sigma^2} \right) = \frac{P_{\text{CNST}} + P_{\text{SYN}} + R_p \rho_t E_p}{P_{\text{COR}}} + M^*, \quad (5.13)$$

where  $e$  is the natural number. As distance increases,  $M^*$  will eventually converge to a particular value as  $\sum_{i=1}^{M^*} |\alpha_i|^2 P_{\text{tx}} / G_d \sigma^2 \rightarrow 0$ . The convergence value of  $M^*$ , which is denoted by  $M_{\text{CONV}}^*$ , can be found from:

$$(\log_2 e) \frac{\sum_{i=1}^{M_{\text{CONV}}^*} |\alpha_i|^2}{|\alpha_{M_{\text{CONV}}^*}|^2} - \frac{P_{\text{CNST}} + P_{\text{SYN}} + R_p \rho_t E_p}{P_{\text{COR}}} = M_{\text{CONV}}^*. \quad (5.14)$$

In general, there is no closed form solution for (5.14), since the distribution of  $|\alpha_i|^2$  directly determines the solvability of this equation and  $M$  has to be chosen in the positive discrete domain. For the doubly exponential decay of  $E[|\alpha_{k,l}|^2]$ ,  $M_{\text{CONV}}^*$  exists and can always be easily found through an exhaustive search.

### 5.3.1.2 Lower bound based on data rate

During the modeling in (5.12), the data rate is bounded by the channel capacity, which implies a linear channel (such as an AWGN channel), infinitely long codewords, and arbitrarily low bit error rates (i.e., no retransmissions). Most of the above assumptions do not hold in practical communication systems. Thus, the practical data rate is usually much lower than the channel capacity. The lower bound from (5.12) is, therefore, a very

loose lower bound. In the following analysis, instead of using channel capacity, we derive the lower bound energy consumption per information bit by using achievable data rates. In particular, the data rate needs to consider the penalty caused by retransmission and packetization.

In this work, we only consider binary modulation schemes. Also, to avoid ISI, the maximum pulse rate is limited by the maximum excess delay of the multipath channel,  $D_s$ . That is, the maximum pulse rate is  $1/D_s$ . In addition, considering the impacts of packetization, retransmission and overheads, the lower bound of the energy consumption per information bit can be further tightened as

$$E_b \geq \left( \frac{P_r + P_{\text{SYN}} + \rho_t E_p R_p}{1/D_s} \right) N \frac{T_{\text{on}} + 2T_{\text{IPS}} + T_{\text{ACK}}}{T_{\text{onL}}}, \quad (5.15)$$

where  $N$  is the total number of transmissions.  $(T_{\text{on}} + 2T_{\text{IPS}} + T_{\text{ACK}})/T_{\text{onL}}$  denotes the rate penalty caused by packetization, where  $T_{\text{IPS}}$  denotes inter packet space and  $T_{\text{ACK}}$  represents the time duration the transmitter listens for acknowledgement from the receiver. The detailed formulas of  $T_{\text{IPS}}$  and  $T_{\text{ACK}}$  can be found in the following section. The above parameters are determined by the detailed packet structure, channel conditions, and modulation schemes. Although (5.15) implies an ideal wideband channel with no multipath and omits the possible energy losses due to circuit start-up, this model tightens the bound from (5.12) and better represents practical scenarios since both  $N$  and  $(T_{\text{on}} + 2T_{\text{IPS}} + T_{\text{ACK}})/T_{\text{onL}}$  are greater than or equal to 1.

### 5.3.2 Practical Average Energy Consumption Per Information Bit

Although (5.15) provides a lower bound on the energy consumption per information bit, it does not consider many practical issues. For example, it does not consider the energy spent on the packet overhead and listening. In practice, we need to consider the detailed procedure of transmitting one packet instead of one bit [52].

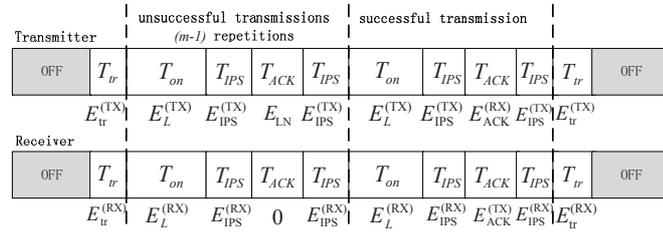


Figure 5.3: The transmission and reception of one packet using  $N$  total transmissions.

### 5.3.2.1 Energy Consumption per Packet with Retransmissions

To guarantee the successful reception of one packet, an automatic repeat request (ARQ) protocol is used. A delivery procedure involving  $N - 1$  retransmissions is shown in Fig. 5.3. The inter packet space (IPS) is denoted by  $T_{IPS}$ . The power consumption during  $T_{IPS}$  is mainly due to the clock generator and synchronizer. Therefore, the corresponding energy consumption at the transmitter is  $E_{IPS}^{(TX)} = P_{SYN}T_{IPS}$ , while the receiver consumes  $E_{IPS}^{(RX)} = \rho_r P_{SYN}T_{IPS}$ .

We assume that before transmission or reception of a packet, the transmitter and receiver spend  $T_{tr}$  amount of time to go from the off (sleep) state to an on (active) state. We refer to this time duration as the “transient session”. During the transient session, the transmitter consumes  $E_{tr}^{(TX)} = P_{SYN}T_{tr}$  amount of energy to start the front end clock generator and synchronizer. Similarly, the receiver consumes  $E_{tr}^{(RX)} = \rho_r P_{SYN}T_{tr}$ .

$T_{on}$  is the time duration for the transmission of one packet. That is

$$T_{on} = T_{SP} + T_{PHR} + T_{onL} = \left( L_{SP} + \frac{L_{PHR}}{R_c} \right) / R_{base} + \frac{L_L}{R_b R_c}. \quad (5.16)$$

The energy consumptions at the transmitter and receiver during  $T_{\text{on}}$  are

$$\begin{aligned} E^{(\text{TX})} &= E_L^{(\text{TX})} + E_O^{(\text{TX})}, \\ E^{(\text{RX})} &= E_L^{(\text{RX})} + E_O^{(\text{RX})}. \end{aligned} \quad (5.17)$$

where  $E_L^{(\text{TX})}$ ,  $E_O^{(\text{TX})}$ ,  $E_L^{(\text{RX})}$  and  $E_O^{(\text{RX})}$  are given in (5.3), (5.4), (5.5) and (5.6), respectively.

$T_{\text{ACK}}$  is the time period when the transmitter listens for an acknowledgement. We set  $T_{\text{ACK}} = T_O$ . Overall, the definitions of the energy consumptions within one transmission are summarized as follows

$$\begin{aligned} E_{\text{IPS}} &= 2E_{\text{IPS}}^{(\text{TX})} + 2E_{\text{IPS}}^{(\text{RX})}, \\ E_{\text{LN}} &= \rho_r P_{\text{SYN}} T_{\text{ACK}}, \\ E_{\text{TRAN}} &= 2E_{\text{tr}}^{(\text{TX})} + 2E_{\text{tr}}^{(\text{RX})}, \\ E_{\text{ACK}}^{(\text{RX})} &= P_r T_{\text{ACK}}, \\ E_{\text{ACK}}^{(\text{TX})} &= (L_{\text{SP}} + L_{\text{PHR}}/R_c)E_p + P_{\text{SYN}} T_{\text{ACK}}, \\ E^{(\text{TX})} &= E_L^{(\text{TX})} + E_O^{(\text{TX})}, \\ E^{(\text{RX})} &= E_L^{(\text{RX})} + E_O^{(\text{RX})}, \end{aligned} \quad (5.18)$$

where  $E_{\text{IPS}}$  is the total energy consumed by the receiver and the transmitter in IPSs within one transmission.  $E_{\text{TRAN}}$  is the total energy consumption of the receiver and the transmitter during the transient sessions. In both IPSs and transient sessions, only the frequency synthesizers consume energy.

$E_{\text{LN}}$  denotes the energy consumption of the transmitter on listening to the media for the ACK from the receiver in the first  $N - 1$  unsuccessful transmissions. Therefore,  $E_{\text{LN}}$  is the energy consumption of idle listening during  $T_{\text{ACK}}$ .  $E_{\text{ACK}}^{(\text{RX})}$ ,  $E_{\text{ACK}}^{(\text{TX})}$  are the energy consumption of the transmitter for receiving the ACK and the energy consumption of the receiver for transmitting the ACK. In this work, we assume the ACK message is simply a packet containing only the PHR and the SP. Since the PHR and the SP always consist of  $\{-1, 1\}$  symbols and only coherent detection with SD combining is used,  $E_{\text{ACK}}^{(\text{RX})}$  and  $E_{\text{ACK}}^{(\text{TX})}$  are constant for a given repetition coding scheme.

The decomposition of the energy consumption during each packet transmission session has been shown in Section 5.3.2.1. Therefore, the average energy consumption for successful delivery of a packet can be expressed as

$$E = (E^{(\text{TX})} + E^{(\text{RX})} + E_{\text{LN}} + E_{\text{IPS}})N - E_{\text{LN}} + E_{\text{TRAN}} + E_{\text{ACK}}^{(\text{TX})} + E_{\text{ACK}}^{(\text{RX})}, \quad (5.19)$$

where  $N$  is the average number of transmissions/receptions required to successfully deliver one packet. The average number of transmissions  $N = 1/(1 - P_b)^{L_L}$ , where  $P_b$  is the average BEP. Note that  $(1 - P_b)^{L_L}$  is the probability that a packet is received correctly. As shown in the following subsection, the average BEP  $P_b$  is closely related to the modulation type, detection schemes, repetition coding/combing schemes, and number of Rake fingers.

### 5.3.2.2 Average BEP over Independent Nakagami Fading Channels

The average BEP can be obtained utilizing the characteristic function of the pdf of the output SNR after the MRC [87][97][98]. The instantaneous SNR at the  $i$ th finger is

$$\gamma_i = \frac{|\alpha_i|^2 P_{\text{tx}}}{G_d \sigma^2}, \quad (5.20)$$

where  $P_{\text{tx}}$  is the transmit power,  $G_d$  denotes the path loss at distance  $d$ , and  $\sigma^2$  represents the noise power at the receiver. Also,  $\alpha_i$  represents the attenuation of the selected path preserving the  $i$ th largest power. The instantaneous SNR at the output of the MRC is  $\gamma = \sum_{i=1}^M \gamma_i$ .

The average bit error probability of DBPSK-CO can be found by averaging the BEP of DBPSK-CO over an AWGN channel, with the pdf of  $\gamma$  which is indicated by  $p(\gamma)$ . By using an alternate representation of the Q-function, we have the BEP of DBPSK-CO over an AWGN channel as [7][99]

$$\begin{aligned} P_{b,\text{DBPSK-CO,AWGN}} &\approx 2Q(\sqrt{2\gamma})[1 - Q(\sqrt{2\gamma})]. \\ &\approx 2Q(\sqrt{2\gamma}), \\ &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{2\gamma}{\sin^2 \phi}} d\phi. \end{aligned} \quad (5.21)$$

The average BEP can be expressed as

$$\begin{aligned}
P_{b,\text{DBPSK-CO}} &= \int_0^\infty P_{b,\text{DBPSK-CO,AWGN}} p(\gamma) d\gamma, \\
&= \int_0^\infty \left[ \frac{2}{\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{2\gamma}{\sin^2\phi}} d\phi \right] p(\gamma) d\gamma, \\
&= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \left[ \int_0^\infty e^{-\frac{2\gamma}{\sin^2\phi}} p(\gamma) d\gamma \right] d\phi,
\end{aligned} \tag{5.22}$$

On the other hand, the moment generating function (MGF) of the pdf of  $\gamma$  is defined as  $\Psi_\gamma(\nu) = \int_0^\infty e^{\nu\gamma} p(\gamma) d\gamma$ . Therefore,

$$P_{b,\text{DBPSK-CO}} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \Psi_\gamma(\nu) \Big|_{\nu=-\frac{2}{\sin^2\phi}} d\phi. \tag{5.23}$$

Since  $\alpha_{k,l}$  follows a Nakagami- $m$  distribution from the channel model, the MGF of the pdf of  $\gamma$  has been established as [87]

$$\Psi_\gamma(\nu) = \prod_{i=1}^M \frac{1}{(1 - \nu\bar{\gamma}_i/m_i)^{m_i}} \tag{5.24}$$

where  $\bar{\gamma}_i = E[|\alpha_i|^2] P_{\text{tx}} G_c / G_d \sigma^2$  is the average received SNR at the  $i$ th finger.  $E[|\alpha_i|^2]$  represents the average power of the  $i$ th selected path, that is  $E[|\alpha_i|^2] = E[|\alpha_{k,l}|^2]$ , where  $|\alpha_{k,l}|^2$  has the  $i$ th largest expectation among all paths.  $m_i$  is the Nakagami  $m$ -factor for the  $i$ th selected path. As shown in [46], for the first ray of each cluster,  $m_i$  is assumed to be deterministic and independent of delay; while for the remaining paths,  $m_i$  follows a delay-dependence lognormal distribution.

Moreover, if we consider repetition coding with parameter  $N_p$  and SD combining, the average SNR will increase  $N_p$  times compared with the corresponding uncoded modulations. That is, the eventual average BEP for DBPSK with coherent detection and SD combining (DBPSK-CO-SD) is

$$P_{b,\text{DBPSK-CO-SD}} = \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \Psi_\gamma(\nu) \Big|_{\nu=-\frac{2N_p}{\sin^2\phi}} d\phi. \tag{5.25}$$

Instead, if HD combining is used, the average BEP for DBPSK with coherent detection and HD combining (DBPSK-CO-HD)  $P_{b,\text{DBPSK-CO-HD}}$  can be expressed as

$$P_{b,\text{DBPSK-CO-HD}} = \sum_{k=\frac{N_p+1}{2}}^{N_p} \binom{N_p}{k} P_{b,\text{DBPSK-CO}}^k (1 - P_{b,\text{DBPSK-CO}})^{N_p-k}, \quad (5.26)$$

where  $P_{b,\text{DBPSK-CO}}$  is given in (5.22). A similar procedure can be directly applied to OOK with coherent detection and DPPM with coherent detection, using either HD or SD combining.

In the case of DBPSK-NC, the average bit error probability of DBPSK-CO can be found by directly utilizing the MGF of the pdf of  $\gamma$ . First, we have the BEP of DBPSK-NC over an AWGN channel as [7]

$$P_{b,\text{DBPSK-NC,AWGN}} \approx \frac{1}{2}e^{-\gamma} \quad (5.27)$$

The average BEP can then be expressed as

$$\begin{aligned} P_{b,\text{DBPSK-NC}} &= \int_0^\infty P_{b,\text{DBPSK-NC,AWGN}} p(\gamma) d\gamma, \\ &= \int_0^\infty \frac{1}{2}e^{-\gamma} p(\gamma) d\gamma, \\ &= \frac{1}{2} \Psi_\gamma(\nu)|_{\nu=-1}, \\ &= \frac{1}{2} \prod_{i=1}^M \frac{1}{(1+\bar{\gamma}_i/m_i)^{m_i}} \end{aligned} \quad (5.28)$$

Correspondingly, we have

$$\begin{aligned} P_{b,\text{DBPSK-NC-SD}} &= \frac{1}{2} \prod_{i=1}^M \frac{1}{(1 + N_p \bar{\gamma}_i / m_i)^{m_i}}, \\ P_{b,\text{DBPSK-NC-HD}} &= \sum_{k=\frac{N_p+1}{2}}^{N_p} \binom{N_p}{k} P_{b,\text{DBPSK-NC}}^k (1 - P_{b,\text{DBPSK-NC}})^{N_p-k}. \end{aligned} \quad (5.29) \quad (5.30)$$

The average bit error probabilities of DPPM and OOK with non-coherent detection, using either HD or SD combining, can be obtained following a similar procedure.

Moreover, to avoid excessive requirements of memory units, only HD combining is considered for noncoherent detections.

### 5.3.2.3 Energy per Information Bit Minimization

From (5.19), the energy consumption per information bit is

$$E_b = \frac{E^{(\text{TX})} + E^{(\text{RX})} + E_{\text{LN}} + E_{\text{IPS}}}{L_L(1 - P_b)^{L_L}} + \frac{E_{\text{TRAN}} + E_{\text{ACK}}^{(\text{TX})} + E_{\text{ACK}}^{(\text{RX})} - E_{\text{LN}}}{L_L}. \quad (5.31)$$

Our goal is to find an optimum combination of the modulation scheme, the detection scheme, the repetition coding parameter  $N_p$ , the combining scheme, the packet length,  $L_L$ , and the number of RAKE fingers at the receiver,  $M$ , over a slow frequency-selective channel for a given transmission distance, that minimizes the effective energy consumption per information bit denoted by (5.31). Removing the integer constraint on  $L_L$ , it is straight forward to find the closed form optimum packet length by solving  $\partial E_b / \partial L_L = 0$ . At high SNR, the result is

$$L_L^* \approx \frac{-P_b(A+B) + \sqrt{P_b^2(A+B)^2 + 4(A+B)CP_b}}{2CP_b}, \quad (5.32)$$

where

$$\begin{aligned} A &= E_{\text{TRAN}} + E_{\text{ACK}}^{(\text{TX})} + E_{\text{ACK}}^{(\text{RX})} - E_{\text{LN}}, \\ B &= E_{\text{IPS}} + E_{\text{LN}} + E_{\text{O}}^{(\text{TX})} + E_{\text{O}}^{(\text{RX})} + E_L^{(\text{RX})}, \\ C &= (\rho_t E_p + P_{\text{SYN}}/R_b)/R_c \\ &\quad + [\rho_t(MP_{\text{COR}} + \rho_c P_{\text{ADC}} + P_{\text{LNA}} + P_{\text{VGA}}) + \rho_r(P_{\text{GEN}} + P_{\text{SYN}})]/(R_b R_c). \end{aligned}$$

As indicated by (5.32),  $L_L^*$  will decrease as BEP increases. In a real application of this model, the packet length can always be selected as the nearest integer of the resulting  $L_L^*$ .

In this work, we have assumed that the data rate is fixed at the maximum allowable data rate that avoids ISI. Correspondingly, the transmit power, as shown in (5.1), is also fixed. Therefore, the average BEP at a given transmission distance for a given

modulation scheme is only determined by the modulation and detection schemes, repetition coding and combining schemes, and the number of RAKE fingers at the receiver. For a given combination of the modulation/detection scheme and the repetition coding/combining scheme, the BEP follows a non-increasing function of the number of RAKE fingers. Thus,  $L_L^*$  follows a non-decreasing function of the number of RAKE fingers.

The optimum number of RAKE fingers reflects the tradeoff between the power consumption cost,  $MP_{\text{COR}}$  and the received power gain,  $E[|\alpha_i|^2]P_{\text{tx}}/G_d$ . The optimum selection of modulation/detection schemes and repetition coding/combining schemes reflects the tradeoff between the performance and the power consumption at the transceiver, since higher performance is often accompanied by higher power consumption. Unlike the optimum packet length, there are no closed form expressions for the optimum modulation/detection schemes, repetition coding/combining schemes and number of RAKE fingers. However, numerical optimizations can be performed over these parameters, and the optimization results will be given in the following section.

## 5.4 Numerical Results

In this section, we demonstrate the results of minimizing the effective energy consumption per information bit modeled by equation (5.31). We assume that  $B = 500\text{MHz}$ ,  $L_{\text{SP}} = 1024$  symbols,  $L_{\text{PHR}} = 16$  symbols [32], coding rate  $R_c = 1/N_p$ , and  $N_p = 1, 3, 5, \dots, 15$ . The maximum excess delay is  $D_S = 40$  ns, which limits the maximum pulse rate to  $R_p \approx 1/D_s = 25$  Mbps to avoid inter symbol interference. The power consumptions of the transmitter and receiver components are as follows [95],[100]-[103]:  $P_{\text{SYN}} = 30.6\text{mW}$ ,  $P_{\text{ADC}} = 2.2\text{mW}$ ,  $P_{\text{GEN}} = 2.8\text{mW}$ ,  $P_{\text{VGA}} = 22\text{mW}$ ,  $P_{\text{LNA}} = 9.4\text{mW}$ ,  $P_{\text{COR}} = 10.08\text{mW}$ . To the best of our knowledge, there is no existing literature providing specific power consumption evaluations of the channel estimation block in an IR-UWB receiver. Therefore, in this work we assume the IR-UWB receiver uses the

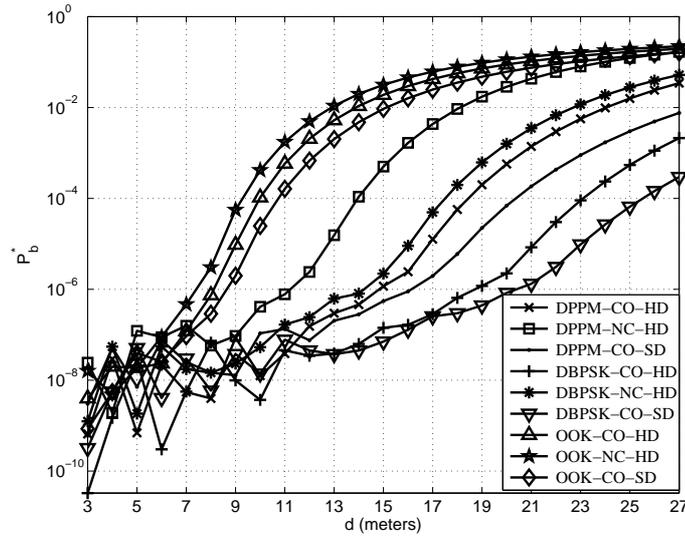


Figure 5.4: The optimum target BEP versus distance ( $N_p = 3$ ).

data-aided estimation method from [104], which can essentially be implemented as a correlator. Thus, we assume  $P_{\text{EST}} = P_{\text{COR}} = 10.08\text{mW}$ .

The fixed emitted energy per pulse is  $E_p = 4.5$  pJ/pulse. Therefore, the maximum amount of transmit power is  $P_{\text{tx}} = E_p R_p = 0.113$  mW. Also, we assume  $T_{\text{IPS}} = 200$   $\mu\text{s}$  and  $T_{\text{tr}} = 400$   $\mu\text{s}$ . Moreover,  $R_{\text{base}} = 1$  Mbps, where  $R_{\text{base}}$  denotes the data rate to transmit the header and preamble symbols, and the path loss parameter is set to  $L_w = 0.7$  dB/m.

The parameters of the channel model for an office environment with no line of sight (NLOS) are used [46]. That is, the path loss exponent  $n = 3.07$ , the frequency dependency decaying factor  $\kappa = 0.71$ , reference path gain  $G_0 = -59.9$  dB, transmission center frequency  $f = 3.1$  GHz, the reference frequency  $f_c = 5$  GHz. The distance range of interest is  $d \in [3, 27]$  meters. We used exhaustive search to solve the optimization model. The quality of service (QoS) is assumed to be error free. That is, the expected number of total transmissions is  $1/(1 - P_b)^{L_L}$ .

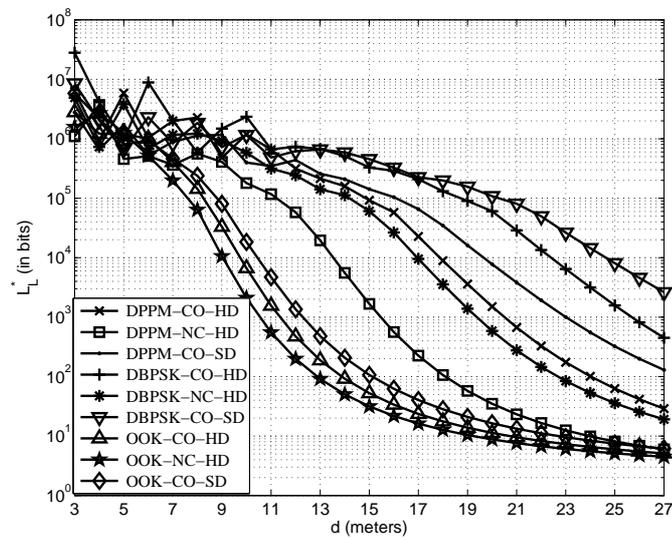


Figure 5.5: The optimum packet length versus distance ( $N_p = 3$ ).

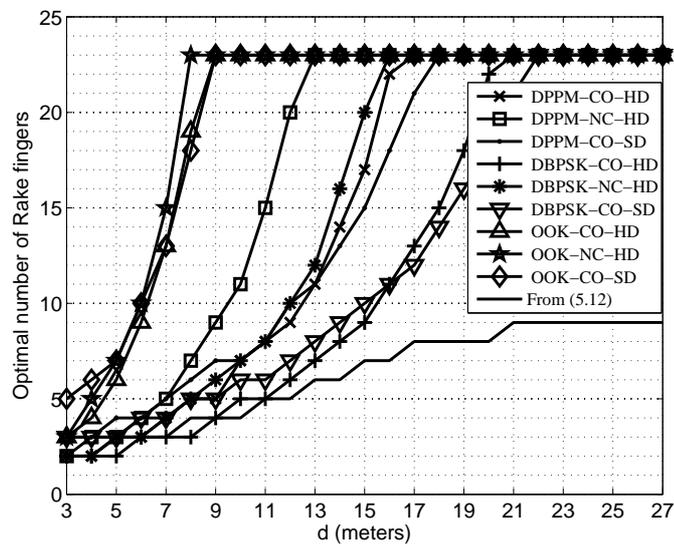


Figure 5.6: The optimum number of RAKE fingers versus distance ( $N_p = 3$ ).

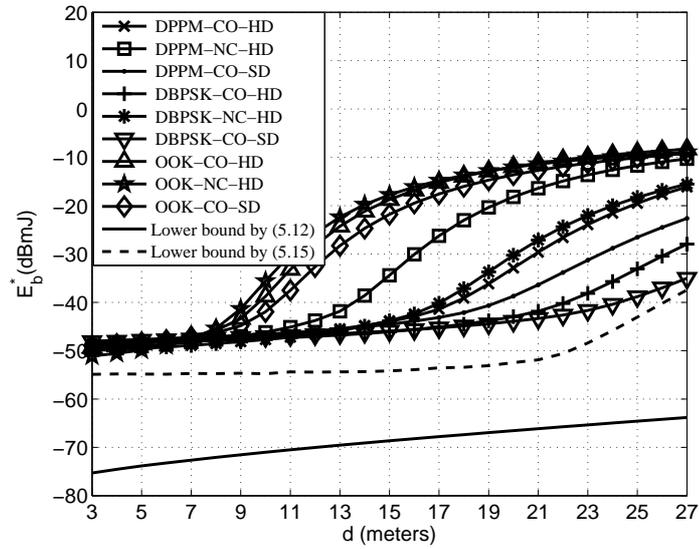


Figure 5.7: The minimum energy consumption per information bit versus distance ( $N_p = 3$ ).

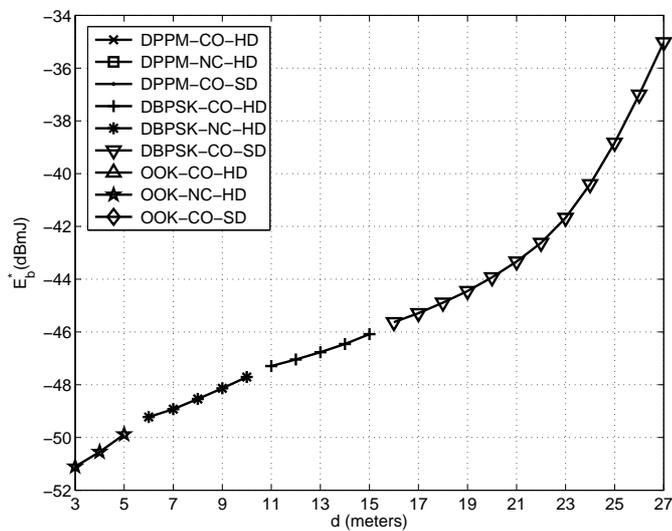


Figure 5.8: The overall minimum energy consumption per information bit and corresponding modulation/decoding/detection schemes ( $N_p = 3$ ).

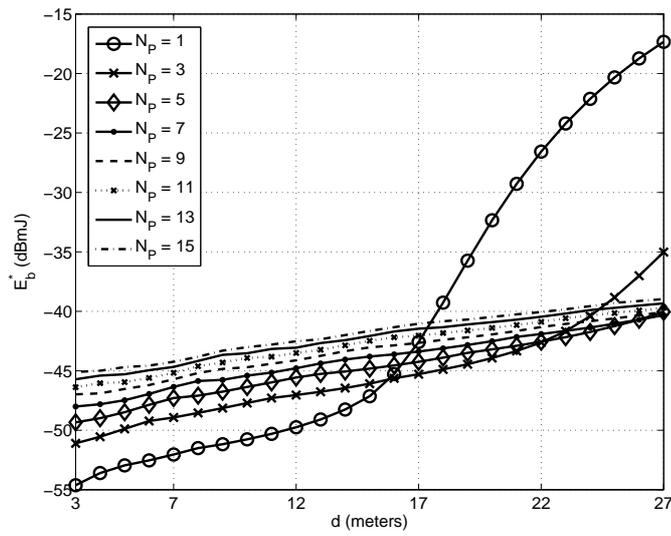


Figure 5.9: The overall minimum energy consumption per information bit for different repetition coding parameters with optimized  $L_L$ ,  $M$ , and modulation/repetition coding/detection schemes.

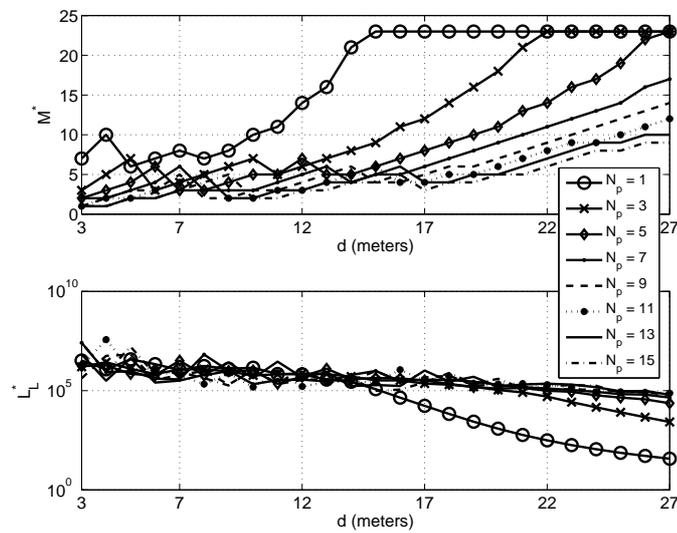


Figure 5.10: The  $M^*$  and  $L_L^*$  with optimized modulation/combining/detection schemes for different  $N_p$ .

### 5.4.1 Optimization With Fixed $N_p$

To better understand the influence of the packet length, the number of RAKE fingers, the modulation, the combining schemes, and the detection schemes, first, we consider a fixed  $N_p = 3$  to isolate the impact of repetition coding from the rest of the parameters.

The optimum BEPs and optimum packet lengths of different modulation/combining/detection schemes with  $N_p = 3$  are shown in Figs. 5.4 and 5.5, respectively. As shown in Fig. 5.4, as the transmission distance increases, the optimum BEP will increase since it will require increasingly more power at the receiver to maintain a low BEP as  $d$  increases. Therefore, the optimum choice is to lower the target BEPs to avoid a dramatic increase in the number of RAKE fingers. Correspondingly, as shown in Fig. 5.5, the optimum packet length will decrease as  $d$  increases to avoid costly retransmissions caused by higher BEP, since a short packet length results in a lower packet error probability.

Note that, at short transmission distances, there are high variations of  $P_b^*$  and  $L_L^*$ . This is because at short distances where  $P_b^*$  is very low,  $P_b^*$  is very sensitive to a change in the number of RAKE fingers. That is, at short distances, additional RAKE fingers will provide a considerable amount of collected power gain. Correspondingly,  $L_L^*$  will change significantly as the number of RAKE fingers changes at short distances. On the other hand, at large distances, additional RAKE fingers only provide a small amount of collected power gain. Therefore,  $P_b^*$  and  $L_L^*$  are not sensitive to a change in the number of RAKE fingers. The curves of  $P_b^*$  and  $L_L^*$  thereby become smooth as distances increase.

Fig. 5.6 shows the optimum number of RAKE fingers at the receiver versus distance. As the transmission distance increases, the optimum number of RAKE fingers will increase and converge to a certain value. This is due to the change of balance between the diversity gain and the power cost induced by each additional RAKE finger. The absolute value of the increase in collected energy by an additional RAKE finger decreases with distance. This diminishing gain incurs a fixed cost, namely  $P_{\text{COR}}$ .

Consequently, as distance increases, to avoid excessive retransmissions, the number of RAKE fingers should increase to collect more received power. However, when distance increases beyond a certain level, path loss becomes very large, and increasing the number of RAKE fingers does not lead to the collection of considerably more received power. However, more RAKE fingers will incur more power consumption at the receiver. Thus, at large transmission distances, increasing the number of RAKE fingers does not improve the energy efficiency. The optimum receiver power consumptions of different modulation schemes at different transmission distances follow the trend of the optimum number of RAKE fingers shown in Fig. 5.6. The optimum number of RAKE fingers from (5.12) is also shown in Fig. 5.6. Since (5.12) does not include the imperfection of repetition coding/modulation and the overhead caused by packetization and retransmissions, the received power that is collected by the RAKE fingers at the receiver reaches the theoretical maximum efficiency (no waste on overhead).

The overall minimized energy consumption per information bit is shown in Fig. 5.7. OOK-NC-HD consumes the least amount of energy when  $d < 6\text{m}$ , while DBPSK-NC-HD offers the lowest energy consumption per information bit when  $6\text{m} \leq d < 11\text{m}$ . DBPSK-CO-HD is the most energy efficient scheme when  $11\text{m} \leq d < 16\text{m}$ . DBPSK-CO-SD is the most energy efficient scheme when the distance is greater than 16 meters. This trend reflects the balance between the transmitter energy consumption and the receiver energy consumption. At a short transmission distance, the less robust schemes (OOK-NC-HD) require less power consumption at the transmitter/receiver and provide a BEP low enough to avoid excessive retransmissions. Therefore, the OOK scheme, noncoherent detection, and HD combining have a high energy efficiency at short transmission distances. However, as transmission distance increases, the above schemes require a large number of RAKE fingers to maintain a low BEP, thereby the receiver power consumption will increase dramatically if using schemes like OOK-NC-HD. On the other hand, the more robust schemes (such as DBPSK-CO-HD, DBPSK-CO-SD) consume much less power at the receiver since they need fewer RAKE fingers

to achieve a low BEP. Thus, as transmission distance increases, these schemes will become the energy efficient schemes. The lower bound on  $E_b$  from (5.12) and the packetized lower bound on  $E_b$  from (5.15) are also shown in Fig. 5.7. The packetized bound from (5.15) is larger than (5.12), especially for large distances. This is caused by the decrease of  $L_L^*$  as distance increases, as shown in Fig. 5.5, which in turn increases the packetization overhead.

The overall minimum  $E_b$  and corresponding modulation, repetition coding and detection schemes are shown in Fig. 5.8. This figure, together with Fig. 5.6, reveals that, at short distances (high SNRs), low complexity and low performance modulation/repetition coding/detection schemes, such as OOK-NC-HD with a small number of RAKE fingers, are energy efficient; while at large distances (low SNRs), higher complexity and higher performance modulation/repetition coding/detection schemes, such as DBPSK-CO-SD with a large number of RAKE fingers, become energy efficient.

#### 5.4.2 Optimization With Variable $N_p$

Besides the optimization on packet length, number of RAKE fingers and modulation/combining/ detection schemes, the repetition coding parameter  $N_p$  in repetition coding should also be adjusted to minimize  $E_b$ . Now we assume that  $N_p$  takes the values 1, 3, 5, ..., 15.

Fig. 5.9 shows the overall minimum energy consumption per information bit for different repetition coding parameters. To show the influence of  $N_p$ , the  $E_b^*$  shown in Fig. 5.9 have been optimized over  $L_L$ ,  $M$  and modulation/repetition coding/detection schemes. Fig. 5.9 shows that the optimum  $N_p$  increases as  $d$  increases (SNR decreases). This observation further confirms that, to guarantee link level reliable communication, high-complexity and high-performance transceiver structures are energy efficient at low SNRs; while low-complexity and low-performance transceiver structures are energy efficient at high SNRs.

Fig. 5.10 shows the optimum number of RAKE fingers and packet lengths for different repetition coding parameters  $N_p$ , with optimized modulation/combining/detection schemes. The curves with different  $N_p$  display the same trend: as  $d$  increases (SNR decreases),  $M^*$  increases and converges to the same level, while  $L_L^*$  decreases. Fig. 5.10 also reveals that the effect of a large  $N_p$  on the packet error probability on the expected number of retransmissions is equivalent to that of a small  $L_L$  or a large  $M$ , and vice versa. Therefore, it is possible to use a long repetition code to minimize energy under some circumstances where a large number of RAKE fingers and adjustable packet lengths are not feasible.

### 5.4.3 Optimum Configuration Table

The results of these optimizations can enable the transceiver to adapt by selecting the overall optimum configurations (including the modulation/detection scheme, the repetition coding/combining scheme, the packet length and the number of RAKE fingers) according to the expected transmission distance through a lookup table. Table 8.1 is an example of such a look-up table obtained with the particular power consumption and channel models assumed in this work.

### 5.4.4 The Effects of Power Consumption Values on the Optimum Configurations

Implementation technologies have a paramount impact on the choice of the optimal communication schemes. For example, should  $P_{\text{SYN}}$  become smaller, the distance range in which coherent detection is energy efficient becomes larger. In fact, considering an extreme case where the transceiver does not consume any power, we shall always use the communication scheme with the highest performance, such as DBPSK with coherent detection, soft decoding and an all-RAKE receiver. In real applications, the

overall optimum configurations should be carefully evaluated using the generic energy consumption model provided in this work as summarized in (5.31) and the power consumption values of the actual circuit components based on the adopted production technologies. For instance, suppose we choose another set of power consumption values where  $P_{\text{SYN}} = 2.2\text{mW}$ , with the rest of the power consumption values stay the same. The resulting overall optimum configurations are summarized in Table 5.2. By comparing Tables 8.1 and 5.2, we find that since the difference in power consumption of coherent and noncoherent schemes becomes smaller, the distance range in which coherent detection is energy efficient becomes larger. However, the general trend of the optimal configurations versus transmission distance stays the same: high-complexity and high-performance transceiver structures are energy efficient at large distances; while low-complexity and low-performance transceiver structures are energy efficient at short distances.

## 5.5 Summary

In this chapter, we provided the power consumption models of typical transmitter and receiver structures of IR-UWB systems. Then, under the assumption of a frequency selective time-invariant channel, a minimization of energy consumption per information bit considering packet overhead, retransmission, repetition coding and number of RAKE fingers is performed. An optimum number of RAKE fingers exists under the assumption of a frequency selective time-invariant channel with a double exponentially decaying power delay profile. The results show that low-complexity, low-performance transmission schemes are energy efficient at high SNRs, while high-complexity, high-performance schemes are energy efficient at low SNRs. Detailed optimum transmission schemes, including packet length, modulations, detection, repetition coding, combining, and number of RAKE fingers, are also provided for given transmission distances for a typical IR-UWB link.

Table 5.1: Overall Optimum Configurations

Distance (m)	Modulation	Detection	$N_p$	Combining	$M^*$	$L_L^*$ (Kbit)				
3	OOK	NC	1	HD	7	~ 1500				
4					10					
5	6									
6	7									
7	8									
8	7									
9	8									
10	10									
11	11				CO		3	SD	11	~ 500
12	14									
13	16									
14	21									
15	23									
16	12									
17	14									
18	16									
19	18									
20	21									
21	23	5	7	SD	16	~ 50				
22	17									
23	19									
24	16									
25	17									
26	17									
27	19									

Table 5.2: Overall Optimum Configurations

Distance (m)	Modulation	Detection	$N_p$	Combining	$M^*$	$L_L^*$ (Kbit)					
3	OOK	NC	1	HD	7	$\sim 3000$					
4					9	$\sim 1500$					
5	DBPSK				4						
6					5						
7					6						
8					7						
9					8						
10					9		$\sim 500$				
11					11						
12					13						
13					DBPSK			3	SD	15	$\sim 100$
14										18	
15										22	
16										10	
17										11	
18			12								
19	14										
20	16										
21	18	$\sim 50$									
22	13										
23	14										
24	16										
25	18	5									
26	14										
27	16										
			7								

## 6 Maximizing Data Gathering in Clustered Short-Range Wireless Networks

In this chapter, we extend our cross-layer optimization to include the network layer. In particular, we investigate the maximization of the amount of gathered data in a clustered short-range wireless network (SRWN). The amount of gathered data is maximized by (i) choosing the optimal transmit power, (ii) selecting the optimal cluster head, and (iii) deciding whether or not to use multi-hop within a cluster. For problem (i), we find closed-form solutions for the optimal or near optimal transmit power of cluster members (CM). For problem (ii), we propose a near optimal cluster head selection (CHS) algorithm. The communication burden and computational complexity of CHS only grow linearly with the size of the cluster. Finally, for problem (iii), we propose an algorithm to decide whether or not to use multi-hop within a cluster to further increase the amount of data gathered by a cluster.

In the proposed algorithms, iterations have been avoided in order to significantly lower the complexity of the algorithms compared with traditional iteration-based numerical optimization algorithms, making these approaches suitable for use in energy-constrained wireless networks. The optimization gain is shown to be significant.

## 6.1 Introduction

In the domain of short-range wireless networks, there has been much research effort aimed at efficiently utilizing the limited energy at the nodes [2][10]. Numerous strategies have been investigated to improve the energy efficiency of SRWNs, including power control, mobile data sink deployment, multiple data sink deployment, nonuniform initial energy assignment, and intelligent node and relay deployment [11].

Another option to improve the energy efficiency of SRWNs, clustering protocols have been broadly adopted due to their effectiveness and simplicity. In clustered networks, neighboring nodes are grouped as clusters. One of the nodes in a cluster is selected as the cluster head, and the remaining nodes are the cluster members. The cluster head is usually in charge of certain local coordinations, such as collecting data from the cluster members and communicating with other clusters and the data sink, while cluster members simply transmit data to the cluster head. The cluster head may be selected in a randomized manner, such as in HEED [57] or LEACH [58]. Such a randomized selection of the cluster head, combined with rotating the cluster head position, can effectively avoid the early drain of the energy of a particular node. However, it cannot guarantee the optimality of the selection.

Transmit power is a very important factor that influences the energy efficiency of SRWNs from the physical layer (PHY) [12]. Power control techniques can be easily adopted in clustering topologies [56]. In this chapter, we focus our attention on a clustered wireless network with power control capability at the cluster members. The goal is to maximize the energy efficiency of a cluster by optimally determining the cluster head and obtaining the optimal transmit power of the cluster members.

Multi-hop can be used to further improve the energy efficiency of a SRWN. By utilizing multi-hop routing, the transmitting nodes are able to use less transmit power and thereby save energy spent on transmitting information. However, in multi-hop routing, there are more nodes involved in delivering information from the source node

to the sink. Therefore, the energy efficiency needs to be carefully evaluated considering the extra energy consumption at the relay nodes. In this chapter, we analyze the use of multi-hop within the scope of a cluster to further improve the energy efficiency of a cluster. That is, we propose a criterion to determine if a transmitting cluster member should choose multi-hop routing to reach the cluster head, and which of its neighbor should be chosen as the relay node, so that the energy efficiency of the cluster can be improved.

The energy efficiency of wireless networks is usually evaluated by network lifetime. Network lifetime can be defined as the time elapsed until the first node in the network depletes its energy. This definition of lifetime can avoid the situation that certain nodes have very high power consumption, such as the common relay nodes of many routes, during network optimization. However, in a clustering-based network, due to the uneven importance of cluster members and cluster heads, the definition of network lifetime needs to be modified.

In this chapter, we define the lifetime of a cluster as the time duration for a cluster to function properly, i.e., the time elapsed until the cluster head dies or all cluster members die, whichever comes first. In data-centric applications, however, lifetime itself is not as meaningful a criterion as the amount of data gathered during the lifetime to evaluate the performance of a cluster. Thus, the goal of our proposed optimization is to maximize the amount of data gathered during the lifetime of a cluster, under the energy constraints and fairness constraints. The optimization parameters are the transmit power of cluster members, the selection of the cluster head, and the decision about whether or not to use multi-hop routing within the cluster.

The rest of the chapter is organized as follows: In section 6.2, we formulate the optimization problem for selecting the cluster members' transmit power to maximize the amount of data gathered during the lifetime of a cluster, given energy and fairness constraints. Moreover, we derive an iteration-free solution to this optimization problem. The complexity of the solution is significantly lower than conventional iterative

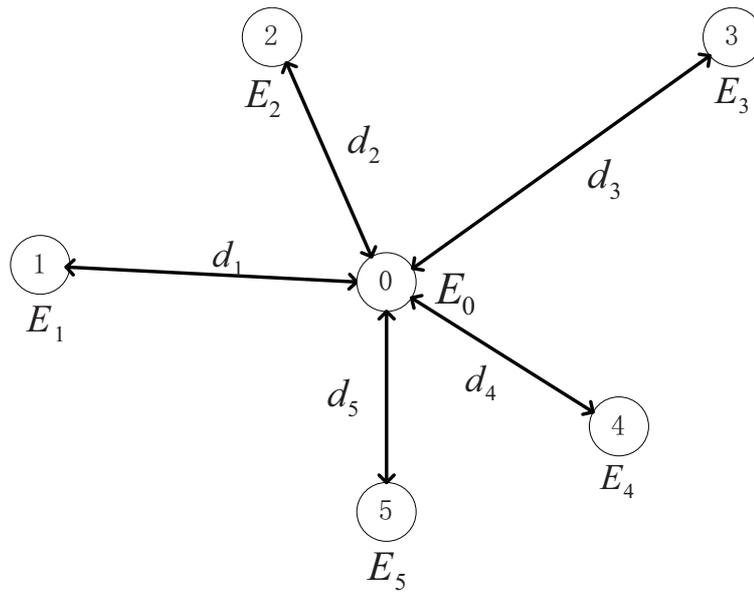


Figure 6.1: A typical cluster topology ( $N_d = 5$ ).

numerical solutions. In section 6.3, we model the cluster head selection problem, and from this model we formulate an optimization problem to determine the near optimal cluster head. In section 6.4, an algorithm is proposed to decide whether or not to use a multi-hop within a cluster to further increase the amount of data gathered during the lifetime of the cluster. The performance of the proposed algorithms is evaluated in section 6.5. Finally, section 6.6 concludes this chapter.

## 6.2 Transmit Power Optimization

### 6.2.1 Optimization Model

Fig. 6.1 shows a typical cluster topology with 5 cluster members (CMs) and 1 cluster head (CH). In this work, cluster members operate under a *fairness constraint*. The fairness constraint of a CM is defined as the constraint that the cluster head should collect an equal number of information bits from each CM. That is, if every CM  $i$  works

for time duration  $T_i$ , then during  $T_i$ , each CM transmits  $D$  bits of data to the cluster head. This fairness constraint is due to the fact that the nodes in a cluster are usually geographically adjacent, and thus they observe the same phenomenon and generate the same amount of samples and, consequently, generate the same amount of data for transmission given every node has the same source coding rate.

Given this fairness constraint and Shannon's channel capacity theorem, we have

$$BT_i \log \left( 1 + \frac{P_i}{d_i^n P_{n,i}} \right) > D, \quad \forall i \in \{1, 2, \dots, N_d\}, \quad (6.1)$$

where  $B$  is the bandwidth,  $P_i$  is the transmit power of cluster member  $i$ ,  $d_i$  is the transmission distance from cluster member  $i$  to the cluster head,  $n$  is the path loss exponent,  $P_{n,i}$  is the additive white gaussian noise (AWGN) power on the link from node  $i$  to the cluster head, and  $N_d$  denotes the number of cluster members.

Rearranging Eq. (6.1), the transmit power of CM  $i$  is

$$P_i > \left( 2^{\frac{D}{BT_i}} - 1 \right) d_i^n P_{n,i}. \quad (6.2)$$

That is, the transmit power of CM  $i$  is determined by its operation time duration  $T_i$  and the total transmitted data  $D$ .

Moreover, the energy constraints at the CMs are

$$T_i(P_i + P_{CT}) \leq E_i, \quad \forall i \in \{1, 2, \dots, N_d\}, \quad (6.3)$$

where  $P_{CT}$  denotes the circuit power consumption at the CMs, and the transmit powers are constrained by Eq. (6.2).

At the cluster head, the energy consumption comes from receiving the data from cluster members. Therefore, the energy constraint at the cluster head is

$$P_{CR} \sum_{i=1}^{N_d} T_i \leq E_0, \quad (6.4)$$

where  $P_{CR}$  is the circuit power consumption to receive data.

## 6.2.2 Problem Formulation

The resulting maximization of the total gathered data given the energy constraints and the fairness constraint can be formulated as

$$\begin{aligned}
& \min \quad -D \\
& \text{s.t.} \quad C_1 : T_i > 0, \\
& \quad \quad C_2 : P_{CR} \sum_{i=1}^{N_d} T_i \leq E_0 \\
& \quad \quad C_3 : T_i [P_i + P_{CT}] \leq E_i, \\
& \quad \quad C_4 : P_i > (2^{\frac{D}{BT_i}} - 1) d_i^n P_{n,i},
\end{aligned} \tag{6.5}$$

where  $i \in \{1, 2, \dots, N_d\}$ . In this model, constraints  $C_2$  and  $C_3$  are the energy constraints at the cluster head and cluster members, respectively.  $C_4$  results from the fairness constraint. Our goal is, for given residual energies and communication environments, to find the optimal cluster member operation time durations  $[T_1, T_2, \dots, T_{N_d}]$  and transmit powers  $P_i$  that maximize the total data gathered at the cluster head.

Problem (6.5) is a minimum feasible set test problem with variables  $T_i$ ,  $P_i$  and  $D$  [105]. Assume that the optimal solution is  $D^*$ ,  $T_i^*$  and  $P_i^*$ , where  $P_i^* = (2^{\frac{D^*}{BT_i^*}} - 1) d_i^n P_{n,i} + \epsilon_i$ ,  $\epsilon_i > 0$ . It is obvious that  $D^*$ ,  $T_i^*$  and  $\hat{P}_i$ , where  $\hat{P}_i = (2^{\frac{D^*}{BT_i^*}} - 1) d_i^n P_{n,i} + \frac{\epsilon_i}{2}$ ,  $\epsilon_i > 0$ , is also an optimal solution, as  $D^*$ ,  $T_i^*$  and  $\hat{P}_i$  satisfy the constraints. Therefore, we can always have the optimal transmit power  $P_i^* \rightarrow (2^{\frac{D^*}{BT_i^*}} - 1) d_i^n P_{n,i}$  from the right hand side. Thus, we can replace the inequality in constraint  $C_4$  with equality. The resulting problem model is

$$\begin{aligned}
& \min \quad -D \\
& \text{s.t.} \quad C_1 : T_i > 0, \\
& \quad \quad C_2 : P_{CR} \sum_{i=1}^{N_d} T_i \leq E_0 \\
& \quad \quad C_3 : T_i [P_i + P_{CT}] \leq E_i, \\
& \quad \quad C_4 : P_i = (2^{\frac{D}{BT_i}} - 1) d_i^n P_{n,i}.
\end{aligned} \tag{6.6}$$

The optimal solutions of the above problem are arbitrarily close to the optimal solutions

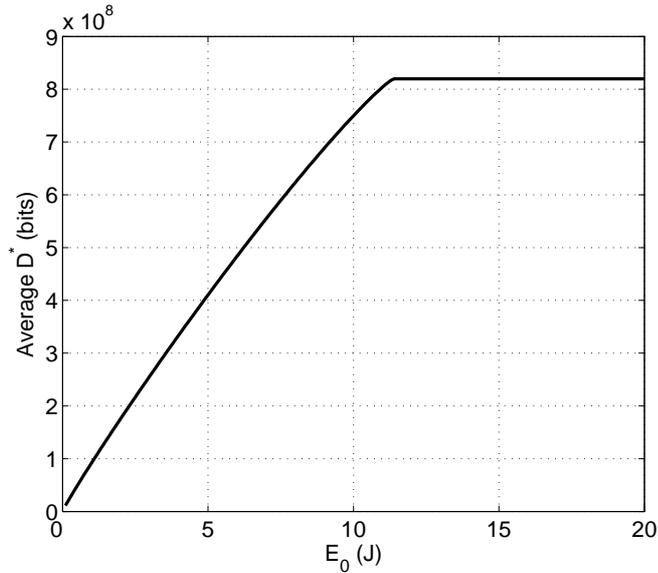


Figure 6.2: Average  $D^*$  as the CH energy increases. of problem (6.5). We determine an iteration-free solution to this problem in the next subsection.

### 6.2.3 Transmit Power Optimization Algorithm

Problem (6.6) is a typical minimum feasible set test of a quasi-convex optimization problem, which can be readily solved through iterative numerical methods [105]. However, performing hundreds of iterations is prohibitively complex in energy-limited SR-WNs, and a simple solution with few or no iterations is desirable. In this section, we propose a simple iteration free solution to find the near optimal transmission time durations and transmit powers of each CM in a cluster.

Consider a cluster with 5 cluster members where the initial energy of the cluster members and the transmission distances are randomly generated. Fig. 6.2 presents the average maximum information bits gathered by the cluster head,  $D^*$ , obtained from the numerical solution of (6.6). We observe that as  $E_0$  increases, the maximum data gathered at the CH increases until it reaches a saturation limit. That is, there are two regions of maximum data gathered,  $D^*$ , as  $E_0$  increases, namely the *CH-constrained region*

and the *CM-constrained region*. In the CH-constrained region, the cluster head energy is the limiting factor in the total amount of data collected, while in the CM-constrained region, the residual energy at the CMs becomes the limiting factor. Our goal is to find an iteration-free solution for  $T_i$  that maximizes  $D$  in both the CH-constrained and the CM-constrained regions. After obtaining the optimal  $T_i$ , the corresponding optimal transmit power  $P_i$  of the CMs can be easily determined.

### 6.2.3.1 CH-constrained region

In the CH-constrained region, the limiting factor on the amount of data collected is the cluster head energy. In this case, the cluster head uses up its energy. The total available operation time for all cluster members can thus be expressed as

$$\sum_{i=1}^{N_d} T_i = \frac{E_0}{P_{CR}}. \quad (6.7)$$

Through linear approximations, an IFS for problem (6.6) when the cluster operates in the CH-constrained region can be achieved.

In the CH-constrained region, since the cluster head is the limiting factor, the cluster head and all cluster members should use up their energy. Otherwise, any cluster member having energy left indicates that at least one of the cluster members limits the increase of the collected data and the cluster is operating in the CM-constrained region. Therefore, both the cluster head and the cluster member energy constraints are active. That is

$$\sum_{i=1}^{N_d} T_i = \frac{E_0}{P_{CR}}, \quad (6.8)$$

$$T_i \left[ \left( 2^{\frac{D}{BT_i}} - 1 \right) d_i^n P_{n,i} + P_{CT} \right] = E_i. \quad (6.9)$$

From (6.9), we have

$$\begin{aligned}
\log_2 \left( \frac{E_i - P_{CT}}{d_i^n P_{n,i}} + 1 \right) &= \frac{D}{BT_i}, \\
\stackrel{a}{\implies} T_i \log_2 \left( \frac{E_i}{d_i^n P_{n,i}} \right) &\approx \frac{D}{B}, \\
\implies T_i \log_2 \left( \frac{E_i}{d_i^n P_{n,i}} \right) - T_i \log_2 T_i &\approx \frac{D}{B}, \\
\stackrel{b}{\implies} T_i &\approx \frac{D/B}{\log_2 \left( \frac{E_i}{d_i^n P_{n,i}} \right) + \frac{6}{\ln 2}},
\end{aligned} \tag{6.10}$$

where step *a* follows by assuming that the signal-to-noise ratio is much greater than 1 and  $\frac{E_i}{T_i} \gg P_{CT}$ . Step *b* follows from Borchardt's algorithm [106], which provides the following approximation which is the linearization of this function around zero:

$$T_i \log_2 T_i \approx -T_i \frac{6}{\ln 2}. \tag{6.11}$$

Therefore, we have

$$\begin{aligned}
\frac{T_i}{\sum_{i=1}^{N_d} T_i} &= \frac{\frac{D/B}{\log_2 \left( \frac{E_i}{d_i^n P_{n,i}} \right) + \frac{6}{\ln 2}}}{D/B \sum_{i=1}^{N_d} \frac{1}{\log_2 \left( \frac{E_i}{d_i^n P_{n,i}} \right) + \frac{6}{\ln 2}}} \\
&= \frac{1/(\log_2(E_i/d_i^n P_{n,i}) + 6/\ln 2)}{\sum_{i=1}^{N_d} 1/(\log_2(E_i/d_i^n P_{n,i}) + 6/\ln 2)}.
\end{aligned} \tag{6.12}$$

Thus, from (6.8) and (6.12), we have

$$T_i^* \approx \frac{E_0}{P_{CR}} \frac{1/(\log_2(E_i/d_i^n P_{n,i}) + 6/\ln 2)}{\sum_{i=1}^{N_d} 1/(\log_2(E_i/d_i^n P_{n,i}) + 6/\ln 2)}. \tag{6.13}$$

The above approximation is accurate around zero. The optimal transmit power for node *i* in the CH-constrained region can be easily calculated by

$$P_i^* = \frac{E_i}{T_i^*} - P_{CT}. \tag{6.14}$$

The maximum collected data from each node in the CH-constrained region is then

$$D^* \approx BT_i^* \log \left( 1 + \frac{P_i^*}{d_i^n P_{n,i}} \right). \quad (6.15)$$

### 6.2.3.2 CM-constrained region

In the CM-constrained region, the maximum possible total data gathered from each cluster member as  $E_0 \rightarrow \infty$  is determined by the minimum value of the following sequence

$$D^* = \min\{D_1, D_2, \dots, D_{N_d}\}, \quad (6.16)$$

where the operator  $\min\{\cdot\}$  returns the minimum element of a sequence, and  $D_i$  are the values of the following unconstrained maximization problems:

$$D_i = \max \left\{ BT_i \log \left( 1 + \frac{E_i - P_{CT}}{d_i^n P_{n,i}} \right) \right\}, \quad (6.17)$$

The above unconstrained maximization problems are conducted over  $T_i$ , and their analytical solutions can be found as

$$T_i^\dagger = \frac{\frac{E_i}{\ln 2(P_{CT} - d_i^n P_{n,i})}}{W\left(\frac{\ln 2}{2d_i^n P_{n,i}}(P_{CT} - d_i^n P_{n,i})\right) + P_{CT} - d_i^n P_{n,i}}, \quad (6.18)$$

where  $W(\cdot)$  is Lambert W function [107]. Although the Lambert W function can be calculated efficiently using numerical methods, it is still prohibitively complicated to calculate in nodes. However, it is possible to further simplify (6.18) when  $P_{CT} \gg d_i^n P_{n,i}$ . In this case,

$$\begin{aligned} T_i^\dagger &\approx \frac{E_i}{\frac{P_{CT} \ln 2}{W\left(\frac{\ln 2}{2d_i^n P_{n,i}} P_{CT}\right)} + P_{CT}}, \\ &\approx \frac{E_i}{\frac{P_{CT} \ln 2}{\ln\left(\frac{\ln 2}{2d_i^n P_{n,i}} P_{CT}\right)^{-3}} + P_{CT}}. \end{aligned} \quad (6.19)$$

Then

$$D^* = \min \left\{ BT_i^\dagger \log \left( 1 + \frac{E_i - P_{CT}}{d_i^n P_{n,i}} \right) \right\}, \quad (6.20)$$

where exact, (6.18) or approximate, (6.19) values of  $T_i^\dagger$  can be used.

Without loss of generality, assume  $D_1 \leq D_i, \forall i \in \{2, 3, \dots, N_d\}$ . Then we have

$$\begin{aligned} D^* &= D_1, \\ T_1^* &= T_1^\dagger. \end{aligned} \quad (6.21)$$

Moreover,  $T_i^* \forall i \in \{2, \dots, N_d\}$  can be any values that satisfy the following constraints:

$$\begin{aligned} T_i^* &> 0, \\ \sum_{i=1}^{N_d} T_i^* &\leq \frac{E_0}{P_{CR}}, \\ T_i^* \left[ \left( 2^{\frac{D^*}{BT_i^*}} - 1 \right) d_i^n P_{n,i} + P_{CT} \right] &\leq E_i. \end{aligned} \quad (6.22)$$

Another important observation is that in the CM-constrained region, an increase in the cluster head energy cannot increase the total number of bits collected from the CMs. Thus, once the cluster enters the CM-constrained region, that is, when the energy at the CH,  $E_0$ , is greater than a critical value,  $\hat{E}_0$ , an increase in the receiver energy becomes redundant. The value of  $\hat{E}_0$  can be determined by the following equation

$$\hat{E}_0 = P_{CR} \sum_{i=1}^{N_d} T_i^*, \quad (6.23)$$

where  $T_1^* = \hat{T}_1$  as defined by (6.21), and

$$T_i^* = \min \left\{ \arg \left[ BT_i \log \left( 1 + \frac{E_i - P_{CT}}{d_i^n P_{n,i}} \right) = D^* \right] \right\}, \quad (6.24)$$

$\forall i \in \{2, \dots, N_d\}$ , where  $D^*$  is from (6.20). Clearly the resulting  $T_i^* \forall i \in \{2, \dots, N_d\}$  also satisfy the constraint set (6.22).

### 6.2.3.3 Transmit power optimization algorithm

Based on the previous analysis, the transmit power optimization algorithm that finds the near optimal solution to model (6.6) is summarized in Fig. 6.3. As shown in Fig. 6.3, the first step is to determine the condition of the cluster of interest (CH-constrained

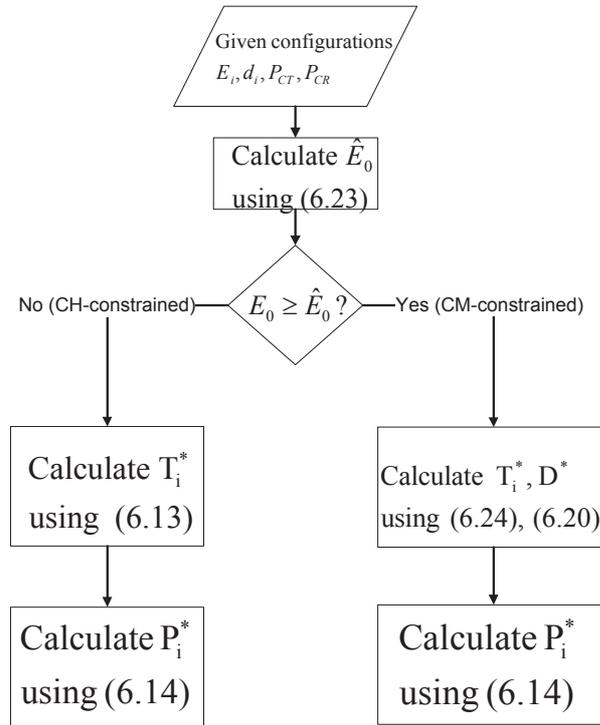


Figure 6.3: The transmit power optimization algorithm flowchart.

region or CM-constrained region); once the operating condition is determined, the results from sections 6.2.3.1 and 6.2.3.2 can be directly used. This algorithm does not require prohibitive numbers of iterations as in conventional numerical algorithms. The performance of the proposed algorithm will be evaluated in section 6.5.

### 6.3 Cluster Head Selection

The transmit power optimization algorithm proposed in the previous selection can also be used to identify the optimal cluster head that can provide the largest amount of total data gathered. An obvious approach is to repeat the transmit power optimization algo-

rithm  $N_d + 1$  times and determine which is the optimal cluster head. However, there are redundancies in this method. For example, some information generated by the proposed algorithm, such as the optimal transmit power, are not necessary in determining the optimal cluster head.

In this section, we propose a simplified method to identify the optimal cluster head, which only involves the calculation of (6.16) and (6.23) in each iteration. The simplified algorithm is based on a linear approximation around zero of the following functions:

$$D_i = BT_i \log \left( 1 + \frac{E_i - P_{CT}}{d_i^n P_{n,i}} \right). \quad (6.25)$$

First, we have the following observation:  $D^*$  has a near-linear relationship with the cluster head energy  $E_0$  in the CH-constrained region. The linearity is strengthened when  $P_{CT} \gg d_i^n P_{n,i}$  and  $T_i \rightarrow 0$ .

*Proof:* From (6.13), the optimal  $T_i$ s and the energy of the cluster head  $E_0$  has an approximately linear relationship in the CH-constrained region. Therefore, to show that  $D^*$  has a near-linear relationship with  $E_0$  in the CH-constrained region, we only need to show the linearity of  $D_i$  and  $T_i$ , for  $i \in \{1, 2, \dots, N_d\}$ :

$$\begin{aligned} D_i &= BT_i \log \left( 1 + \frac{E_i - P_{CT}}{d_i^n P_{n,i}} \right), \\ \implies D_i &= BT_i \left[ \log \left( \frac{P_{CT} - d_i^n P_{n,i}}{d_i^n P_{n,i}} \right) + \log \left( -1 + \frac{E_i}{(-d_i^n P_{n,i} + P_{CT})T_i} \right) \right], \\ \stackrel{a}{\implies} D_i &\approx BT_i \left[ \log \left( \frac{P_{CT} - d_i^n P_{n,i}}{d_i^n P_{n,i}} \right) + \log \left( \frac{E_i}{P_{CT}} \right) - \log T_i + \log \left( 1 - \frac{P_{CT} T_i}{E_i} \right) \right], \\ \stackrel{b}{\implies} D_i &\approx BT_i \left[ \log \left( \frac{P_{CT} - d_i^n P_{n,i}}{d_i^n P_{n,i}} \right) + \log \left( \frac{E_i}{P_{CT}} \right) + \frac{6}{\ln 2} \right]. \end{aligned}$$

The approximation  $a$  becomes accurate when  $P_{CT} \gg d_i^n P_{n,i}$ , and the approximation  $b$  becomes accurate when  $T_i \rightarrow 0$ . As shown by the above derivation,  $D_i$  and  $T_i$  have a near-linear relationship when  $P_{CT} \gg d_i^n P_{n,i}$  and  $T_i \rightarrow 0$ .  $\square$

The assumption  $P_{CT} \gg d_i^n P_{n,i}$  is usually true in small-scale narrowband systems where the noise power is low and transmission distances are in general short. However,  $T_i \rightarrow 0$  is not usually the case when the cluster operates in the CH-constrained region. In fact,  $T_i \rightarrow 0$  is true only when the cluster head has very limited energy. Therefore, in

most cases, the linear approximation is a coarse approximation to  $D^*$  as a function of  $E_0$ . Although the linear approximation cannot describe  $D^*$  very accurately, it provides good performance in the cluster head selection and greatly reduces the computation complexity.

Let  $D'(i)$  denote the linear approximation of the maximum data collected for the cluster with node  $i$  as the cluster head. The procedure for the proposed cluster head selection algorithm is straightforward: for each node  $i$ , calculate  $D'(i)$ ; then choose the node with the largest  $D'(i)$  as the cluster head.

In the proposed cluster head selection (CHS) algorithm,  $D'(i)$  can be expressed as

$$D'(i) = \begin{cases} D^* & E_i \geq \hat{E}_0 \\ D^* - \frac{D^*}{\hat{E}_0}(\hat{E}_0 - E_i) & E_i < \hat{E}_0 \end{cases} \quad (6.26)$$

where  $D^*$  can be obtained from (6.16) and  $\hat{E}_0$  can be calculated from (6.23).

The CHS algorithm can be distributed as follows (assuming each node knows the distances to its neighbors): first, each node uses one broadcast to inform the other nodes of its residual energy, and each node can then find its own  $D'$  through (6.26). The node with the most residual energy (assume it is node  $i$ ) broadcasts to the other nodes to declare its  $D'(i)$ . The rest of the nodes will compare the received  $D'(i)$  with their own  $D'(j)$ ,  $j \neq i$ . If  $D'(j) > D'(i)$ , then node  $j$  will notify node  $i$  of its  $D'(j)$ . Otherwise, node  $j$  does not take any action. At last, node  $i$  will compare the received information and use one broadcast to inform the rest of the nodes about the selected cluster head which provides the largest  $D'$ . The performance of the proposed CHS algorithm will be evaluated in section 6.5.

## 6.4 To Hop or Not to Hop

So far, we have obtained the optimal transmit power and the optimal selection of the cluster head in a single-hop cluster. To further improve the performance of a cluster,

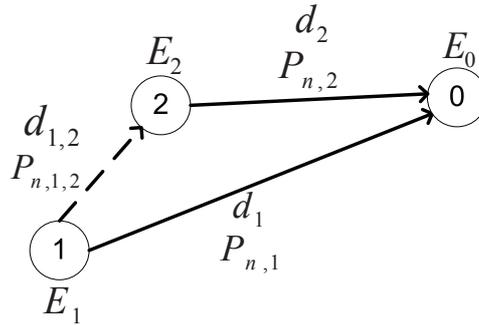


Figure 6.4: A 3-node cluster with a potential multi-hop route.

we investigate the use of multi-hop relaying within a cluster during the CM-constrained region.

We assume that the nodes have the ability to choose either direct communication or one-hop relaying to send information to the cluster head. Fig. 6.4 shows a 3-node cluster with multi-hop relaying. Node 1 can communicate to the cluster head (node 0) directly, or by relaying its data through node 2. As shown in Fig. 6.4, a multi-hop link may imply shorter transmission distance for node 1 than a direct link. However, a multi-hop link also incurs additional energy cost at the relay node. The energy consumption tradeoff is an important issue that deserves to be investigated.

It is also obvious that the multi-hop technique can only reduce the energy dissipation at the source node. Thus, multi-hop does not help when the energy storage at the CMs does not affect the cluster's performance, such as in a cluster operating in the CH-constrained region. In the CM-constrained region, the residual energy at the CMs determines the performance of the cluster and needs to be preserved to the maximum extent, and thus multi-hop relaying has enormous potential to improve the cluster's performance.

Let us start the analysis by considering the simple case of a 3-node cluster as depicted in Fig. 6.4. Assume that the cluster is operating in the CM-constrained region,

and the unconstrained maximum amount of data directly transmitted from node 1 and node 2 to the cluster head are  $D_1^*$  and  $D_2^*$  defined by (6.17), and  $D_1^* < D_2^*$ .

Under a strict fairness constraint as defined in section 6.2.2, a simple rule of thumb is that if  $D_1^* \geq \frac{D_2^*}{2}$ , then a single-hop cluster must be better than a multi-hop link; if  $D_1^* < \frac{D_2^*}{2}$ , a multi-hop link has the potential to outperform a single-hop cluster. This is because, in the multi-hop link, the relay node (node 2) has to transmit both the data received from node 1 and its own data. According to the strict fairness constraint, these two transmissions need to be of equal amount. That is, the maximum amount of data collected from each node cannot exceed  $\frac{D_2^*}{2}$  in a multi-hop link. On the other hand, the maximum amount of data collected from each node is  $D_1^*$  in a single-hop cluster. By comparing these two values, we have the above simple rule.

However,  $\frac{D_2^*}{2}$  is simply an upper bound on the maximum amount of data collected from each node in a multi-hop link. The detailed evaluation of clustering and multi-hop requires the knowledge of the exact value of the maximum amount of data collected from each node in a multi-hop link.

In the multi-hop link depicted in Fig. 6.4, it is observed that, when the energy at node 1 is smaller than a certain value, denoted by  $E_{1,2,\text{critical}}$ , the data collected through this multi-hop link is determined by node 1. That is, node 1 transmits the unconstrained maximum amount of data it could send to node 2,  $D_{1,2}^* = \max \left\{ T_1 B \log_2 \left[ 1 + \frac{E_1/T_1 - P_{CT}}{d_{1,2}^n P_{n,1,2}} \right] \right\}$ , where  $P_{n,1,2}$  represents the noise power on the link between node 1 and node 2. When the energy at node 1 is greater than  $E_{1,2,\text{critical}}$ , the data collected through this multi-hop route is determined by the relay node, node 2. That is, the energy storage at the relay node becomes the threshold for maximizing the data collected through the multi-hop route.

The exact expression of the transmitted data of the multi-hop route  $\tilde{D}_{1,2}$  can be found through the following rule: If  $E_1 \geq E_{1,2,\text{critical}}$ , then  $\tilde{D}_{1,2} = \hat{D}_{1,2}$ ; If  $E_1 < E_{1,2,\text{critical}}$ , then  $\tilde{D}_{1,2} = D_{1,2}^*$ , where  $D_{1,2}^* = \max \left\{ T_1 B \log_2 \left[ 1 + \frac{E_1/T_1 - P_{CT}}{d_{1,2}^n P_{n,1,2}} \right] \right\}$ ,  $\hat{D}_{1,2} =$

$$\frac{E_1}{\hat{P}_1 + P_{CT}} \log_2 \left( 1 + \frac{\hat{P}_1}{d_{1,2}^n P_{n,1,2}} \right),$$

$$\hat{P}_1 = -\frac{W \left( -\frac{a \ln 2}{c} 2^{b - \frac{ad}{c}} \right)}{a \ln 2} - \frac{d}{c}, \quad (6.27)$$

where  $a = \frac{D_2^*}{2E_1}$ ,  $b = \frac{D_2^*}{2} \left( \frac{P_{CT}}{E_1} - P_{CR} \right)$ ,  $c = 1$ , and  $d = \frac{1}{d_{1,2}^n P_{n,1,2}}$ .  $W(\cdot)$  is the lambert function. The value of  $E_{1,2,\text{critical}}$  can also be obtained analytically. That is

$$E_{1,2,\text{critical}} = \frac{\hat{D}_{1,2}}{\frac{B}{P_1^* + P_{CT}} \log_2 \left[ 1 + \frac{P_1^*}{d_{1,2}^n P_{n,1,2}} \right]}, \quad (6.28)$$

where

$$\begin{aligned} P_1^* &= \operatorname{argmax} \left\{ \frac{1}{P_1 + P_{CT}} \log_2 \left[ 1 + \frac{P_1}{d_{1,2}^n P_{n,1,2}} \right] \right\}, \\ &= \left\{ \frac{\left( \frac{P_{CT}}{d_{1,2}^n P_{n,1,2}} - 1 \right) \ln 2}{W \left[ \left( \frac{P_{CT}}{d_{1,2}^n P_{n,1,2}} - 1 \right) \right]} - 1 \right\} d_{1,2}^n P_{n,1,2}. \end{aligned} \quad (6.29)$$

Once the exact value of the maximum transmitted data of the multi-hop route  $\tilde{D}_{1,2}$  has been found, we can compare it with the maximum data using single-hop links  $D_1^*$ . That is, if  $\tilde{D}_{1,2} \geq D_1^*$ , then a multi-hop route is better than a single hop link, and vice versa.

This algorithm can be easily extended to a general clustering topology with  $N_d + 1$  nodes. Without loss of generality, we assume that  $D_1^* = \min \{D_i^*\}$ , then the algorithm can be stated as:

- Step 1** Compute the maximum amount of data that can be transmitted by node 1 though all possible relaying nodes  $i$ ,  $\tilde{D}_{1,i}$
- Step 2** Use node  $i$  with the highest  $\tilde{D}_{1,i}$  as the relay node of node 1;
- Step 3** Determine the analytical solution of the optimal transmit power  $P_1^*$  from (6.27).

**Step 4** Utilize the transmit power optimization algorithm to find the optimal transmit powers for the new topology.

**Step 5** Assume if the current maximum amount of gathered data is  $D^*$ , then if  $D^* \leq D_1^*$ , the original single-hop topology should be kept; otherwise, the multi-hop topology should be used.

The proposed algorithm to decide whether or not multi-hop should be used within a cluster can be applied together with the previously proposed transmit power optimization algorithm and CHS algorithm to further increase the amount of data collected by a cluster, when the cluster of interest is working in the CM-constrained region. The performance of this algorithm will be evaluated in the next section.

## 6.5 Results

### 6.5.1 Transmit Power Optimization Algorithm

In the transmit power optimization algorithm, the main result is the approximation made on the optimal operation time assignment  $T_i$  through (6.13) in the CH-constrained region. Compared with the solutions of  $T_i$  in the CM-constrained region (which are exact optimal solutions), (6.13) provides a near optimal approximation. Therefore, the effectiveness of the approximation needs to be evaluated. We assume that 5 nodes are uniformly placed within a disk with a radius of 100m, centered at the cluster head. The path loss exponent is  $n = 4$ . The circuit powers are  $P_{CR} = 30$  mW, and  $P_{CT} = 10$  mW. The signal bandwidth is  $B = 100$  KHz. The AWGN power is  $-116.5$  dBmW and is equal on all links.

Fig. 6.5 shows the maximum data gathered when all CMs have random residual energy  $E_i$  chosen from a uniform distribution between 1 J and 25 J. The cluster head has energy  $E_0$  varying from 0.01 to 0.5 J. This setup guarantees that the cluster works

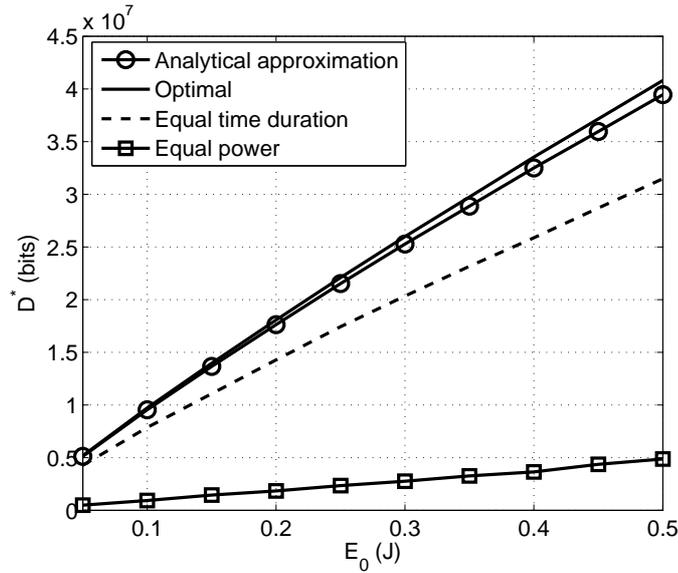


Figure 6.5: The maximum data gathered by clusters operating in the CH-constrained region.

in the CH-constrained region. We compare four scenarios: (i) the optimal solution to problem (6.6) through numerical methods; (ii) the proposed analytical approximation (iteration free solution); (iii) each node has an equal transmission time duration constraint, with numerically optimized transmit power and time duration; (iv) each node has an equal transmit power constraint, with numerically optimized transmit power and time duration.

The maximum amount of data collected is shown in Fig. 6.5. From Fig. 6.5, we can see that the proposed transmit power optimization algorithm provides a close approximation to the numerical optimization maximizing the amount of data collected in a CH-constrained cluster. The proposed transmit power optimization algorithm achieves a significant gain compared to the cluster with equal power and equal transmission time duration constraints. For instance, when  $E_0 = 0.3$  J, the cluster using the proposed transmit power optimization algorithm gathers an average of 1.2 times the maximum gathered data bits in the equal time duration case, and 8.9 times the number of the max-

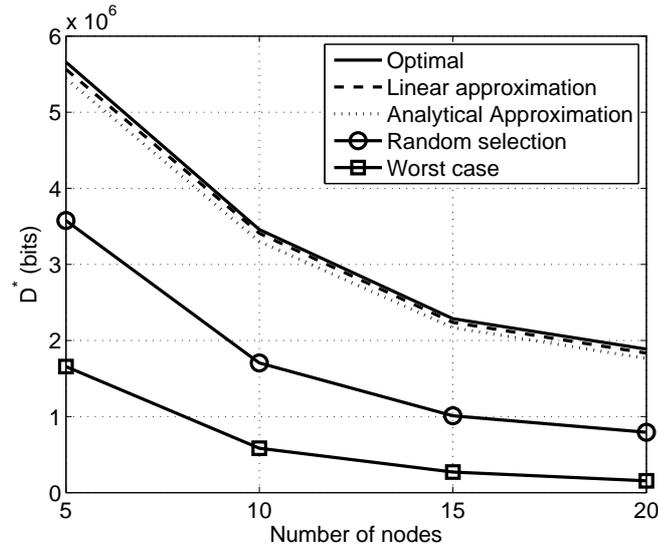


Figure 6.6: The maximum data gathered by clusters with different cluster head selection strategies.

imum gathered data bits in the equal transmit power case. Also shown in the figure is that the analytical approximation becomes more accurate as the cluster head energy decreases, which agrees with the analysis in section 6.2.

### 6.5.2 CHS Algorithm

The performance of the proposed CHS algorithm is evaluated, and the results are shown in Fig. 6.6. The residual energy of the nodes is generated using a random variable  $s$ , which is uniformly distributed between 0 and 1, through

$$E_i = 2s, \forall i \in \{0, 1, \dots, N_d\}, \quad (6.30)$$

where  $N_d \in \{4, 9, 14, 19\}$ , and the nodes are placed within a disk with a radius 100 m following a uniform distribution.

In Fig. 6.6, the term *linear approximation* means that the proposed CHS algorithm by (6.26) is used to select the cluster head; while the term *analytical approximation*

refers to a brute force cluster head selection approach, which uses the proposed transmit power optimization algorithm in each iteration, to select the optimal cluster head. The term *Optimal* refers to a brute force cluster head selection approach, which uses a numerical solution of model (6.6) through an interior-point method in each iteration. Moreover, the performance of a random selection and a worst case selection of the cluster head is provided. After cluster head selection, the proposed transmit power optimization algorithm is used to configure the transmit powers of the CMs in all cases.

Fig. 6.6 shows that the proposed CHS algorithm (linear approximation) provides a sound approximation to the optimal cluster head selection in terms of performance. That is, the maximum amount of data collected by the cluster using the CHS algorithm is almost as much as that of a cluster using the numerically optimized cluster selection, while the CHS algorithm avoids the iterations in the numerical optimization. Therefore, the proposed CHS algorithm has great potential in SRWN applications. Also shown in Fig. 6.6, the analytical approximation method also provides a sound performance, although compared with the proposed CHS algorithm, it has slightly worse performance and higher complexity. The proposed CHS algorithm has a significant performance gain over the random selection and worst case selection. For example, when there are ten nodes in the cluster ( $N_d = 9$ ), the cluster using the proposed CHS algorithm with linear approximation can collect 1.9 times the average number of maximum data bits collected by the clusters with the random selection, and 5.4 times the average number of maximum data bits collected by the clusters with the worst case selection.

### 6.5.3 To Hop or Not to Hop

We assume that two nodes are placed within a disk with a radius 100 m, according to a uniform distribution, centered at the cluster head. The cluster head has a large amount of energy. The cluster members have energy uniformly distributed over  $[1, E_{i,max}]$ , where  $E_{i,max} \in \{5, 10, 15, 20\}$ . We compare the maximum amounts of data gathered

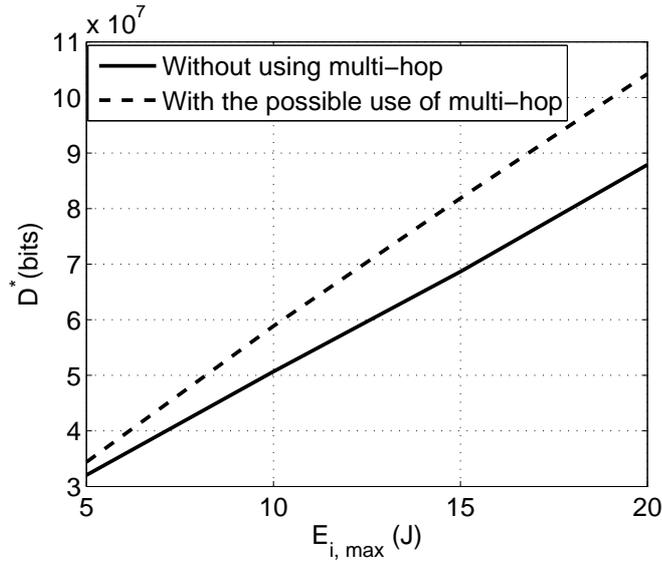


Figure 6.7: Performance comparison of optimization with/without the possible use of a multi-hop.

at the cluster head in a cluster with and without the possible use of multi-hop. The results are shown in Fig. 6.7.

As shown in Fig. 6.7, the possible use of multi-hop always provides a performance gain. Moreover, the gain of the possible use of multi-hop increases as  $E_{i,max}$  increases. For instance, the clusters with the possible use of multi-hop can collect 1.09 times the average number of maximum data bits collected by the clusters without using multi-hop when  $E_{i,max} = 5$  J, and 1.21 times the average number of maximum data bits collected by the clusters without using multi-hop when  $E_{i,max} = 20$  J. This is because a large  $E_{i,max}$  introduces a large variance in the energy distribution and thereby a large variance on the nodes' ability to transmit information bits to the cluster head. Therefore, as  $E_{i,max}$  increases, it becomes more likely that multi-hop improves the performance (when  $D_i^* < \frac{D_j^*}{2}$ , for some  $i$  and  $j$ ).

### 6.5.4 Overall Performance Gain

The optimal transmit power, CHS algorithm, and the possible use of multi-hop can be naturally combined together to increase the maximum amount of collected data in a cluster. That is, for a given group of nodes, we can construct an optimal cluster with the optimal cluster head, the optimal transmit power at the cluster members, and the possible use of multi-hop. In comparison, we refer to an unoptimized cluster as a cluster with randomly selected cluster head, equal transmit time duration at the cluster members, and no multi-hop. We compare the amount of data collected by an optimal cluster and an unoptimized cluster and show the results in Fig 6.8. In Fig. 6.8, we also show the performance of clusters with the optimal cluster head and the optimal transmit power at the cluster members, as well as the performance of clusters with only the optimal transmit power at the cluster members.

The residual energy of the nodes is generated from a random variable  $s$  uniformly distributed between 0 and 1, through

$$E_i = m(s + 1), \quad \forall i \in \{0, 1, \dots, N_d\}, \quad (6.31)$$

and  $m \in \{5, 10, 15, 20, 25\}$ . The nodes are placed within a disk with a radius 100 m following a uniform distribution, and there are 5 nodes in the network (i.e.,  $N_d = 4$ ).

As shown in Fig. 6.8, the performance gain of the optimization (the optimal cluster head, the optimal transmit power at the cluster members, and the possible use of multi-hop) is significant, especially when the residual energy variation between the nodes is large. For instance, when  $m = 25$ , the optimal cluster can collect an average amount of bits of data throughout its lifetime that is 2.6 times the average number of maximum data bits collected by the unoptimized cluster. The clusters with the optimal transmit power at the CMs, the optimal CH, and no multi-hop can collect twice the average number of maximum data bits collected by the unoptimized cluster, while the clusters with only the optimal transmit power at the CMs can collect 1.33 times the average number of maximum data bits collected by the unoptimized cluster. The clusters with

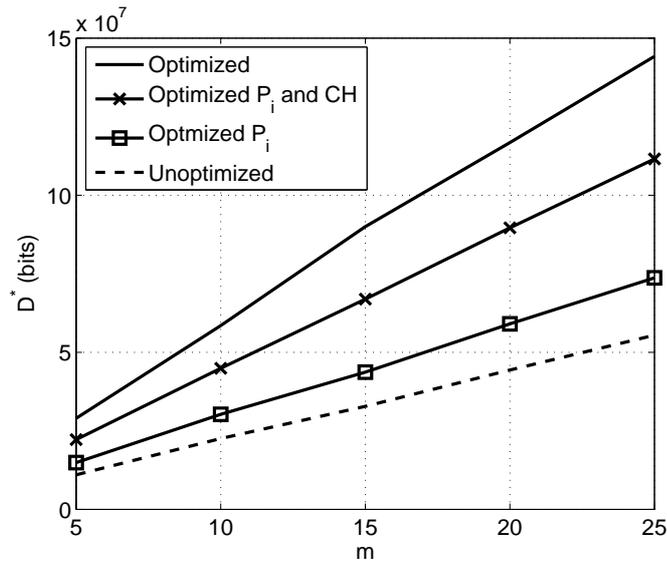


Figure 6.8: Performance comparison of the optimized and unoptimized cluster.

the optimal transmit power at the CMs, the optimal CH, and no multi-hop can collect 1.51 times the average number of maximum data bits collected by the clusters with only the optimal transmit power at the CMs. Therefore, the optimal selection of the cluster head is very important in improving the cluster's performance, as the cooperation of the optimal cluster head selection and the optimal transmit power at CMs provides a significant additional gain over the use of optimal transmit power at CMs alone.

This performance gain increases with an increase in the variation of the residual energy of the nodes. This is because, first, the possible use of multi-hop will provide more performance gain when there is large energy variation. Second, when there are large differences between nodes' conditions (distances and residual energy), the importance of the optimization of transmit power and cluster head selection becomes more significant, as a random choice of clustering in this case may be poor, e.g., by random selection, a node with little energy left and long distances to the rest of the nodes may be chosen as a cluster head. Since usually there is a wide range of conditions for different nodes, the optimization of the clustering can be highly beneficial.

## 6.6 Summary

In this chapter, for the transmit power optimization, we proposed an iteration-free algorithm for maximizing the amount of data gathered by a cluster throughout its lifetime, considering energy constraints and a strict data fairness constraint. This framework can be easily adapted to a fixed-rate constraint case. Moreover, we developed a cluster head selection (CHS) algorithm to determine the optimal cluster head that provides the largest amount of collected data for a cluster. Furthermore, we present an algorithm to decide whether or not multi-hop should be used within a cluster to further improve the performance of a cluster. The performance gain by using the optimal transmit power, the optimal cluster head selection, and the possible use of multi-hop has been shown to be significant.

## **7 Maximizing Gathered Samples in Wireless Networks with Slepian-Wolf Coding**

In this chapter, we look at the cross-layer design of the data link and application layers, considering an energy constrained data-gathering wireless network. We consider an energy constrained wireless sensor network, with arbitrary number of nodes, where source nodes utilize Slepian-Wolf (SW) coding before transmission to a joint decoder. We investigate optimal and near-optimal SW coding rates, transmit powers, and transmit durations that maximize the number of collected samples during the network lifetime, subject to channel capacity, SW rate region, and residual energy constraints. We find optimal (near-optimal) closed-form solutions in the absence (presence) of an energy constraint at the joint decoder. We take into account the energy consumption of SW encoding and decoding and communication circuitry. Numerical results demonstrate the effectiveness of the proposed optimization, especially when the joint decoder is not energy constrained.

## 7.1 Introduction

Exploiting the spatial correlation of sensed data, often present in wireless sensor networks (WSNs), is crucial for the energy efficient operation of a battery powered WSN. Distributed Source Coding (DSC) is a promising compression approach that can reduce the energy consumption in WSNs by extracting the spatial redundancy in the sensed data at the nodes, without requiring the nodes to share their measurements with each other [62]-[65]. The Slepian-Wolf (SW) rate region determines the minimum DSC sum-rate required for lossless recovery of the discrete correlated sources. In general there are infinitely many sets of DSC rates that achieve this minimum sum-rate. In practice, sensors have different conditions (i.e., distance to the destination and residual energy). Moreover, the cost of transmitting one bit from each sensor is affected by the communication parameters chosen (i.e., transmission power and duration). Therefore, the utility of the network can be significantly improved if the DSC rates and communication parameters are jointly optimized.

Researchers have considered the use of theoretical DSC results, in particular SW coding, to enable energy efficient gathering of correlated data in WSNs [78]-[73]. The utility metric considered in these existing studies varies. Oftentimes, the goal is to minimize the total energy consumption, subject to certain constraints. Several works use a *bit-hop* utility metric to measure the energy expenditure [108][79], while others use the sum of the number of bits transmitted over a link, weighted by a function of the distance between two sensors on that link [72]-[109][74]. Total energy consumption, however, is not the best utility metric to consider since it does not take into account the fact that different sensors cannot share their energy. Rather, they likely have different energy reserves (residual energy) and energy costs (distance from the destination).

Alternatively, one can assume some residual energy in each node and consider the total utility of the network in terms of lifetime [110][111] or the number of collected samples [73]. The total number of collected samples is a reasonable utility metric since

it reflects the performance of a data-gathering wireless network. However, in using this utility metric, the existing literature generally assume that the cost of transmitting one bit over a link is independent of transmit power or transmission rate, thus assuming a linear relationship between transmission rate and power consumption. In reality, the logarithmic nature of channel capacity tells us that as the transmission rate over one link increases, the cost of transmitting one bit increases as well.

The existing literature also vary in terms of the network model and optimization variables. The majority of the existing works consider a multi-hop setting and seek to jointly optimize the DSC rates and the routing paths [79]-[76][74] or the flow rates [109]-[73] at the network layer. Among these, only [110] has considered physical layer parameters for optimization. The results, however, are numerical. In contrast, [108] focuses on a single hop setting, and studies the tradeoff between sum-rate and loss factor, and does not seek to optimize the system parameters. With a more practical point of view, [112] and [113] study the joint optimization of DSC rates and communication parameters in a lossy setting. Other works [114]-[117] aim to reduce the computational complexity of the optimal DSC rate and power allocation problem, generally by further constraining the optimization problem and enforcing pair-wise DSC among nodes.

Another important issue, arising in practice, is the energy consumption in the circuitry for compression and transmission. None of the existing literature use energy consumption models that take these energy costs into account.

In this work, we consider this problem with an analytical approach, while considering a more detailed and realistic optimization problem. We consider a data-gathering WSN, where each sensor (source node) generates a fixed number of samples and transmits the encoded samples to a joint decoder. We assume that the medium is shared by the sources through an interference-free multi-access scheme. This maximizes the sum-rate when considering circuit power consumption [118]. We choose the total number of collected samples as the utility metric. Our optimization problem takes as input the distances of the sensors from the joint decoder, the residual energy of each node,

the circuit power consumption, the cost of SW encoding and decoding, the channel model (AWGN or Rayleigh fading), and the SW coding rate region constraints<sup>1</sup>. We employ the Shannon channel capacity to establish a realistic relationship between the cost of transmitting each bit, transmit power and rate. We then maximize the utility and provide jointly optimal SW coding rates, transmit powers and transmit durations that achieve this maximum value, all in closed-form expressions. Our optimization incurs little communication overhead. For the scenarios where the nodes have fixed communication distances and the sources have fixed correlations, the proposed optimization requires the nodes to initially communicate with the joint decoder about their residual energy and distances to the joint decoder. The joint decoder broadcasts the jointly optimal solutions to the nodes. Therefore, only one round of communication is needed to implement our solution.

Major contributions of this work are (i) considering a more detailed optimization problem by including circuit power consumption, energy constraint at the joint decoder, SW encoding and decoding costs, channel capacity constraints, and different communication channel models, (ii) extending the optimization parameter space beyond SW coding rates (via enabling adaptive transmit powers and transmit durations), thus further enhancing the network energy efficiency, and (iii) providing closed-form solutions to the optimization problem, with low computational complexity. To do this, we assume that a cluster consisting of a number of sensors and a joint decoder (cluster head) has already been formed. To form the cluster, any existing clustering technique, such as linked cluster algorithm (LCA) and random competition based clustering (RCC) [122], may be used.

---

<sup>1</sup>In this work, as generally assumed in the literature [72]-[109][73][74], we assume that the joint entropy of the sources are known. The problems of estimating the field entropy and joint probability have been previously addressed in the literature [119]-[121].

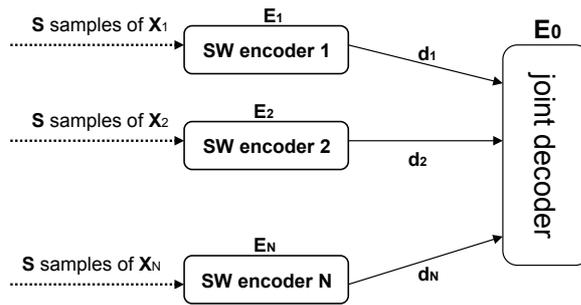


Figure 7.1: A data-gathering WSN with  $N$  source nodes where each node employs SW coding for distributed compression and the joint decoder reconstructs all the collected samples losslessly.

## 7.2 Problem Formulation

Consider a network with  $N$  sensors. Each sensor collects  $S$  samples<sup>2</sup> during the network lifetime and transmits the encoded samples over orthogonal channels to a joint decoder (Fig. 7.1). Let  $X_i$  and  $R_i$  (bits/sample) denote the random sample and the SW coding rate at node  $i$ , respectively. The SW theorem [63] states that  $X_i$  can be perfectly reconstructed at the joint decoder, if and only if  $R(G) \geq H(X(G)|X(G^c)), \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset$ , where  $H(\cdot)$  is the entropy function,  $R(G) = \sum_{i \in G} R_i$ ,  $X(G) = \{X_j | j \in G\}$ , and  $G^c$  is the complement of  $G$ .

Assume that the transmit duration of node  $i$  is  $T_i$ , during which it transmits  $SR_i$  data bits to the joint decoder. Invoking the AWGN channel capacity constraint we obtain  $BT_i \log_2 \left( 1 + \frac{P_i}{d_i^n N_0} \right) \geq SR_i$ , where  $B$  is the channel bandwidth,  $P_i$  is the transmit power of node  $i$ ,  $d_i$  is the distance from node  $i$  to the joint decoder,  $n$  is the path loss exponent, and  $N_0$  is the noise power. Therefore, we can find a lower bound on  $P_i$

$$P_i \geq (2^{\frac{SR_i}{BT_i}} - 1) d_i^n N_0.$$

The energy constraints at sensor  $i$  and at the joint decoder are  $T_i(P_i + P_{CT}) \leq E_i$ ,

---

<sup>2</sup>DSC exploits the correlation between samples generated by different sources [64][65], and therefore, it requires the same number of samples generated by each source node.

and  $P_{CR} \sum_{i=1}^N T_i \leq E_0$ , where  $E_0$  and  $E_i, i = 1, \dots, N$ , denote the residual energy at the joint decoder and at node  $i$ , respectively,  $P_{CT}$  and  $P_{CR}$  denote the circuit power consumption during transmission and reception, respectively. Maximization of the total gathered samples can be formulated as

$$\begin{aligned}
 (\mathbf{P7.1}) \quad & \min -S \\
 \text{s.t.} \quad & C'_0 : T_i > 0, S > 0, \quad C'_1 : P_{CR} \sum_{i=1}^N T_i \leq E_0, \\
 & C'_2 : T_i(P_i + P_{CT}) \leq E_i, \\
 & C'_3 : P_i \geq (2^{\frac{SR_i}{BT_i}} - 1)d_i^n N_0, \\
 & C'_4 : R(G) \geq H(X(G)|X(G^c)), \\
 & \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset.
 \end{aligned}$$

Here,  $C'_1$  and  $C'_2$  are the energy constraints at the nodes.  $C'_3$  is the channel capacity constraint, and  $C'_4$  is the SW coding rate region constraint. Note that  $S$  is allowed to be non-integer. Problem (P7.1) is a minimum feasible set test problem with variables  $S$ ,  $R_i$ ,  $P_i$ , and  $T_i$  [105].

Let  $(S^*, R_i^*, \hat{P}_i, T_i^*)$  be an optimal solution of (P7.1), where  $\hat{P}_i = (2^{S^* R_i^* / BT_i^*} - 1)d_i^n N_0 + \epsilon_i$  for  $\epsilon_i > 0$ . Then  $(S^*, R_i^*, P_i^*, T_i^*)$ , where  $P_i^* = (2^{S^* R_i^* / BT_i^*} - 1)d_i^n N_0$ , is also an optimal solution of (P7.1). Thus, we can replace  $C'_3$  with an equality, which yields

$$\begin{aligned}
 S^* R_i^* &= BT_i^* \log_2 \left( 1 + \frac{P_i^*}{d_i^n N_0} \right) \\
 \Rightarrow S^* \sum_{i=1}^N R_i^* &= \sum_{i=1}^N BT_i^* \log_2 \left( 1 + \frac{P_i^*}{d_i^n N_0} \right) \\
 \Rightarrow S^* &= \frac{\sum_{i=1}^N BT_i^* \log_2 \left( 1 + \frac{P_i^*}{d_i^n N_0} \right)}{\sum_{i=1}^N R_i^*}.
 \end{aligned}$$

We observe that, given  $P_i^*$  and  $T_i^*$ ,  $S^*$  is a decreasing function of  $\sum_{i=1}^N R_i^*$ . Moreover, the SW rate region dictates  $\sum_{i=1}^N R_i^* \geq H(X_1, X_2, \dots, X_N)$ . Hence, given  $P_i^*$  and  $T_i^*$ ,  $S^*$  attains its maximum if and only if  $\sum_{i=1}^N R_i^* = H(X_1, X_2, \dots, X_N)$ . Consequently

(P7.1) reduces to

$$\begin{aligned}
(\text{P7.2}) \quad & \min -S \\
\text{s.t. } & C_0 : T_i > 0, S > 0, \quad C_1 : P_{CR} \sum_{i=1}^N T_i \leq E_0, \\
& C_2 : T_i [(2^{\frac{SR_i}{BT_i}} - 1) d_i^n N_0 + P_{CT}] \leq E_i, \\
& C_3 : R(G) \geq H(X(G)|X(G^c)), \\
& \quad \forall G \subset \{1, 2, \dots, N\}, G \neq \{1, 2, \dots, N\}, G \neq \emptyset, \\
& C_4 : \sum_{i=1}^N R_i = H(X_1, X_2, \dots, X_N)
\end{aligned}$$

where variables  $P_i$  have been eliminated from the optimization problem.

Here we assume that the orthogonal channels between sensors and the joint decoder are modeled as AWGN. Our formulation can be easily extended to a Rayleigh fading case, by changing the channel capacity expression from  $\mathcal{C}_{\text{AWGN}} = \log_2 \left( 1 + \frac{P_i}{d_i^n N_0} \right)$  to  $\mathcal{C}_{\text{Rayleigh}} \approx \log_2 \left( 1 + \frac{\sigma_i P_i}{2d_i^n N_0} \right)$ , where  $\sigma_i$  is the mean of the channel amplitude, assuming that channel state information is available at the joint decoder [123]. This approximation is accurate at high SNR (if  $\sigma_i P_i > 8d_i^n N_0$ ).

### 7.3 maximizing collected samples without joint decoder energy constraint

Let us first assume that the joint decoder is not energy constrained, e.g. when it has a plugged-in power supply (See Section 7.4 for the energy constrained joint decoder). Then  $C_1$  in (P7.2) can be omitted. Proposition 1 provides closed-form solutions for the optimal  $R_i^*$  and  $T_i^*$  that produce  $S^*$ .

**Proposition 1:** Let us define  $D_i^\dagger = \max_{T_i} \left\{ BT_i \log_2 \left( 1 + \frac{E_i - P_{CT} T_i}{T_i d_i^n N_0} \right) \right\}$  as the maximum number of information bits that node  $i$  can send, in the absence of an energy constraint at the joint decoder, and let  $D^\dagger(G) = \sum_{i \in G} D_i^\dagger$ . The optimal  $R_i^*$  and  $T_i^*$  are

as follows.

$$R_1^* = \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\},$$

$$T_1^* = E_1 \left( \frac{\ln 2(P_{CT} - d_1^n N_0)}{\mathcal{W}\left(\frac{\ln 2}{2d_1^n N_0}(P_{CT} - d_1^n N_0)\right)} + P_{CT} - d_1^n N_0 \right)^{-1}.$$

$R_j^*$  for  $j \geq 2$  can be any feasible solution that satisfies  $\sum_{j=2}^N R_j^* = H(X_1, X_2, \dots, X_N) - R_1^*$ . Given  $R_j^*, T_j^*$  must satisfy  $(P_j + P_{CT})T_j^* \leq E_j$ , where  $P_j = d_j^n N_0(2^{D_1^\dagger R_j^*/(BT_j^* R_1^*)} - 1)$ . In particular, if  $(P_j + P_{CT})T_j^* = E_j$ , then  $T_j^*$  is

$$T_j^* = \left( -\frac{\mathcal{W}\left(-\frac{a \ln 2}{b} 2^{-\frac{ad}{b}}\right)}{a \ln 2} - \frac{d}{b} \right)^{-1}, \quad (7.1)$$

where  $\mathcal{W}(\cdot)$  is the Lambert function,  $a = D_1^\dagger \frac{R_j^*}{BR_1^*}$ ,  $b = \frac{E_j}{d_j^n N_0}$ , and  $d = 1 - \frac{P_{CT}}{d_j^n N_0}$ .

Before proving Proposition 1 we need the following lemma.

**Lemma 1:** Define  $D_i^* = S^* R_i^*$  as the optimal number of information bits that node  $i$  can send. For at least one node we have  $D_i^* = D_i^\dagger$ .

**Proof of Lemma 1:** Considering **(P7.2)** we note that  $C_3$  is independent of  $S^*$ . Hence,  $S^*$  is only constrained by  $C_2$ . From  $C_2$  we have  $S^* \leq (BT_i/R_i^*) \log_2(1 + (E_i - P_{CT}T_i)/(T_i d_i^n N_0))$ . Therefore,  $S^*$  can be expressed as

$$S^* = \min_i \left\{ \frac{1}{R_i^*} \max_{T_i} \left[ BT_i \log_2 \left( 1 + \frac{E_i - P_{CT}T_i}{T_i d_i^n N_0} \right) \right] \right\}$$

$$= \min_i \left\{ \frac{D_i^\dagger}{R_i^*} \right\}, \quad (7.2)$$

implying that at least one of  $D_i^*$  is equal to  $D_i^\dagger$ .  $\square$

Without loss of generality we assume that  $i = 1$  satisfies the minimization in (7.2), i.e.  $D_1^* = D_1^\dagger$ .

**Proof of Proposition 1:** Assume that node  $i$  sends  $D_i^* = S^* R_i^*$  bits and  $D_1^* = D_1^\dagger$ .

Then

$$\begin{aligned}
\sum_{i=1}^N D_i^* &= S^* \sum_{i=1}^N R_i^* \\
\stackrel{\text{I}}{\implies} S^* &= \frac{1}{H(X_1, \dots, X_N)} \sum_{i=1}^N D_i^* \\
\stackrel{\text{II}}{\implies} S^* &= \frac{D_1^\dagger}{H(X_1, \dots, X_N)} \left( 1 + \frac{1}{R_1^*} \sum_{j=2}^N R_j^* \right).
\end{aligned} \tag{7.3}$$

Step I results from  $\sum_{i=1}^N R_i^* = H(X_1, \dots, X_N)$ , and step II follows from  $\frac{D_i^*}{R_i^*} = S^*$ . Examining (7.3), we see that  $S^*$  attains its maximum if  $R_1^*$  assumes its smallest feasible value and  $\sum_{j=2}^N R_j^*$  assumes its largest feasible value, subject to  $R_1^*$  taking its smallest feasible value. First, we find the smallest feasible value of  $R_1^*$ . We have

$$D_1^\dagger \frac{R^*(G)}{R_1^*} = D^*(G) \leq D^\dagger(G) \Rightarrow R^*(G) \leq \frac{D^\dagger(G)R_1^*}{D_1^\dagger}.$$

where  $R^*(G) = \sum_{i \in G} R_i^*$  and  $D^*(G) = \sum_{i \in G} D_i^*$ . Combining with the SW coding rate region constraint  $R^*(G) \geq H(X(G)|X(G^c))$  we have

$$\frac{D^\dagger(G)R_1^*}{D_1^\dagger} \geq H(X(G)|X(G^c)) \Rightarrow R_1^* \geq \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)},$$

implying that the smallest feasible value of  $R_1^*$  is  $R_1^* = \max_G \left\{ (H(X(G)|X(G^c))D_1^\dagger) / D^\dagger(G) \right\}$ . Furthermore,  $T_1^* = \arg \max_{T_1} \left\{ BT_1 \log_2 \left( 1 + \frac{E_1 - P_{CT}T_1}{T_1 d_1^n N_0} \right) \right\}$  leading to the  $T_1^*$  expression given in the proposition statement. Knowing  $R_1^*$  and  $D_1^\dagger$  we can find  $S^* = D_1^\dagger / R_1^*$ . The largest feasible value of  $\sum_{j=2}^N R_j^*$ , subject to  $R_1^*$  taking its smallest feasible value given above is  $\sum_{j=2}^N R_j^* = H(X_1, X_2, \dots, X_N) - R_1^*$ . We note that optimal  $R_i^*$  are not unique for  $N > 2$  and any non-negative values that satisfy

$$\begin{aligned}
R_1^* &= \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\}, \\
\sum_{j=2}^N R_j^* &= H(X_1, X_2, \dots, X_N) - R_1^* \\
R^*(G) &\geq H(X(G)|X(G^c)), \\
\forall G &\subseteq \{1, 2, \dots, N\}, G \neq \{1, 2, \dots, N\}, G \neq \emptyset,
\end{aligned} \tag{7.4}$$

are optimal. In particular, if

$$\begin{aligned}
R_1^* &= \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\} \\
&= \frac{H(X(g)|X(g^c))D_1^\dagger}{D^\dagger(g)}, \quad \text{where } g \in G,
\end{aligned} \tag{7.5}$$

then  $R_i^* = D_i^\dagger R_1^* / D_1^\dagger$  for  $i \in g$  and  $R^*(g^c) = \sum_{i \in g^c} R_i^* = H(X_1, X_2, \dots, X_N) - R^*(g)$ .

For  $N > 2$ , the optimal  $T_j^*$  for  $j \geq 2$  are also not unique, as long as  $T_j^*$  guarantees node  $j$  can transmit  $D_j^* = D_1^\dagger R_j^* / R_1^*$  information bits to the joint decoder and the energy constraint  $C_2$  in (P7.2) is met. One optimal solution can be obtained from solving

$$BT_j^* \log_2 \left( 1 + \frac{E_j - P_{CT} T_j^*}{T_j^* d_j^n N_0} \right) = D_1^\dagger \frac{R_j^*}{R_1^*}.$$

assuming that all nodes use up their entire energy, where the solutions are given in (7.1).

□

We note that the complexity of obtaining the optimal set  $R_i^*$  and  $T_i^*$  (through Proposition 1) is low for the following reasons: First, as long as the system parameters are fixed, this calculation is performed only once. All the SW coding rates, transmit powers and durations are calculated and sent to the sensors for use throughout the network lifetime. Second, the Lambert function can be efficiently calculated [124]. Third, while the complexity of finding  $R_1^*$  in (7.5) grows exponentially with  $N$ , in practice  $N$  will not be large. This is due to the fact that (i) the number of sensors with highly correlated measurements (where DSC is effective) is not expected to be large (ii) often local clustering techniques divide the nodes into groups with smaller numbers of nodes.

## 7.4 maximizing collected samples with joint decoder energy constraint

In general the energy storage at the joint decoder may be finite, leading to one of two scenarios: (i) the joint decoder's energy is sufficiently high such that it does not impose any constraint, and (ii) the joint decoder's energy is a limiting factor. We refer to such a network as a *joint decoder limited network*. Scenario (i) is equivalent to the case where there is no energy constraint at the joint decoder, which was addressed in Section 7.3.

In this section we consider scenario (ii) and in Section 7.5 we will study the delineation of these two scenarios.

In a joint decoder limited network, the total transmit duration is limited by the energy storage at the joint decoder, i.e.,  $\sum_{i=1}^N T_i^* = E_0/P_{CR}$ . The optimal transmit power of sensor  $i$  can be upper bounded as  $P_i^* \leq E_i/T_i^* - P_{CT}$ . To maximize the number of transmit information bits in a given duration, the sensors will spend all their energy, i.e.,  $P_i^* = E_i/T_i^* - P_{CT}$ . This means that both the joint decoder and the sensors use up their energy. Based on this observation and a linear approximation of problem (P7.2), Propositions 2 and 3 provide closed-form solutions for the near-optimal  $R_i^*$  and  $T_i^*$  that produce near-optimal  $S^*$ , respectively.

**Proposition 2.** In a joint decoder limited network, the near-optimal  $R_i^*$  are  $R_1^* = H(X_1)$ ,  $R_2^* = H(X_2|X_1)$ ,  $R_3^* = H(X_3|X_1, X_2)$ , ...,  $R_N^* = H(X_N|X_1, X_2, \dots, X_{N-1})$ , where we have  $E_1/d_1^n \geq E_2/d_2^n \geq E_3/d_3^n \geq \dots \geq E_N/d_N^n$ .

**Proof:** Both the joint decoder and the sensors' residual energy constraints are active

$$P_{CR} \sum_{i=1}^N T_i^* = E_0, \quad T_i^* \left( \left( 2^{\frac{S^* R_i^*}{B T_i^*}} - 1 \right) d_i^n N_0 + P_{CT} \right) = E_i. \quad (7.6)$$

From (7.6), we have

$$\begin{aligned} & \log_2 \left( \frac{E_i/T_i^* - P_{CT}}{d_i^n N_0} + 1 \right) = \frac{S^* R_i^*}{B T_i^*} \\ \stackrel{\text{I}}{\implies} & T_i^* \log_2 \frac{E_i}{T_i^* d_i^n N_0} \approx \frac{S^* R_i^*}{B} \\ \implies & T_i^* \log_2 \frac{E_i}{d_i^n N_0} - T_i^* \log_2 T_i^* \approx \frac{S^* R_i^*}{B}, \\ \stackrel{\text{II}}{\implies} & T_i^* \approx \frac{S^* R_i^* / B}{\log_2 \frac{E_i}{d_i^n N_0} + \frac{6}{\ln 2}}, \end{aligned} \quad (7.7)$$

where step I assumes that  $(E_i/T_i - P_{CT})/(d_i^n N_0) \gg 1$  and  $E_i T_i \gg P_{CT}$ . The first approximation is equivalent to assuming a high SNR. The second approximation means that much more energy is consumed by transmission, rather than the circuit components. This approximation is reasonable in the joint decoder limited network, where

the transmission durations are mostly determined by the limited energy of the joint decoder and are reasonably short. In this scenario, short transmission durations and large power need to be used.

Step II employs the linear approximation  $T_i^* \log_2(T_i^*) \approx -6T_i^*/\ln 2$ , which is accurate for small  $T_i^*$ , a reasonable assumption if  $E_0$  is small. Defining  $U_i = [\log_2(E_i/d_i^n N_0) + 6/\ln 2]^{-1}$  and using (7.7) we find  $T_i^* \approx S^* U_i R_i^*/B$  and consequently  $\sum_{i=1}^N T_i^* \approx \frac{S^*}{B} \sum_{i=1}^N R_i^* U_i$ . On the other hand, we have  $\sum_{i=1}^N T_i^* = E_0/P_{CR}$ . Combining these and solving for  $S^*$  we obtain

$$S^* \approx \frac{BE_0}{P_{CR}} \left( \sum_{i=1}^N R_i^* U_i \right)^{-1}. \quad (7.8)$$

Examining (7.8) we see that  $S^*$  attains its maximum value if  $\sum_{i=1}^N R_i^* U_i$  assumes its minimum feasible value. Hence, the near-optimal  $R_i^*$  can be found by solving the following inequality form linear programming (we refer to the solutions  $R_i^*$  of (P7.3) as *near-optimal*, since the linear approximation we used to form (P7.3) is accurate for small  $T_i^*$ )

$$\begin{aligned} \text{(P7.3)} \quad & \min \sum_{i=1}^N R_i U_i \\ & \text{s.t. } R(G) \geq H(X(G)|X(G^c)), \\ & \quad \forall G \subseteq \{1, 2, \dots, N\}, \end{aligned}$$

This problem has been solved in [76] by invoking the contra-polymatroid feature of the SW coding rate region. We thus directly conclude that  $R_1^* = H(X_1)$ ,  $R_2^* = H(X_2|X_1)$ ,  $R_3^* = H(X_3|X_1, X_2)$ , ...,  $R_N^* = H(X_N|X_1, X_2, \dots, X_{N-1})$  is near-optimal, if and only if  $U_1 \leq U_2 \leq \dots \leq U_N$ , or equivalently  $E_1/d_1^n \geq E_2/d_2^n \geq \dots \geq E_N/d_N^n$ .

□

**Proposition 3:** In joint decoder limited networks, the near-optimal  $T_i^*$  are

$$T_i^* = \frac{E_0}{P_{CR}} \frac{R_i^* U_i}{\sum_{j=1}^N R_j^* U_j}. \quad (7.9)$$

**Proof:** Recall  $T_i^* \approx S^* U_i R_i^* / B$ . Hence,  $T_i / T_j = (R_i U_i) / (R_j U_j)$ . We have

$$\begin{aligned} \sum_{i=1}^N T_i &= \frac{E_0}{P_{CR}} \Rightarrow T_1 \left( 1 + \frac{T_2}{T_1} + \dots + \frac{T_N}{T_1} \right) = \frac{E_0}{P_{CR}}, \\ \Rightarrow T_1 \left( 1 + \frac{R_2 U_2}{R_1 U_1} + \dots + \frac{R_N U_N}{R_1 U_1} \right) &= \frac{E_0}{P_{CR}} \\ \Rightarrow T_1 \frac{1}{R_1 U_1} \sum_{i=1}^N R_i U_i &= \frac{E_0}{P_{CR}} \\ \Rightarrow T_1 &= \frac{E_0}{P_{CR}} \frac{R_1 U_1}{\sum_{i=1}^N R_i U_i}. \end{aligned}$$

Therefore, given near-optimal rates  $R_i^*$ , the near-optimal transmit duration of node 1 is  $T_1^* = E_0 R_1^* U_1 / (P_{CR} \sum_{i=1}^N R_i^* U_i)$ . Repeating this procedure yields (7.9).  $\square$

The following remark highlights the differences between the optimal and near-optimal  $R_i^*$  provided in Propositions 1 and 2 and the literature.

**Remark:** Neglecting energy constraints at the joint decoder and the nodes, the authors in [72][76][74] assigned rates to the nodes according to  $R_1^* = H(X_1)$ ,  $R_2^* = H(X_2|X_1)$ ,  $R_3^* = H(X_3|X_1, X_2)$ , ...,  $R_N^* = H(X_N|X_1, X_2, \dots, X_{N-1})$ , where  $1/d_1^n \geq 1/d_2^n \geq 1/d_3^n \geq \dots \geq 1/d_N^n$ . This solution can lead to early energy drain of nodes that are closer to the joint decoder, since it assigns highest (lowest) rate to the node that is closest to (farthest from) the joint decoder. In contrast, Proposition 2 considers both distance and residual energy for rate assignment and assigns highest (lowest) rate to the node that has the largest (smallest) ratio of residual energy to the  $n$ th power of distance.

The rate assignment in [72][76][74] is also significantly different from the one provided in Proposition 1. The solutions provided in Proposition 1 indicate that there always exists at least one limiting source node whose jointly optimal transmit duration and SW coding rate are the ones that enable the limiting node(s) to send as much data as possible under the SW coding rate region constraints. Meanwhile the remaining nodes have the freedom to choose any feasible transmit durations and SW coding rates. The jointly optimal transmit durations and SW coding rates depend on the residual energy, the distances, the circuit power consumption, and the channel capacity constraints.

## 7.5 The critical energy of the joint decoder

We define the critical energy of the joint decoder as the minimum energy storage at the joint decoder such that constraint  $C_1$  in (P7.2) remains inactive. When  $E_0$  is greater than or equal to (less than) the critical energy,  $C_1$  in (P7.2) is inactive (active) and we can use Proposition 1 (Propositions 2 and 3). In this section we investigate this critical energy.

Suppose the  $R_i^*$  and  $T_i^*$  are solutions provided by Proposition 1. We have

$$\begin{aligned} T_i^* B \log_2 \left( 1 + \frac{P_i^*}{N_0 d_i^n} \right) &= \frac{D_1^\dagger R_i^*}{R_1^*} \\ \Rightarrow T_i^* &= \frac{\frac{D_1^\dagger R_i^*}{BR_1^*}}{\log_2 \left( 1 + \frac{P_i^*}{N_0 d_i^n} \right)} \\ \stackrel{\text{I}}{\Rightarrow} T_i^* &\geq \frac{\frac{D_1^\dagger R_i^*}{BR_1^*}}{\log_2 \left( 1 + \frac{E_i/T_i^* - P_{CT}}{N_0 d_i^n} \right)} \\ \stackrel{\text{II}}{\Rightarrow} T_i^* &\geq f_i(R_i^*) \end{aligned}$$

where  $f_i(R_i^*) = \left( -\mathcal{W} \left( -\frac{a_i \ln 2}{c_i} 2^{-\frac{a_i d_i}{c_i}} \right) / (a_i \ln 2) - d_i/c_i \right)^{-1}$ ,  $a_i = D_1^\dagger R_i^* / (BR_1^*)$ ,  $c_i = E_i / (d_i^n N_0)$ , and  $d_i = 1 - P_{CT} / (d_i^n N_0)$ . Step I results from  $P_i^* \leq E_i/T_i^* - P_{CT}$  and step II is due to the Lambert function being monotonically increasing over  $\mathbb{R}$ . Therefore, if for an optimal set  $R_i^*$  the joint decoder energy storage  $E_0 = P_{CR} \sum_{i=1}^N T_i^*$  is greater than or equal to the critical energy  $P_{CR} \sum_{i=1}^N f_i(R_i^*)$  then  $C_1$  in (P7.2) is inactive. Since the optimal set  $R_i^*$  is not unique (in fact, based on Proposition 1 there is an infinite number of optimal sets), we can formulate the following linear constrained optimization problem to find the critical energy

$$\begin{aligned} \text{(P7.4)} \quad \hat{E}_0 &= \min P_{CR} \sum_{i=1}^N f_i(R_i^*), \\ \text{s.t.} \quad C_a &: R_1^* = \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\}, \\ C_b &: \sum_{i=2}^N R_i^* = H(X_1, X_2, \dots, X_N) - R_1^*, \\ C_c &: R^*(G) \geq H(X(G)|X(G^c)), \\ &\forall G \subseteq \{1, 2, \dots, N\}, G \neq \{1, 2, \dots, N\}, G \neq \emptyset. \end{aligned}$$

Solving (P7.4) gives us the exact value for the critical energy. However, this is rather complicated. A simple suboptimal approach is to minimize an approximation of the lower bound on  $P_{CR} \sum_{i=1}^N f_i(R_i^*)$ . Examining  $f_i(R_i^*)$  we find

$$\begin{aligned} f_i(R_i^*) &= \left( -\frac{\mathcal{W}\left(-\frac{a_i \ln 2}{c_i} 2^{-\frac{a_i d_i}{c_i}}\right)}{a_i \ln 2} - \frac{d_i}{c_i} \right)^{-1} \\ &\approx \left( -\frac{-\frac{a_i \ln 2}{c_i} 2^{-\frac{a_i d_i}{c_i}}}{a_i \ln 2} - \frac{d_i}{c_i} \right)^{-1} \\ &= \frac{c_i}{2^{-\frac{a_i d_i}{c_i}} - d_i} = \frac{E_i}{N_0 d_i^n 2^{V_i R_i^*} + P_{CT} - N_0 d_i^n}, \end{aligned}$$

where  $V_i = D_1^\dagger(P_{CT} - N_0 d_i^n)/(E_i B R_1^*)$ . The approximation above is based on the linear approximation of the Lambert function  $\mathcal{W}(-x) \approx -x$ , when  $x \rightarrow 0^+$ , and it becomes accurate when  $N_0 d_i^n \ll P_{CT}$ . This is equivalent to assuming a high SNR when the circuit and transmit power consumptions are comparable. Consequently, we have

$$\begin{aligned} \sum_{i=1}^N f_i(R_i^*) &\approx \sum_{i=1}^N \frac{E_i}{N_0 d_i^n 2^{V_i R_i^*} + P_{CT} - N_0 d_i^n} \\ &= 2^{\log_2 \left( \sum_{i=1}^N \frac{E_i}{N_0 d_i^n 2^{V_i R_i^*} + P_{CT} - N_0 d_i^n} \right)} \\ &\geq 2^{L(R_i^*)}, \end{aligned}$$

where

$$\begin{aligned} L(R_i^*) &= \sum_{i=1}^N \frac{1}{N} \log_2 E_i + \log_2 N \\ &\quad - \sum_{i=1}^N \frac{1}{N} \log_2 (N_0 d_i^n 2^{V_i R_i^*} + P_{CT} - N_0 d_i^n), \end{aligned} \tag{7.10}$$

and the inequality results from the concavity of the logarithm function and Jensen's inequality. Thus,  $\mathcal{L}(R_i^*) = P_{CR} 2^{L(R_i^*)}$  is a lower bound on the critical energy and minimization of this bound is equivalent to minimizing the third term in (7.10). We can rewrite this term as

$$\begin{aligned} &\frac{1}{N} \sum_{i=1}^N \left[ \log_2 \left( \frac{N_0 d_i^n 2^{V_i R_i^*}}{P_{CT} - N_0 d_i^n} + 1 \right) + \log_2 (P_{CT} - N_0 d_i^n) \right] \\ &\approx \sum_{i=1}^N \frac{N_0 d_i^n 2^{V_i R_i^*}}{N(P_{CT} - N_0 d_i^n)} + \frac{1}{N} \sum_{i=1}^N \log_2 (P_{CT} - N_0 d_i^n) \\ &\approx \sum_{i=1}^N \frac{N_0 d_i^n D_1^\dagger R_i^*}{N E_i B R_1^*} + \sum_{i=1}^N \frac{\log_2 (P_{CT} - N_0 d_i^n)}{N}. \end{aligned} \tag{7.11}$$

The first approximation in (7.11) is based on the linear approximation  $\log_2(1+x) \approx x$ , when  $x \rightarrow 0$ . The second approximation in (7.11) results from the Maclaurin expansion and discarding higher-order components. These approximations become accurate when  $N_0 d_i^n \ll P_{CT}$ . We note that minimizing the critical energy lower bound is equivalent to maximizing  $\sum_{i=1}^N (R_i^* N_0 d_i^n D_1^\dagger) / (N E_i B R_1^*)$ . Therefore, the minimization of the critical energy lower bound can be approximated by

$$\begin{aligned}
(\mathbf{P7.5}) \quad & \max \sum_{i=1}^N \frac{N_0 d_i^n D_1^\dagger}{N E_i B R_1^*} R_i^*, \\
\text{s.t.} \quad & C'_a : R_1^* = \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\}, \\
& C'_b : \sum_{i=2}^N R_i^* = H(X_1, X_2, \dots, X_N) - R_1^*, \\
& C'_c : R^*(G) \geq H(X(G)|X(G^c)), \\
& \forall G \subseteq \{1, 2, \dots, N\}, G \neq \{1, 2, \dots, N\}, G \neq \emptyset.
\end{aligned}$$

Let  $\tilde{R}_i^*$  denote the SW coding rate allocation obtained by solving (P7.5). Using similar techniques we used to prove Proposition 2, we find  $\tilde{R}_1^* = \max_G \left\{ \frac{H(X(G)|X(G^c))D_1^\dagger}{D^\dagger(G)} \right\}$ ,  $\tilde{R}_2^* = H(X_1, X_2|X_3, \dots, X_N) - \tilde{R}_1^*$ ,  $\tilde{R}_3^* = H(X_1, X_2, X_3|X_4, \dots, X_N) - \tilde{R}_1^* - \tilde{R}_2^*$ , ..., and  $\tilde{R}_N^* = H(X_1, X_2, \dots, X_N) - \sum_{i=1}^{N-1} \tilde{R}_i^*$ , where we have  $d_2^n/E_2 \geq d_3^n/E_3 \geq \dots \geq d_N^n/E_N$ . Thus, the critical energy must satisfy

$$P_{CR} \sum_{i=1}^N f_i(\tilde{R}_i^*) \geq \hat{E}_0 \geq \mathcal{L}(\tilde{R}_i^*).$$

When the joint decoder's residual energy satisfies  $E_0 \geq P_{CR} \sum_{i=1}^N f_i(\tilde{R}_i^*)$ , the constraint  $C_2$  in (P7.2) is inactive. When  $E_0 \leq \mathcal{L}(\tilde{R}_i^*)$  this constraint is active. When  $P_{CR} \sum_{i=1}^N f_i(\tilde{R}_i^*) \geq E_0 \geq \mathcal{L}(\tilde{R}_i^*)$ , we cannot determine whether or not the network is joint decoder limited. In this case, we assume constraint  $C_2$  is active and find a suboptimal solution using Propositions 2 and 3.

## 7.6 Impact of SW Encoding and Decoding Energy Costs

So far we have assumed that the SW encoding and decoding energy costs are negligible. In the following, we investigate how we can incorporate these energy costs into (P7.2)

and how the results of Propositions 1, 2, and 3 can be used to find the optimal resource allocation when these costs are incorporated. We modify constraints  $C_1$  and  $C_2$  in (P7.2) as follows

$$\begin{aligned} C_1 : P_{CR} \sum_{i=1}^N T_i &\leq E_0 - E_{DE}SN, \\ C_2 : T_i [(2^{\frac{SR_i}{BT_i}} - 1)d_i^n N_0 + P_{CT}] &\leq E_i - E_{EN,i}S, \end{aligned} \quad (7.12)$$

where  $E_{DE}$  and  $E_{EN,i}$ , measured both in Joules/sample, denote the energy cost of SW decoding at the joint decoder and encoding at node  $i$ , respectively.

Let (P7.6) denote the new optimization problem with modified constraints  $C_1$  and  $C_2$  given in (7.12). Also, let  $S_E^*$ ,  $R_{E,i}^*$  and  $T_{E,i}^*$ , respectively, denote the optimal parameters obtained from solving (P7.6). Unfortunately, solving (P7.6) in closed-form is intractable. However, in the following we describe an iterative numerical method to solve (P7.6).

Let  $S_m^*$ ,  $R_{i,m}^*$  and  $T_{i,m}^*$ , represent the solution of (P7.2) and  $\mathcal{E}_{i,m}$  and  $\mathcal{E}_m$ , respectively, indicate the residual energy of sensor  $i$  and the joint decoder, in the  $m$ th iteration. We set  $\mathcal{E}_{i,0} = E_i$  and  $\mathcal{E}_0 = E_0$ . In the  $m$ th iteration, we first update the residual energy as  $\mathcal{E}_{i,m} = \mathcal{E}_{i,m-1} - E_{EN}S_{m-1}^*$  and  $\mathcal{E}_m = \mathcal{E}_{m-1} - NE_{DE}S_{m-1}^*$ , then solve (P7.2) using these updated residual energies and obtain  $S_m^*$ ,  $R_{i,m}^*$  and  $T_{i,m}^*$ . Clearly,  $S_0^*$  is an upper bound of  $S_E^*$ . It is easy to show that in general the resulting  $S_m^*$  is either a new upper bound or a new lower bound of  $S_E^*$ , depending on whether  $m$  is even or odd. As  $m$  increases, these upper and lower bounds approach to  $S_E^*$ . The convergence of the algorithm will be demonstrated in Section 7.8 through simulations. The iterations will be stopped when the average decreases of the lower and upper bounds  $0.5 (|S_m^* - S_{m-2}^*| + |S_{m-1}^* - S_{m-3}^*|) / S_m^*$  is less than a small value  $\epsilon$ . Note that at least 4 iterations are required to use this stopping criterion. At each iteration as we update the residual energies, we need to apply the results in Section 7.5 and re-evaluate whether or not the network is joint decoder limited and hence use the appropriate proposition to find the optimal or near-optimal set for that specific iteration. The iterative procedure to solve (P7.6) is summarized in the pseudocode shown in Table. 7.1.

Table 7.1: The proposed optimal rate allocation with coding energy cost

---

**Algorithm 7.6.1:** OPTIMAL RATE ALLOCATION ( $E_{\text{EN},i}, E_{\text{DE}}$ )

---

Initialize  $\mathcal{E}_{i,0} = E_i, \mathcal{E}_0 = E_0$

calculate  $\{S_0^*, R_{i,0}^*, T_{i,0}^*\}$  based on  $\mathcal{E}_{i,0}, \mathcal{E}_0$

$m \leftarrow 1$ ; **while**  $m \leq 4$

**do**  $\left\{ \begin{array}{l} m \leftarrow m + 1 \\ \text{update residual energy } \mathcal{E}_{i,m} = \mathcal{E}_{i,m-1} - S_{m-1}^* E_{\text{EN},i}, \\ \mathcal{E}_m = \mathcal{E}_{m-1} - N S_{m-1}^* E_{\text{DE}} \\ \text{calculate}\{S_m^*, R_{i,m}^*, T_{i,m}^*\} \text{ based on } \mathcal{E}_{i,m}, \mathcal{E}_m \end{array} \right.$

$m \leftarrow 4; \epsilon \leftarrow$  a small positive number

**while**  $0.5 (|S_m^* - S_{m-2}^*| + |S_{m-1}^* - S_{m-3}^*|) / S_m^* \geq \epsilon$

**do**  $\left\{ \begin{array}{l} \text{update residual energy } \mathcal{E}_{i,m} = \mathcal{E}_{i,m-1} - S_{m-1}^* E_{\text{EN},i}, \\ \mathcal{E}_m = \mathcal{E}_{m-1} - N S_{m-1}^* E_{\text{DE}} \\ \text{calculate}\{S_m^*, R_{i,m}^*, T_{i,m}^*\} \text{ based on } \mathcal{E}_{i,m}, \mathcal{E}_m \\ m \leftarrow m + 1 \end{array} \right.$

let  $S_{\text{E}}^* = S_m^*, R_{\text{E},i}^* = R_{i,m}^*, T_{\text{E},i}^* = T_{i,m}^*$

---

## 7.7 Extensions to Multi-hop Scenarios

So far we have assumed that the sensors are a single hop away from the joint decoder. This is a reasonable assumption assuming that the network is clustered, and the cluster

heads are the joint decoders. Many algorithms for clustering and cluster head selection have been proposed and well studied in the literature [112][125]-[127].

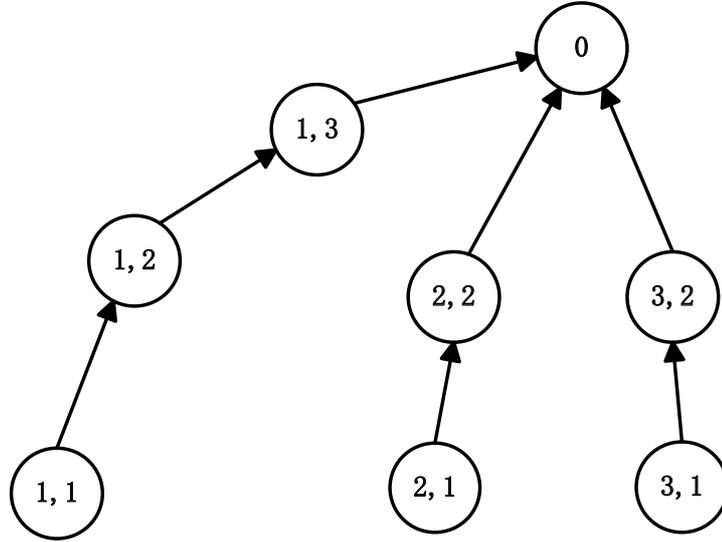


Figure 7.2: Application of the proposed algorithms in a data-gathering wireless network with multi-hop routes.

In the case that the network is not clustered, or that the cluster heads cannot act as joint decoders, it is desirable to perform the optimization in a multi-hop setting. The proposed methodology in the single hop scenario can be extended to a specific multi-hop scenario, where all of the multi-hop routes are distinct (non-overlapping). An example of such a multi-hop network is shown in Fig. 7.2, where three source nodes send data through a 3-hop, a 2-hop, and a 2-hop route to the sink, respectively.

Assume that we have  $N$  sources and correspondingly  $N$  distinct routes. The source on the  $i$ th route is labeled as  $(i, 1)$ . Furthermore, assume that there are  $K_i$  relay nodes on each route, and the relay nodes are denoted by  $(i, k)$ , where  $k = 2, \dots, K_i + 1$ . Let  $d_{i,k}$  be the distance of node  $(i, k)$  to the next node on the route,  $(i, k + 1)$ , and  $E_{i,k}$  be the energy of node  $(i, k)$ . When there is no joint decoder energy constraint, the energy

constraints of the source and relay nodes are

$$\begin{aligned} T_{i,1} \left[ (2^{\frac{SR_i}{BT_{i,1}}} - 1) d_{i,1}^n N_0 + P_{CT} \right] &\leq E_{i,1}, \\ T_{i,k-1} P_{CR} + T_{i,k} \left[ (2^{\frac{SR_i}{BT_{i,k}}} - 1) d_{i,k}^n N_0 + P_{CT} \right] &\leq E_{i,k}, \\ \forall k = 2, \dots, K_i + 1, i = 1, 2, \dots, N, \end{aligned}$$

where  $T_{i,k}$  is the transmit duration of node  $(i, k)$ . Note that the above energy constraints correspond to constraints  $C_2$  of (P7.2) in the single-hop scenario. The SW coding region constraints stay the same in the multi-hop scenario. Equivalently, the above constraints can be written as

$$\begin{aligned} SR_i &\leq BT_{i,1} \log_2 \left( \frac{E_{i,1}/T_{i,1} - P_{CT}}{d_{i,1}^n N_0} + 1 \right), \\ SR_i &\leq BT_{i,k} \log_2 \left( \frac{(E_{i,k} - T_{i,k-1} P_{CR})/T_{i,k} - P_{CT}}{d_{i,k}^n N_0} + 1 \right), \\ \forall k = 2, \dots, K_i + 1. \end{aligned}$$

Therefore, maximizing the number of samples is equivalent to finding the optimal transmit durations  $T_{i,k}$  that maximize the tightest constraint in the above constraint set. That is  $S^* R_i \leq D_i^\dagger$ , where

$$\begin{aligned} D_i^\dagger &= \max_{T_{i,k}} \min \left\{ BT_{i,1} \log_2 \left( \frac{E_{i,1}/T_{i,1} - P_{CT}}{d_{i,1}^n N_0} + 1 \right), \right. \\ &\quad \left. BT_{i,k} \log_2 \left( \frac{(E_{i,k} - T_{i,k-1} P_{CR})/T_{i,k} - P_{CT}}{d_{i,k}^n N_0} + 1 \right) \right\}. \end{aligned} \quad (7.13)$$

Assume that the solutions to the above max-min problem are  $T_{i,k}^\dagger$ . Based on Proposition 1 and Lemma 1, without loss of generality, we assume  $D_1^* = D_1^\dagger$ . Correspondingly, the optimal transmit durations for route one is  $T_{1,k}^* = T_{1,k}^\dagger, \forall k = 1, \dots, K_1 + 1$ , and the optimal SW coding rate of node (1,1) is the same as in the single hop case, i.e.,  $R_1^* = \max_G \left\{ H(X(G)|X(G^c)) D_1^\dagger / D^\dagger(G) \right\}$ . The maximum number of samples is  $S^* = D_1^\dagger / R_1^*$ .  $R_j^*$  for  $j \geq 2$ , can be any feasible solution that satisfies  $\sum_{j=2}^N R_j^* = H(X_1, \dots, X_N) - R_1^*$ . The optimal transmit durations for route  $j$  are the ones that enable route  $j$  to deliver  $S^* R_j^*$  amount of data to the joint decoder. Unfortunately, unlike the single-hop case, there are no analytical solutions for the optimal  $T_{i,k}$ .

In the presence of an energy constraint at the joint decoder, a similar linearization method as in Section 7.4 can be applied to the last hop to the decoder, i.e.,  $k = K_i + 1$ . That is,

$$\begin{aligned} SR_i &= BT_{i,K_i+1} \log_2 \left( \frac{(E_{i,K_i+1} - T_{i,K_i} P_{CR})/T_{i,K_i+1} - P_{CT}}{d_{i,K_i+1}^n N_0} + 1 \right) \\ &\approx BT_{i,K_i+1} \log_2 \frac{(E_{i,K_i+1} - T_{i,K_i} P_{CR})/T_{i,K_i+1} - P_{CT}}{d_{i,K_i+1}^n N_0} \\ &\approx B \left[ \log_2 \left( \frac{E_{i,K_i+1}}{d_{i,K_i+1}^n N_0} \right) + \frac{6}{\ln 2} \right] T_{i,K_i+1}, \end{aligned}$$

where we assume  $[(E_{i,K_i+1} - T_{i,K_i} P_{CR})/T_{i,K_i+1} - P_{CT}]/(d_{i,K_i+1}^n N_0) \gg 1$ , i.e. a high SNR, and  $E_{i,K_i} \gg T_{i,K_i-1} P_{CR} + T_{i,K_i} P_{CT}$ . Thus, we have a similar problem that provides the near-optimal  $R_i^*$  and  $T_{i,K_i}^*$  by using Propositions 2 and 3. The remaining near-optimal transmit durations  $T_{i,k}^*$ ,  $\forall k \neq K_i$  can be any feasible solutions that satisfy the constraints.

Slightly more complicated multi-hop topologies may be treated as above with some further detail. However, not much can be said for an arbitrary multi-hop topology.

## 7.8 Numerical Results

In this section, we numerically verify the analytical optimal and near-optimal solutions provided in Propositions 1, 2, and 3, and we quantify the increase in the number of collected samples due to our optimization. We consider two cases with  $N = 2$  and 10 source nodes and assume that the nodes are uniformly placed within a disk with a radius of 200 m, centered at the joint decoder. In our simulations, the path loss exponent is  $n = 3.5$ , the bandwidth is  $B = 50$  kHz, the circuit power consumption at the transmitter and receiver, respectively, are  $P_{CT} = 84.8$  mW and  $P_{CR} = 107.8$  mW, and the power spectral density of noise is  $-174$  dBm/Hz [112].

### 7.8.1 Network with $N = 2$ source nodes

We compare the average of the maximum number of collected samples  $S^*$  as the spatial correlation varies for six different approaches. The average is calculated over 1000 random network deployments. The six approaches are: (i) Each node codes its information with a rate equal to its unconditioned entropy  $R_i = H(X_i)$  and uses optimized  $T_i$ . We refer to this approach as “No DSC”. (ii) The SW rate allocation among nodes is specified by a corner of the SW coding rate region and nodes use optimized  $T_i$ . We refer to this approach as “Corner  $R_i$ ”. (iii) Nodes use equal SW coding rates of the joint entropy divided by  $N$  and optimized  $T_i$ , which we refer to as “Equal  $R_i$ ”. (iv) Nodes determine their transmit powers according to a target SNR of 20dB and numerically optimize the SW coding rates, which we refer to as “Fixed  $P_i$ ”. (v) The SW coding rates and transmit durations are obtained from Proposition 1 (when  $C_2$  in (P7.2) is inactive) or Propositions 2 and 3 (when it is active), which we refer to as “Analytical Results”, and (vi) The SW coding rates and transmit durations are found via numerically solving (P7.2) without any approximations or simplifications, which we refer to as “Numerical Optimization.”

We let  $H(X_1) = H(X_2) = 1$  and  $H(X_1|X_2) = H(X_2|X_1) = h$ ,  $0 \leq h \leq 1$ . Note that a larger  $h$  indicates a lower spatial correlation between  $X_1$  and  $X_2$ , and vice versa. For the “Corner  $R_i$ ” approach we choose the corner  $R_1 = H(X_1|X_2)$ ,  $R_2 = H(X_2)$ . For the case when constraint  $C_2$  in (P7.2) is inactive we assume the joint decoder and the sensors’s residual energies are  $E_0 = 11880$  J (approximately the energy capacity of a AA battery [128]) and  $E_1 = E_2 = 118.8$  J, (selected to be 1% of  $E_0$ ), respectively. For the case when constraint  $C_2$  in (P7.2) is active we assume  $E_0 = 118.8$  J and  $E_1 = E_2 = 11880$  J.

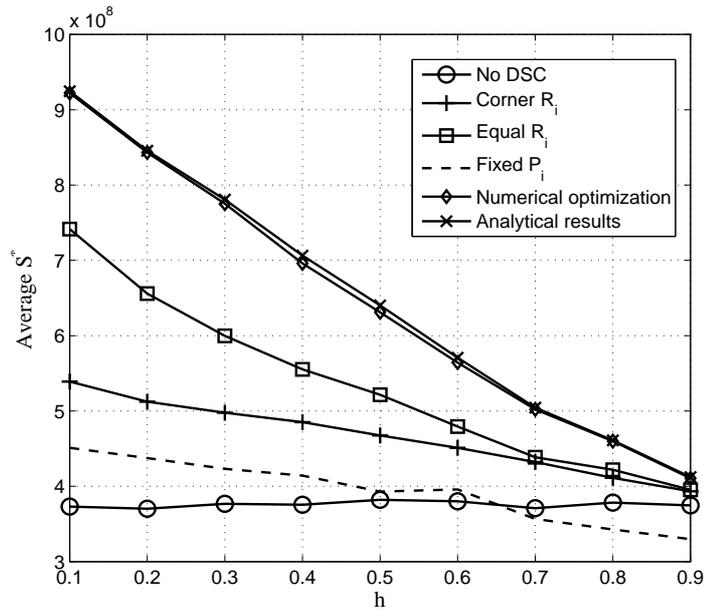


Figure 7.3: Performance comparison when the network is not joint decoder limited: average  $S^*$  versus  $h$  for  $N = 2$ .

#### 7.8.1.1 Constraint $C_2$ in (P7.2) is inactive

Fig. 7.3 shows that the “Analytical Results” and “Numerical Optimization” approaches have similar performance. Also, the “Analytical Results” approach offers a significant performance gain over the “Fixed  $P_i$ ”, “No DSC”, “Corner  $R_i$ ”, and “Equal  $R_i$ ” approaches. For instance, for  $h = 0.3$ , the average of  $S^*$  in the “Analytical Results” approach is 2.1 times that of the “No DSC” approach, 1.8 times that of the “Fixed  $P_i$ ” approach, 1.6 times that of the “Corner  $R_i$ ” approach, and 1.3 times that of the “Equal  $R_i$ ” approach. The relative importance of the optimization over transmit durations and the SW coding rates differs according to the conditional entropy. This is evident by comparing the “Fixed  $P_i$ ” and the “No DSC” approaches. The “Fixed  $P_i$ ” approach optimizes only over the SW coding rates, while the “No DSC” approach optimizes only over the transmit durations. When  $h$  is small, the “Fixed  $P_i$ ” approach outperforms the “No DSC” approach, since the optimization over the SW coding rates is relatively more important than the optimization over the transmit durations. When  $h$  is large, the “No

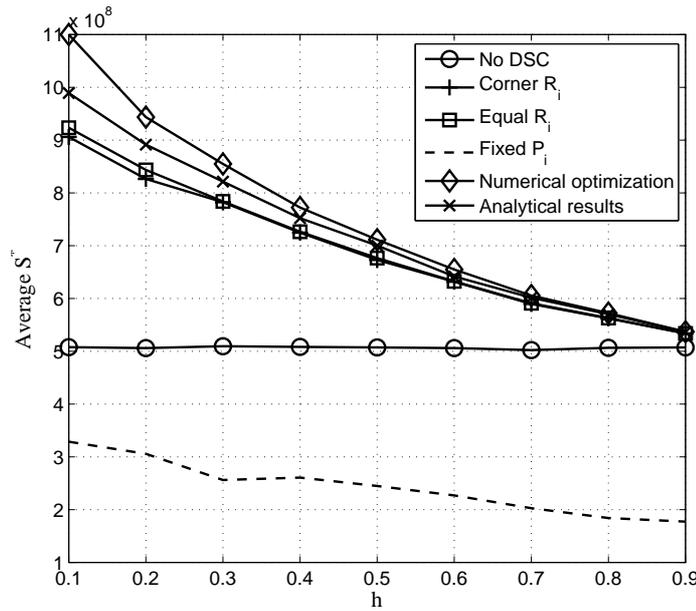


Figure 7.4: Performance comparison when the network is joint decoder limited: average  $S^*$  versus  $h$  for  $N = 2$ .

DSC” approach outperforms the “Fixed  $P_i$ ”, approach since in this case the optimization over the transmit durations is the dominating factor due to the SW coding being less effective in absence of high spatial correlation between sources.

### 7.8.1.2 Constraint $C_2$ in (P7.2) is active

Fig. 7.4 shows that the performance of the “Analytical Results” and the “Numerical Optimization” approaches are close, although the computational complexity of the former approach is almost negligible, compared with that of the latter. Comparing Figs. 7.3 and 7.4, we observe that joint optimization in networks with inactive joint decoder energy constraints is more important than that of networks with active joint decoder energy constraints. Considering Fig. 7.3, we observe that when  $h \approx 0.1$ , the “Analytical Results” approach provides a gain as large as 171% over the “Corner  $R_i$ ” and the “Equal  $R_i$ ” approaches, while considering Fig. 7.4, we see that this gain is only about 22%. This is because when constraint  $C_2$  is inactive, the total transmission duration  $T_1 + T_2$

is determined by  $E_1$  and  $E_2$ , and a non-optimal SW coding rate allocation accelerates the drain of energy at the nodes and reduces  $T_1 + T_2$ . However, when constraint  $C_2$  is active,  $T_1 + T_2$  is mainly determined by  $E_0$  and it is less affected by the adopted SW coding rate allocation (compared with the case where  $C_2$  is inactive). Thus, when the nodes' residual energy is limited and the joint decoder energy storage is abundant, SW coding rate optimization becomes more crucial. This is also evident by the fact that the performance of the "No DSC" approach is always better than that of the "Fixed  $P_i$ " approach. We also note that the performance gap between different approaches, excluding the "Fixed  $P_i$ " approach, diminishes as  $h$  increases since SW coding becomes less effective.

### 7.8.2 Network with $N = 10$ source nodes

Our joint entropy model for the 10-source node case is based on the entropy approximation proposed in [79] where the  $i$ th source contributes an amount of uncorrelated data equal to  $\frac{r_i}{c+r_i}H(X_1)$ , where  $r_i$  is the minimum distance of the  $i$ th source to the  $1, \dots, i-1$  sources and  $c$  is a constant representing the extent of the spatial correlation of the data and varies with respect to the data of interest. Similar to [79] we let  $c = 25$  and  $r_j < r_i, \forall j > i$ . For the sake of simplicity, a complete symmetry of the entropy is assumed.

We assume the joint decoder and sensors' residual energies are uniformly distributed within  $(0, E_{max}]$ , with  $E_{max} = 1000\text{J}, 2000\text{J},$  and  $3000\text{J}$ . For each  $E_{max}$ , 1000 trials are simulated. We let  $r_i$  be 10, 20, ..., 90, and  $H(X_i) = 1, \forall i$ . For the "Corner  $R_i$ " approach we randomly choose one corner of the SW coding region. In addition to the previously described six approaches, we include two more approaches: the "Proposition 1" approach and "Propositions 2 and 3" approach. In the "Proposition 1" approach, we assume the network is never joint-decoder-limited and only use Proposition 1 to determine the optimal transmit durations and the SW coding rates, while in the "Propositions 2 and 3" approach we assume the network is always joint-decoder-limited and

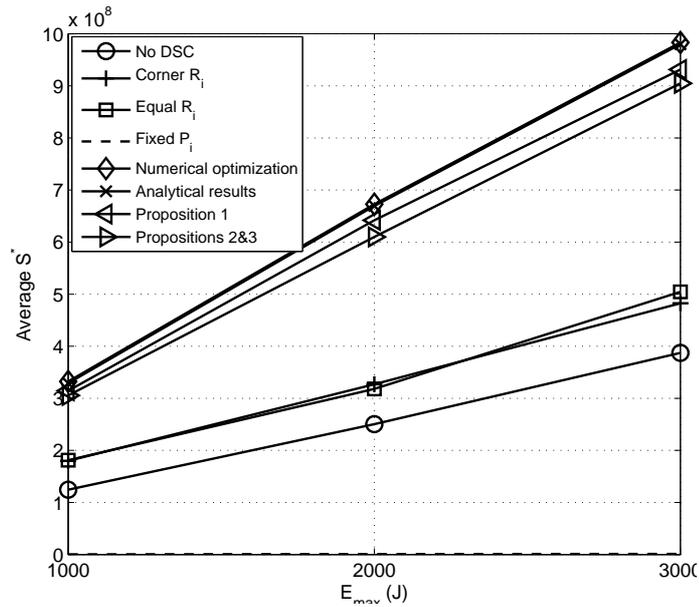


Figure 7.5: Performance comparison: average  $S^*$  versus  $E_{max}$  for  $N = 10$ .

only use Propositions 2 and 3 to determine the optimal transmit durations and the SW coding rates. These two additional approaches illustrate the importance of the choice of appropriate solutions according to the calculated critical energy. In the “Analytical Results” approach, for each realization, we apply our results in Section 7.5 to find the critical energy and hence determine whether constraint  $C_2$  is inactive or active.

Fig. 7.5 shows that the performance of the “Analytical Results” approach (provided by either Proposition 1 or Propositions 2 and 3, depending on whether  $C_2$  is inactive or active) and the “Numerical Optimization” approaches are very close, although the computational complexity of the former approach is almost negligible, compared with that of the latter approach. For instance, for  $E_{max} = 3000$  J the average  $S^*$  for the “Analytical Results” and the “Numerical Optimization” approaches has only 1% difference. Furthermore, the “Analytical Results” approach offers a significant performance gain over the “No DSC”, “Corner  $R_i$ ”, “Equal  $R_i$ ”, and “Fixed  $P_i$ ” approaches. For instance, for  $E_{max} = 3000$  J the average of  $S^*$  in the “Analytical Results” approach is 2.5 times that of the “No DSC” approach, 2.0 times that of the “Corner  $R_i$ ” approach, 1.9

times that of the “Equal  $R_i$ ” approach, and 777.8 times that of the “Fixed  $P_i$ ” approach.

In our calculations, the percentage of using Proposition 1, i.e.,  $C_2$  is inactive according to the calculated critical energy, is approximately 70% (2093 out of 3000 trials). The percentage of using Proposition 2 and 3, i.e.,  $C_2$  is active according to the calculated critical energy, is about 30% (907 out of 3000 trials). Thus, as shown in Fig. 7.5, the performance of the “Analytical Results” approach which is based on critical energy, is better than either “Proposition 1” or “Propositions 2 and 3” approaches, and the “Proposition 1” approach outperforms the “Propositions 2 and 3” approach, since oftentimes constraint  $C_2$  is inactive. The results show the importance of choosing the appropriate proposition(s) based on the critical energy calculation.

### 7.8.3 Impact of SW Encoding and Decoding Energy Costs

To demonstrate the effectiveness of the proposed algorithm in Section 7.6, we apply it to a 10-source node network with residual energies 1000J for all nodes and transmission distances  $d_i = 100\text{m}$  for all sources. The remaining parameters remain unchanged.

It has been shown in the literature that low-density parity-check (LDPC) coding is a promising implementation technique for DSC [129][130]. Thus, we assume the energy consumption of the DSC decoder is the same as that of a state-of-the-art LDPC decoder [131]. In [131], the author states that the normalized energy consumption efficiency of the LDPC decoder implementation is 243 pJ/bit/iteration, and the number of iterations is 10. Therefore, the normalized energy consumption of the decoder is 2.43 nJ/bit. Since there are  $H(X_1, \dots, X_N)$  bits/sample to be decoded, we let the DSC decoder energy consumption be  $2.43H(X_1, \dots, X_N)$  nJ/sample. On the other hand, the DSC encoder is similar to the LDPC encoder, which consists of simple matrix multiplications and is much simpler than the LDPC decoder. We assume that the energy consumption of the encoder is 5% of that of the decoder, i.e.  $0.12H(X_1, X_2, \dots, X_N)$  nJ/sample.

Fig. 7.6 shows  $S_E^*$ ,  $S_m^*$  and  $S^*$ . We observe that the addition of SW encoding and decoding energy costs decreases the total collected samples by about 3%. The impact

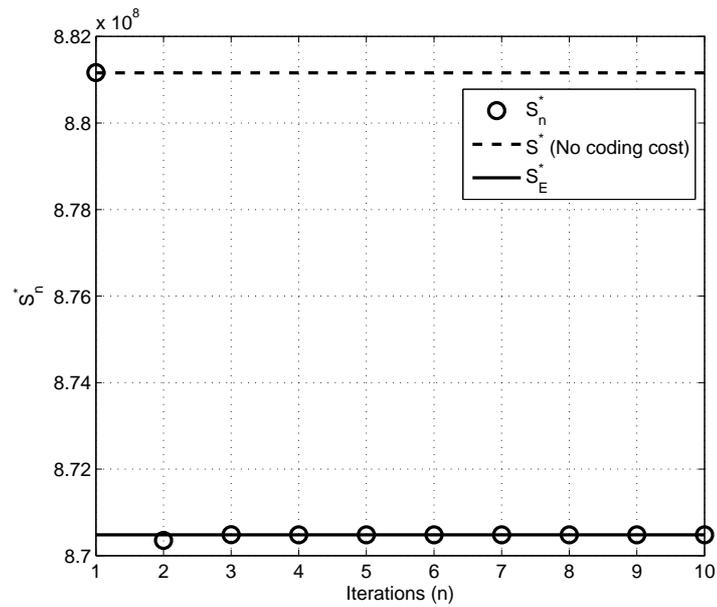


Figure 7.6: Performance of the proposed iterative algorithm:  $S_n^*$  approaches  $S_E^*$  as  $n$  increases for  $N = 10$ .

of encoding and decoding energy consumption is thus not very significant. We also note that the proposed algorithm converges in only 4 iterations with  $\epsilon = 10^{-3}$ . Of course, the number of iterations required for the the algorithm to converge depends on the choice of  $\epsilon$ . Table 7.2 illustrates the average number of iterations required for convergence in terms of  $\epsilon$ . These are the results of averaging over  $10^4$  random network deployments and random initial residual energies. Our simulation results indicate that the proposed algorithm converges in every trial, and the required number of iterations for convergence decreases as  $\epsilon$  increases.

## 7.9 Conclusions and Discussion

In this chapter, we considered a cluster in an energy constrained WSN, where each source node utilizes SW source coding to transmit its encoded information bits to a joint decoder. To maximize the collected samples during the network lifetime, we formulated

Table 7.2: The average number of iterations  $\bar{n}$  w.r.t.  $\epsilon$ 

$\epsilon$	$\bar{n}$
$10^{-10}$	8.7
$10^{-9}$	8.1
$10^{-8}$	7.7
$10^{-7}$	7.0
$10^{-6}$	6.7
$10^{-5}$	6.0
$10^{-4}$	5.7
$10^{-3}$	5.0
$10^{-2}$	4.7
$10^{-1}$	4.0

a detailed optimization problem, that takes into account circuit power consumption, energy constraints at the joint decoder and the nodes, the SW coding rate region, and the capacity constraints imposed by different communication channel models. The optimization parameters are the SW coding rates, the transmit powers, and the transmit durations.

For the case where the joint decoder is not energy constrained, we found low complexity optimal closed-form solutions. For the case where the joint decoder is energy constrained, we approximated the original nonlinear optimization problem with an inequality form linear program, and we found near-optimal closed-form solutions. Our proposed optimal and near-optimal rate assignments are fundamentally different from the ones in literature, which assume the communication cost per information bit depends only on distances. We also provided a simple criterion for determining whether

or not the joint decoder is energy constrained. Our numerical results validated our analytical solutions and also demonstrated the effectiveness of the proposed optimization to increase the number of collected samples, especially when the joint decoder is not energy constrained. Lastly, we proposed an iterative algorithm that employs our analytical results to solve the problem when the SW encoding and decoding energy costs are also taken into account. Numerical results indicate that this algorithm has a fast convergence rate.

In this work we have considered a single, pre-determined, cluster. We note that the clustering algorithm can have a significant impact on the performance of the algorithm. Consequently, further gains are potentially possible using cross layer optimization that includes our approach described in this work and the clustering algorithm.

## **8 Optimal Rate Allocation for Distributed Source Coding over Gaussian Multiple Access Channels**

In this chapter, we continue the work of Chapter 7 by studying the problem of joint optimization of Slepian-Wolf (SW) source coding and transmission rates over a Gaussian multiple access channel with the considerations of circuit power consumption and average transmit power constraint. The goal is to maximize the sample rate at the source nodes. We first derive a criterion to determine the optimality of different multiple access schemes such that the highest sample rate can be achieved at the source nodes when SW coding is used. Based on the derived optimality criterion, we propose a rate allocation procedure to determine the jointly optimal SW coding and transmission rates corresponding to code division multiple access (CDMA), frequency division multiple access (FDMA) and time division multiple access (TDMA) schemes. Several demonstrative numerical examples are provided to show the performance gain of the proposed joint rate allocation scheme.

## 8.1 Introduction

Studies suggest that most of the observations made in modern sensing systems are spatially correlated due to the fact that the sensing devices are densely deployed [132][133]. An efficient exploitation of this spatial correlation is needed to meet the stringent constraints on these systems, in terms of bandwidth usage and energy consumption. The exploitation of spatial correlation has been investigated in the research community for the past decades and different approaches have been proposed, including: distributed source coding (DSC) [64][134], data-correlation-aware routing [135], minimum-energy data gathering [136], and cross-layer optimization [137][138]. Among these techniques, DSC is particularly attractive due to its unique features. DSC enables compression of correlated sources without requiring any communication between the sources, given the joint distribution of the sources at the encoders and the decoder [62]. Moreover, DSC can be seamlessly incorporated with many other techniques to further exploit the spatial correlation of the sources, such as the aforementioned data-correlation-aware routing [136] and cross-layer optimization [137][138].

Although the theoretical results of DSC do not require coordination between the distributed encoders at the sources and the joint decoder at the destination, the information delivery between different sensing devices and the joint decoder does require some degree of coordination [36]. For example, in clustered wireless sensor networks (WSNs) where a cluster consisting of several sensors and a joint decoder (cluster head) has already been formed, the data sensed by sensors should be sent to the cluster head through multiple direct links. Therefore, the multiple access channel problem arises naturally in a typical distributed sensing system since the communication of the sensing devices must be coordinated to prevent possible data loss.

The performance of orthogonal multiple access (MA) schemes, such as code division multiple access (CDMA), time division multiple access (TDMA) and frequency division multiple access (FDMA), have been extensively studied within the context of

multiple access channel problems, where an average power constraint is imposed at the transmitting nodes, albeit without the considerations of circuit power consumption and SW coding [63][139]. A recent study shows that, considering the circuit power consumption can significantly alter the rate region of a MA scheme [118]. This suggests that the performance analysis and evaluation of different MA schemes needs to be carefully revised when circuit power consumption is not negligible.

The impact of the circuit power consumption on the rate regions of different MA schemes is illustrated in Fig. 8.1, which shows the rate regions of CDMA, FDMA and TDMA over a Gaussian multiple access channel with two transmitters. When circuit power consumption is not considered (Fig. 8.1(a)), it has been shown that CDMA is optimal in terms of providing the highest sum rate, while TDMA and FDMA are equivalent [63]. When the circuit power consumption is considered, however, the rate region of TDMA is no longer contained in that of CDMA [118].

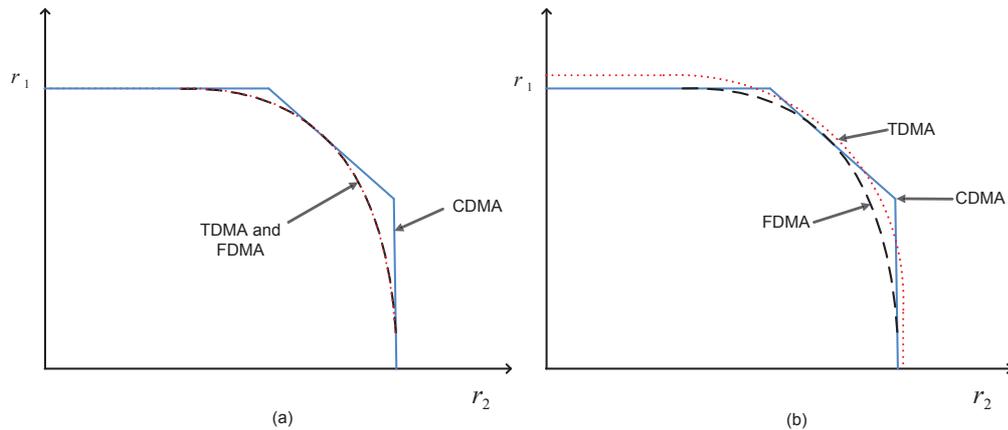


Figure 8.1: The rate regions of TDMA, CDMA, and FDMA for  $N = 2$  without (a) and with (b) the consideration of circuit power consumption.

The performance analysis of different MA schemes is critical in distributed sensing systems, since the transmission rates provided by a particular MA scheme directly determine the sample rate that can be supported at the sensing device. The optimization of MA schemes to improve the energy efficiency of sensing systems has been studied in

the literature [140]-[142]. Most of these works do not take DSC into the consideration. In the context of DSC, and in particular, Slepian-Wolf (SW) coding, the performance of the MA technique needs to be carefully evaluated, since the adoption of SW coding imposes additional constraints on the transmission rates. That is, we need to consider not only the rate regions of the MA schemes but also the SW coding rate region [62].

The source-channel separation theorem has been shown to be valid for the transmission of correlated sources over a Gaussian multiple access channel under certain conditions [143][144], e.g., side information is available at the decoder. However, this theorem does not hold for general multiple access channels with correlated sources. The optimization problem of multiple access channels with correlated sources has been extensively studied from the joint source-channel coding perspective since the pioneer work in [145]. Over the past decade, substantial research efforts have been made in the information theoretical study of multiple access channel rate regions with correlated sources and joint source-channel coding [146][147]. Although joint source-channel coding has been proven to be capacity achieving, a joint source-channel coding scheme is still not immediately available for real world implementations. On the other hand, cross-layer design has proven to possess practical importance in resource allocation problems in WSNs due to its merits, such as simplicity and scalability [148][149].

Maximizing the sample rate with an average power constraint is of significant practical importance. Moreover, in WSNs the circuit power consumption of the sensing devices must be taken into account as the transmit power is usually very low and comparable to the circuit power consumption. Thus the maximization of the sample rate results from an understanding of the requirements imposed by a practical sensing system.

In [150], we provided some preliminary results on maximizing the sample rate for SW coding of two correlated source nodes over Gaussian multiple access channels. We showed that the optimal MA scheme maximizing the sample rate is the one that provides the highest sum rate within a certain region determined solely by the source

entropies. Moreover, we proved that either CDMA or TDMA may be the optimal MA scheme under certain conditions governed by the source entropies, average power and circuit power consumption constraints, noise power, and transmission distances. Although the work in [150] laid the foundation for our subsequent research, it provides neither a detailed criterion to test the optimality of CDMA and TDMA nor a concrete algorithm to generate the jointly optimal rate allocation for CDMA and TDMA.

In this work, we study the problem of joint optimization of the SW coding rates and the MA transmission rates to maximize the sample rates at the sensing devices, when circuit power consumption and an average power constraint are considered. First, we provide a criterion to determine the optimality of CDMA, FDMA and TDMA, and also find the optimal transmission rate allocation for the MA scheme with an arbitrary number of sources. We will show that FDMA cannot be the optimal MA scheme and neither CDMA nor TDMA can guarantee optimality. Based on the optimal transmission rates of the MA scheme, we also find the corresponding optimal SW coding rates at the sources. In our work, we adopt the cross-layer optimization methodology. That is, the channel coding and source coding are implemented separately, while optimized jointly. Our proposed algorithm is easy to implement and provides the jointly optimal allocation of MA transmission rates and SW coding rates.

## 8.2 Problem Formulation

Consider a group of  $N$  source nodes that gather information about the environment of interest and send their information to the destination (node 0) through a multiple access channel (Fig. 8.2). We denote the common sample rate of the nodes by  $S$  (samples/second). The observed samples at node  $i$ , denoted by  $X_i$ , are discrete-valued random variables. At the nodes, SW coding is used, and the corresponding source code rate at node  $i$  is represented by  $R_i$  (bits/sample). After source coding, the nodes apply channel coding to the resulting bit streams and send the coded bits to the destination.

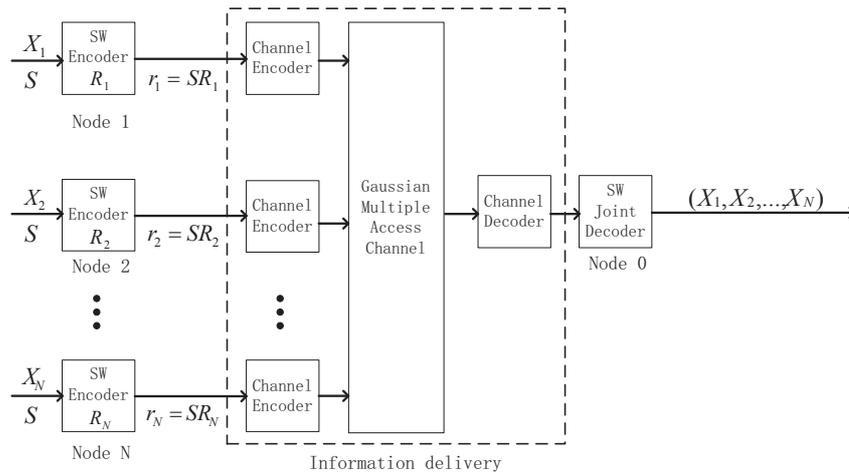


Figure 8.2: An illustration of a WSN with SW coding and a Gaussian multiple access channel.

The transmission rate of the link between node  $i$  and the destination is denoted by  $r_i$  (bits/second).

Our goal is to maximize the sample rate,  $S$ . We formulate this sample rate maximization problem with the following considerations. First, node  $i$  has an average power of  $P_i$ , a circuit power consumption of  $P_{CT}$ , and a transmission distance  $d_i$  from the destination. These constraints are typical of energy constrained, short-range WSNs where the circuit power consumption is not negligible. Second, the signal is transmitted over an additive white Gaussian noise (AWGN) multiple access channel. The signal bandwidth and the noise power are denoted as  $B$  and  $P_N$ , respectively.

To maximize the sample rate, we must jointly consider the rate regions of the multiple access channel and the SW coding. The transmission rates of the communication links are bounded by the achievable rate region of the multiple access channel, which can be expressed as follows

$$r(G) \leq \mathcal{C}(G), \quad \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset, \quad (8.1)$$

where  $r(G) = \sum_{i \in G} r_i$  and  $\mathcal{C}(G)$  represents the maximum achievable sum rate of the

communication links in  $G$ . In other words,  $\mathcal{C}(G)$  for all  $G$  define the boundary of the multiple access channel capacity region. For example, when  $N = 2$ , the above region is described by  $r_1 \leq \mathcal{C}(\{1\})$ ,  $r_2 \leq \mathcal{C}(\{2\})$ , and  $r_1 + r_2 \leq \mathcal{C}(\{1, 2\})$ , where  $\mathcal{C}(\{1\})$ ,  $\mathcal{C}(\{2\})$ , and  $\mathcal{C}(\{1, 2\})$  depend on the particular MA scheme.

According to the Slepian-Wolf theorem [62][63], the achievable rate region of SW coding is

$$R(G) \geq \mathcal{H}(X(G)|X(G^c)), \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset, \quad (8.2)$$

where  $\mathcal{H}(\cdot)$  represents the entropy function,  $R(G) = \sum_{i \in G} R_i$ , and  $X(G) = \{X_j | j \in G\}$ . For instance, when  $N = 2$ , the SW coding rate region is described by  $R_1 \geq \mathcal{H}(X_1|X_2)$ ,  $R_2 \geq \mathcal{H}(X_2|X_1)$ , and  $R_1 + R_2 \geq \mathcal{H}(X_1, X_2)$ .

For each source node, there is a flow constraint, i.e., the input and output information rates should be equal. Together with the equal sample rate constraint, this flow constraint can be expressed as  $r_i = SR_i$ . Thus, considering the SW coding rate region, we have

$$r(G) \geq S\mathcal{H}(X(G)|X(G^c)), \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset. \quad (8.3)$$

Therefore, the sample rate maximization problem under the constraints imposed by the achievable multiple access channel and SW coding rate regions is

$$\begin{aligned} (\mathbf{P8.1}) \quad & \min -S \\ \text{s.t.} \quad & C_1 : r(G) \leq \mathcal{C}(G), \quad \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset \\ & C_2 : r(G) \geq S\mathcal{H}(X(G)|X(G^c)), \quad \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset. \end{aligned} \quad (8.4)$$

The optimization parameters are the transmission rates  $r_i$ , from which we can determine the optimal SW coding rates  $R_i^* = r_i^*/S^*$ . Problem (P8.1) is in fact a rate allocation problem which provides the jointly optimal  $r_i^*$  and  $R_i^*$  that support the maximum sample rate  $S^*$  at the nodes. Note that  $\mathcal{C}(G)$  in (P8.1) depends on the particular MA scheme we employ. Problem (P8.1) enables us to answer the following questions: For a given set of candidate MA schemes, i.e., CDMA, FDMA, and TDMA, which one can produce the maximum sample rate? (We refer to this as the optimal MA scheme.) How

is the optimal MA scheme related to the system design parameters? What are the corresponding optimal transmission and SW coding rates? To facilitate answering these questions, we first provide an equivalent optimization problem to (P8.1) in Section 8.3.

### 8.3 Performance Evaluation for MA Schemes with SW coding

In the following, we show that (P8.1) is equivalent to maximizing the sum of achievable transmission rates within a polyhedral cone rate region determined by the source entropies.

**Proposition 1:** Problem (P8.1) is equivalent to problem (P8.2) where

$$\begin{aligned}
 \text{(P8.2)} \quad & \min - \sum_{i=1}^N r_i \\
 \text{s.t. } & C_1 : r(G) \leq \mathcal{C}(G), \quad \forall G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset \\
 & C_2 : \mathbf{M}\mathbf{r} \leq \mathbf{0},
 \end{aligned} \tag{8.5}$$

where the column vector  $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_N]^T$  and the matrix  $\mathbf{M}$  is uniquely defined by the source entropies.

**Proof:** To prove proposition 1, we first note that given any feasible sample rate  $S$  in (P8.1), there exists a set of feasible transmission rates  $r_i$  such that

$$\sum_{i=1}^N r_i = S\mathcal{H}(X_1, X_2, \dots, X_N). \tag{8.6}$$

The proof of this statement follows the proof of Proposition 1 in [150] and is therefore omitted here for the sake of conciseness. This statement implies that (P8.1) is equivalent to maximizing the sum of transmission rates within the rate region of a MA scheme, and  $r_i^*$ s always reside in a region of the hyperplane in  $\mathbb{R}^N$  determined by (8.6). This region is bounded by the half-spaces defined by the constraints  $r^*(G) \geq$

$S^*\mathcal{H}(X(G)|X(G^c))$ , for all  $G \subset \{1, 2, \dots, N\}$ ,  $G \neq \emptyset$ . That is,  $r_i^*$ 's satisfy the following relations

$$\begin{aligned} r^*(G) &\geq S^*\mathcal{H}(X(G)|X(G^c)), r^*(G^c) \geq S^*\mathcal{H}(X(G^c)|X(G)), \\ r^*(G) + r^*(G^c) &= S^*\mathcal{H}(X_1, X_2, \dots, X_N). \end{aligned} \quad (8.7)$$

Since  $\mathcal{H}(X(G^c)|X(G)) + \mathcal{H}(X(G)) = \mathcal{H}(X_1, X_2, \dots, X_N)$ , we have the following inequalities

$$S^*\mathcal{H}(X(G)|X(G^c)) \leq r^*(G) \leq S^*\mathcal{H}(X(G)). \quad (8.8)$$

Using the fact that  $\sum_{i=1}^N r_i^* = S^*\mathcal{H}(X_1, X_2, \dots, X_N)$ , equivalently we have

$$\frac{\mathcal{H}(X(G)|X(G^c))}{\mathcal{H}(X_1, \dots, X_N)} \leq \frac{r^*(G)}{\sum_{i=1}^N r_i^*} \leq \frac{\mathcal{H}(X(G))}{\mathcal{H}(X_1, \dots, X_N)}. \quad (8.9)$$

The above expressions describe multiple half-spaces in  $\mathbb{R}^N$  passing through the origin. In other words, they define a polyhedral cone. For example, when  $N = 2$ , the polyhedral cone is described by

$$\frac{\mathcal{H}(X_1|X_2)}{\mathcal{H}(X_1, X_2)} \leq \frac{r_1}{r_1 + r_2} \leq \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_1, X_2)}, \quad \frac{\mathcal{H}(X_2|X_1)}{\mathcal{H}(X_1, X_2)} \leq \frac{r_2}{r_1 + r_2} \leq \frac{\mathcal{H}(X_2)}{\mathcal{H}(X_1, X_2)}. \quad (8.10)$$

The polyhedral cone in (8.9) can be represented in a more compact form

$$\mathbf{M}\mathbf{r} \leq \mathbf{0}, \quad (8.11)$$

where the column vector  $\mathbf{r} = [r_1 \ r_2 \ \dots \ r_N]^T$ . Matrix  $\mathbf{M}$  has  $2^N - 2$  rows and  $N$  columns, and its entries are determined by the source entropies. For example, for  $N = 2$ , and 3, we find

$$\mathbf{M} = \begin{bmatrix} \mathcal{H}(X_2|X_1) & -\mathcal{H}(X_1) \\ -\mathcal{H}(X_2) & \mathcal{H}(X_1|X_2) \end{bmatrix} \quad (8.12)$$

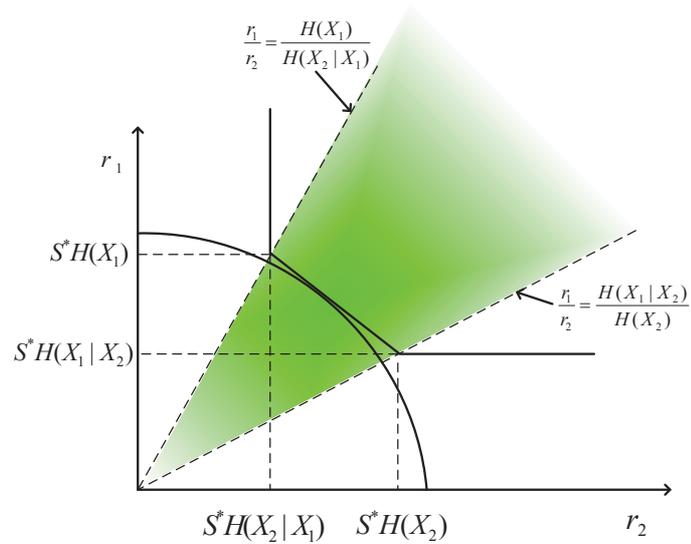


Figure 8.3: A visualization of Proposition 1 for  $N = 2$ .

and

$$\mathbf{M} = \begin{bmatrix} -\mathcal{H}(X_3) & -\mathcal{H}(X_3) & \mathcal{H}(X_1, X_2|X_3) \\ \mathcal{H}(X_3|X_1, X_2) & \mathcal{H}(X_3|X_1, X_2) & -\mathcal{H}(X_1, X_2) \\ \mathcal{H}(X_2, X_3|X_1) & -\mathcal{H}(X_1) & -\mathcal{H}(X_1) \\ -\mathcal{H}(X_2, X_3) & \mathcal{H}(X_1|X_2, X_3) & \mathcal{H}(X_1|X_2, X_3) \\ -\mathcal{H}(X_2) & \mathcal{H}(X_1, X_3|X_2) & -\mathcal{H}(X_2) \\ \mathcal{H}(X_2|X_1, X_3) & -\mathcal{H}(X_1, X_3) & \mathcal{H}(X_2|X_1, X_3) \end{bmatrix} \quad (8.13)$$

□.

Fig. 8.3 illustrates Proposition 1 for  $N = 2$ . The bottom left and upper right regions in Fig. 8.3, respectively, are defined by the achievable multiple access channel and SW coding rate regions. For a sample rate  $S$  to be feasible, the two rate regions must intersect. For a given MA scheme, as  $S$  increases, the SW coding rate region will shift to the upper right. Thus, for a sample rate  $S$  to be optimal, these two regions must be tangent as shown in Fig. 8.3, and  $r_i^*$  on the tangent point must satisfy  $r_1^* + r_2^* = S^* \mathcal{H}(X_1, X_2)$  according to (8.6). Moreover, Fig. 8.3 shows that  $r_i^*$ , i.e., the tangent

point, falls in the shaded polyhedral cone (which is simply a triangular region in the case of  $N = 2$ ) described by  $\mathbf{Mr} \leq \mathbf{0}$ .

Based on Proposition 1, for a given set of candidate MA schemes, we can find the optimal MA scheme by (i) solving the optimization problem in (P8.2) and finding the maximum of the sum of achievable transmission rates within the polyhedral cone rate region characterized by (8.9), that each MA scheme can provide, and (ii) letting the optimal MA scheme be the MA that provides the largest maximum of the sum of transmission rates. Note that at the end of step (i) we can calculate the maximum sample rate  $S^*$  that can be supported by each candidate MA scheme via  $S^* = \frac{\sum_{i=1}^N r_i^*}{H(X_1, \dots, X_N)}$  and the corresponding optimal SW coding rates via  $R_i^* = \frac{r_i^*}{S^*}$ .

## 8.4 Optimal Rate Allocation for MA Schemes with SW coding

Considering CDMA, FDMA, and TDMA as the MA scheme candidates, in the following we solve the optimization problem in (P8.2) for each scheme.

### 8.4.0.1 Optimal rate allocation for CDMA

We need to solve (P8.2) where  $\mathcal{C}(G)$  is substituted with  $\mathcal{C}_{\text{CDMA}}(G)$ , where [63]

$$\mathcal{C}_{\text{CDMA}}(G) = B \log_2 \left( 1 + \sum_{i \in G} \frac{P_i - P_{\text{CT}}}{d_i^n P_N} \right).$$

We realize that (P8.2) becomes a linear optimization problem, which can be easily solved. We observe that there exists at least one set  $G$ , denoted by  $G^*$ , for which the optimal transmission rates satisfy  $r^*(G^*) = \mathcal{C}_{\text{CDMA}}(G^*)$ . This is because (P8.2) becomes equivalent to finding a supporting hyperplane of the linear region defined by constraints  $C_1$  and  $C_2$  in  $\mathbb{R}^N$ . The gradient of the supporting hyperplane is the vector  $[1, 1, \dots, 1]$ . This supporting hyperplane is tangent with the multiple access channel rate

region defined by  $C_1$ , since the polyhedral cone defined by  $C_2$  is unbounded in the direction determined by the vector  $[1, 1, \dots, 1]$ . In the following we show

$$G^* = \arg \min_G \left\{ \frac{\mathcal{C}_{\text{CDMA}}(G)}{\mathcal{H}(X(G)|X(G^c))} \right\}. \quad (8.14)$$

For any set of transmission rates,  $r_i$ , the achievable sample rate is upper bounded by

$$S \leq \frac{r(G)}{\mathcal{H}(X(G)|X(G^c))} \leq \min_G \left\{ \frac{r(G)}{\mathcal{H}(X(G)|X(G^c))} \right\}, G \subseteq \{1, 2, \dots, N\}, G \neq \emptyset. \quad (8.15)$$

We can always increase the set of transmission rates with the minimal value of  $\frac{r(G)}{\mathcal{H}(X(G)|X(G^c))}$  within the boundary defined by  $\mathcal{C}_{\text{CDMA}}(G)$ . One can repeat this process until no further increase in  $\min_G \left\{ \frac{r(G)}{\mathcal{H}(X(G)|X(G^c))} \right\}$  is possible, in which case the set of transmission rates with the minimum value of  $\frac{r(G)}{\mathcal{H}(X(G)|X(G^c))}$  reaches the boundary defined by  $\mathcal{C}_{\text{CDMA}}(G)$ . The resulting maximum sample rate is

$$S^* = \frac{\mathcal{C}_{\text{CDMA}}(G^*)}{\mathcal{H}(X(G^*)|X(G^{*c}))}, \exists G^* \subseteq \{1, 2, \dots, N\}, G^* \neq \emptyset. \quad (8.16)$$

Consider two given sets  $G, G' \subseteq \{1, 2, \dots, N\}$ , such that

$$\frac{\mathcal{C}_{\text{CDMA}}(G)}{\mathcal{H}(X(G)|X(G^c))} > \frac{\mathcal{C}_{\text{CDMA}}(G')}{\mathcal{H}(X(G')|X(G'^c))}. \quad (8.17)$$

If we assume  $G^* = G$ , we have

$$\begin{aligned} r(G') &\geq S^* \mathcal{H}(X(G')|X(G'^c)), \\ &= \mathcal{H}(X(G')|X(G'^c)) \frac{\mathcal{C}_{\text{CDMA}}(G)}{\mathcal{H}(X(G)|X(G^c))}, \\ &> \mathcal{C}_{\text{CDMA}}(G'), \end{aligned} \quad (8.18)$$

which violates constraint  $C_1$ . Therefore, we reach  $G^*$  in (8.14).  $\square$

We conclude that  $r_i^*$ s are any feasible points of the following linear region

$$\begin{aligned} r^*(G) &\leq \mathcal{C}_{\text{CDMA}}(G), \\ r^*(G^*) &= \mathcal{C}_{\text{CDMA}}(G^*), \text{ where } G^* = \arg \min_G \left\{ \frac{\mathcal{C}_{\text{CDMA}}(G)}{\mathcal{H}(X(G)|X(G^c))} \right\}. \end{aligned} \quad (8.19)$$

### 8.4.0.2 Optimal rate allocation for TDMA

We need to solve (P8.2) where  $\mathcal{C}(G)$  is substituted by  $\mathcal{C}_{\text{TDMA}}(G)$ . By definition  $\mathcal{C}_{\text{TDMA}}(G)$  is the solution of the following optimization problem

$$\begin{aligned}
 \text{(P8.3)} \quad \mathcal{C}_{\text{TDMA}}(G) &= \max_{\theta_i} \sum_{i \in G} B\theta_i \log_2 \left( 1 + \frac{P_i/\theta_i - P_{\text{CT}}}{d_i^n P_N} \right) \\
 \text{s.t.} \quad \sum_{i \in G} \theta_i &\leq 1; \quad \theta_i > 0, \forall i \in G
 \end{aligned} \tag{8.20}$$

We recognize that to solve (P8.2) we first need to solve (P8.3).

Let  $\theta_i^*$  denote the optimal  $\theta_i$  obtained from solving (P8.3). For any given set  $G$ , introducing Lagrange multipliers  $\lambda_i^*$  for the inequality constraints  $\theta_i > 0, \forall i \in G$ , and a multiplier  $\mu^*$  for the inequality constraint  $\sum_{i \in G} \theta_i \leq 1$ , we obtain the Karush-Kuhn-Tucker (KKT) conditions [105]

$$\left\{ \begin{array}{l} \theta_i^* > 0, \lambda_i^* \geq 0, \mu^* \geq 0 \\ \sum_{i \in G} \theta_i^* \leq 1, \lambda_i^* \theta_i^* = 0 \\ \mu^* (\sum_{i \in G} \theta_i^* - 1) = 0 \\ \mathcal{R}'(G, \theta_i^*) - \lambda_i^* + \mu^* = 0 \end{array} \right. \tag{8.21}$$

where  $\mathcal{R}(G, \theta_i^*) = \sum_{i \in G} B\theta_i^* \log_2 \left( 1 + \frac{P_i/\theta_i^* - P_{\text{CT}}}{d_i^n P_N} \right)$  and  $\mathcal{R}'(G, \theta_i^*)$  represents the first order derivative of  $\mathcal{R}(G, \theta_i^*)$  with respect to  $\theta_i^*$ .

Clearly  $\theta_i^* > 0$  implies  $\lambda_i^* = 0, \forall i \in G$ . Thus

$$\mathcal{R}'(G, \theta_i^*) = -\mu^*. \tag{8.22}$$

Solving (8.22) for  $\theta_i^*$ , we find

$$\theta_i^* = \frac{P_i}{e a_i d_i^n P_N [\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} + P_{\text{CT}} - d_i^n P_N}, \tag{8.23}$$

where  $a_i = (P_{\text{CT}}/(d_i^n P_N) - 1)/e$ ,  $e$  is the base of natural logarithm and  $\mathcal{W}(\cdot)$  is the Lambert function. Equation (8.23) shows that to calculate  $\theta_i^*$  we need to first find  $\mu^*$ . Considering the KKT condition  $\mu^* (\sum_{i \in G} \theta_i^* - 1) = 0$ , two cases are possible: either

$\mu^* = 0$  and the resulting  $\theta_i^*$ s must satisfy  $\sum_{i=1}^N \theta_i^* \leq 1$ , or  $\mu^* > 0$  and the resulting  $\theta_i^*$ s must satisfy  $\sum_{i=1}^N \theta_i^* = 1$ . Thus, the main complexity in calculating  $\theta_i^*$  is to find  $\mu^* > 0$  such that the resulting  $\theta_i^*$ s satisfy  $\sum_{i \in G} \theta_i^* = 1$ . If such a solution does not exist, then  $\mu^* = 0$ . Unfortunately, this is not analytically tractable. To facilitate the numerical search for  $\mu^* > 0$ , we derive the following lower and upper bounds for  $\mu^*$ , denoted by  $\mu_L^*$  and  $\mu_U^*$ , respectively

$$\begin{aligned}\mu_L^* &= -B \log_2 \left[ \nu e^\nu \left( \frac{P_{CT}}{P_N d_{\max}^n} - 1 \right)^{-1} \right] - \frac{B}{\ln 2}, \\ \mu_U^* &= -B \log_2 \left[ \nu e^\nu \left( \frac{P_{CT}}{P_N d_{\min}^n} - 1 \right)^{-1} \right] - \frac{B}{\ln 2}.\end{aligned}\quad (8.24)$$

where  $\nu = \left( \sum_{i=1}^N P_i / (P_{CT} - P_N d_i^n) - 1 \right)^{-1}$ . Also,  $d_{\max}$  and  $d_{\min}$  are the maximum and the minimum of  $d_i$ , respectively. The detailed derivations of these bounds are given as follows.

Suppose there exists  $\mu^* > 0$  such that the resulting  $\theta_i^*$  satisfy  $\sum_{i \in G} \theta_i^* = 1$ . Therefore, using (8.23) we have

$$\begin{aligned}1 &= \sum_{i \in G} \frac{P_i}{e a_i d_i^n P_N [\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} + P_{CT} - d_i^n P_N} \\ &\leq \sum_{i \in G} \frac{P_i}{e a_i d_i^n P_N \min\{[\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1}\} + P_{CT} - d_i^n P_N} \\ &= \frac{1}{\min\{[\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1}\} + 1} \sum_{i \in G} P_i / (P_{CT} - d_i^n P_N) \\ \Rightarrow \sum_{i \in G} P_i / (P_{CT} - d_i^n P_N) - 1 &\geq \min_i \left\{ [\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} \right\}.\end{aligned}\quad (8.25)$$

Similarly, we have

$$\sum_{i \in G} P_i / (P_{CT} - d_i^n P_N) - 1 \leq \max_i \left\{ [\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} \right\}.\quad (8.26)$$

Since  $[\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} = \left[ \mathcal{W}\left(\frac{P_{CT}/(d_i^n P_N) - 1}{e} 2^{-\mu^*/B}\right) \right]^{-1}$  is a monotonic function of  $d_i$ , there exists a  $d \in [d_{\min}, d_{\max}]$  where  $d_{\min} = \min_i d_i$  and  $d_{\max} = \max_i d_i$ , such that

$$\sum_{i \in G} P_i / (P_{CT} - d_i^n P_N) - 1 = \left[ \mathcal{W}\left(\frac{P_{CT}/(d^n P_N) - 1}{e} 2^{-\mu^*/B}\right) \right]^{-1}.\quad (8.27)$$

Define  $\nu = \left( \sum_{i \in G} P_i / (P_{\text{CT}} - P_N d_i^n) - 1 \right)^{-1}$ , solving (8.27) for  $\mu^*$  we find

$$\mu^* = -B \log_2 \left[ \nu e^\nu \left( \frac{P_{\text{CT}}}{P_N d^n} - 1 \right)^{-1} \right] - \frac{B}{\ln 2}, \quad (8.28)$$

which is a monotonically decreasing function of  $d$ . Using the fact that  $d_{\min} \leq d \leq d_{\max}$ , we reach the lower and upper bounds presented in (8.24). Since  $\theta_i^*$  is a monotonic function of  $\mu$ ,  $\mu^*$  can be found numerically through a binary search between the bounds.

Once  $\mu^*$  and  $\theta_i^*$  are found, From (P8.3) we have

$$\mathcal{C}_{\text{TDMA}}(G) = \sum_{i \in G} B \theta_i^* \log_2 \left( 1 + \frac{P_i / \theta_i^* - P_{\text{CT}}}{d_i^n P_N} \right). \quad (8.29)$$

Having solved (P8.3), now we return to solving (P8.2) when  $\mathcal{C}(G)$  is substituted with  $\mathcal{C}_{\text{TDMA}}(G)$  in (8.29). Unfortunately, an analytical solution to this optimization problem in general remains elusive. For the special case where the transmission rates obtained from (8.29) lie within the polyhedral cone defined by  $\mathbf{M}\mathbf{r} \leq \mathbf{0}$ , these rates are in fact the optimal transmission rates  $r_i^*$ . However, when the transmission rates obtained from (8.29) do not lie within the polyhedral cone, we resort to exhaustive search to solve (P8.2).

### 8.4.0.3 Optimal rate allocation for FDMA

We need to solve (P8.2) where  $\mathcal{C}(G)$  is substituted with  $\mathcal{C}_{\text{FDMA}}(G)$ . By definition,  $\mathcal{C}_{\text{FDMA}}(G)$  is the solution of the following optimization problem

$$\begin{aligned} \text{(P8.4)} \quad \mathcal{C}_{\text{FDMA}}(G) = & \max_{\theta_i} \sum_{i \in G} B \theta_i \log_2 \left( 1 + \frac{P_i - P_{\text{CT}}}{\theta_i d_i^n P_N} \right) \\ & \text{s.t.} \sum_{i \in G} \theta_i \leq 1; \quad \theta_i > 0, \forall i \in G \end{aligned} \quad (8.30)$$

We recognize that to solve (P8.2) we first need to solve (P8.4). To solve (P8.4), we take a similar approach as we did for solving (P8.3). By invoking the KKT conditions, we find the optimal  $\theta_i^*$  as follows

$$\theta_i^* = \frac{P_i - P_{\text{CT}}}{e a_i d_i^n P_N [\mathcal{W}(a_i 2^{-\mu^*/B})]^{-1} - d_i^n P_N}, \quad (8.31)$$

where  $a_i = -1/e$ , such that  $\sum_{i \in G} \theta_i^* \leq 1$  when  $\mu^* = 0$  or  $\sum_{i \in G} \theta_i^* = 1$  for some  $\mu^* > 0$ . Similar to the TDMA case, once  $\mu^*$  and  $\theta_i^*$  are found, from (P8.4) we have

$$\mathcal{C}_{\text{FDMA}}(G) = \sum_{i \in G} B \theta_i^* \log_2 \left( 1 + \frac{P_i - P_{\text{CT}}}{\theta_i^* d_i^n P_N} \right). \quad (8.32)$$

For the special case where the transmission rates obtained from (8.32) lie within the polyhedral cone defined by  $\mathbf{M}\mathbf{r} \leq \mathbf{0}$ , these rates are in fact the optimal transmission rates  $r_i^*$ . However, when the transmission rates obtained from (8.32) do not lie within the polyhedral cone, we resort to exhaustive search to solve (P8.2).

One important result is that, considering the circuit power consumption, the rate region of CDMA contains that of FDMA. This can be easily shown as follows

$$\begin{aligned} \mathcal{C}_{\text{CDMA}}(G) &= B \log \left( 1 + \sum_{i \in G} \frac{P_i - P_{\text{CT}}}{d_i^n P_N} \right) \\ &= B \log \left( \sum_{i \in G} \theta_i + \sum_{i \in G} \theta_i \frac{P_i - P_{\text{CT}}}{d_i^n P_N \theta_i} \right) = B \log \left( \sum_{i \in G} \theta_i \left( 1 + \frac{P_i - P_{\text{CT}}}{d_i^n P_N \theta_i} \right) \right) \\ &\geq \sum_{i \in G} \theta_i B \log \left( 1 + \frac{P_i - P_{\text{CT}}}{d_i^n P_N \theta_i} \right), \forall G \in \{1, 2, \dots, N\}, G \neq \emptyset. \end{aligned} \quad (8.33)$$

where the inequality in (8.33) is based on Jensen's inequality [63]. This inequality holds for any set of  $\theta_i$  where  $\theta_i > 0$  and  $\sum \theta_i \leq 1$ , including the set which maximizes the right side of the inequality. Hence  $\mathcal{C}_{\text{CDMA}}(G) \geq \mathcal{C}_{\text{FDMA}}(G)$ , implying that the rate region of CDMA contains that of FDMA, i.e., comparing CDMA and FDMA, the former always provides a larger sum of the achievable transmission rates. Combined this with Proposition 1, we find that CDMA can always support a higher sample rate than FDMA. Thus, in the subsequent analysis we only consider CDMA and TDMA.

## 8.5 Optimal MA and Rate Allocation: Two-Source Node Case

In this section, we show that, unlike the general case, for  $N = 2$  the optimality of TDMA and CDMA can be directly evaluated without the need to solve (P8.2) for both

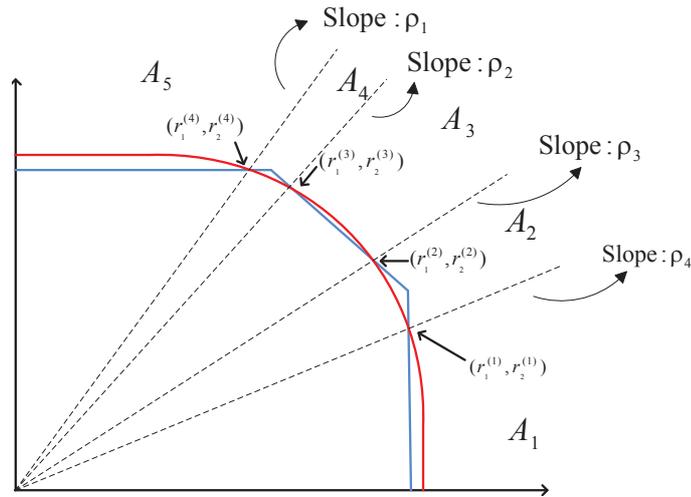


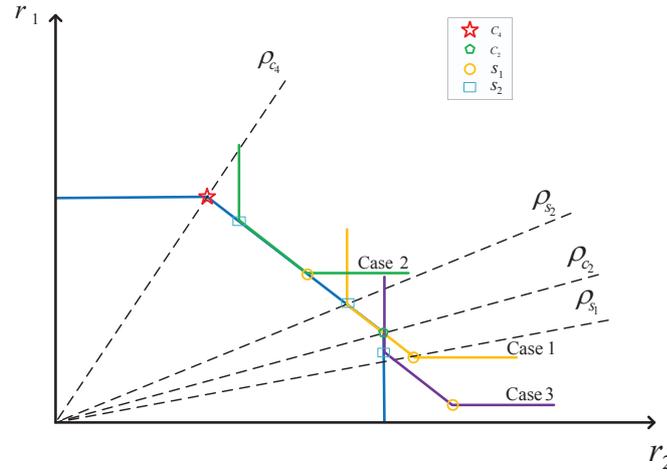
Figure 8.4: Region selection of TDMA and CDMA for  $N = 2$ .

TDMA and CDMA.

Fig. 8.4 depicts the rate regions of CDMA and TDMA with the consideration of circuit power consumption. We observe that the CDMA and TDMA rate region boundaries intersect in at most four points<sup>1</sup>, denoted by  $[r_1^{(i)}, r_2^{(i)}]$  for  $i = 1, 2, 3, 4$ . Consequently, the non-negative quadrant is divided into five regions denoted by  $\mathcal{A}_i$ , for  $i = 1, 2, 3, 4, 5$ . Each region can be described as  $\mathcal{A}_i = \{(r_1, r_2) | \rho_{i-1} < \frac{r_2}{r_1} < \rho_i\}$ , where  $\rho_0 = 0$ ,  $\rho_i = \frac{r_2^{(i)}}{r_1^{(i)}}$ , for  $i = 1, 2, 3, 4$ , and  $\rho_5 = \infty$ . The expressions for  $r_1^{(i)}, r_2^{(i)}$  are derived in Appendix A. It is clear that TDMA is optimal in  $\mathcal{A}_1, \mathcal{A}_3$  and  $\mathcal{A}_5$ , and CDMA is optimal in  $\mathcal{A}_2$  and  $\mathcal{A}_4$ .

As suggested by Proposition 1, the MA scheme that provides the highest sum rate in the polyhedral cone  $\mathcal{B} = \{(r_1, r_2) | \frac{\mathcal{H}(X_1|X_2)}{\mathcal{H}(X_2)} < \frac{r_2}{r_1} < \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2|X_1)}\}$  is optimal. Therefore, if  $\mathcal{B}$  overlaps with  $\mathcal{A}_3$ , TDMA is optimal and the optimal point denoted by  $[r_1^*, r_2^*]$  is in  $\mathcal{A}_3$ . If  $\mathcal{B}$  does not overlap with  $\mathcal{A}_3$ , but it overlaps with  $\mathcal{A}_2$  (or  $\mathcal{A}_4$ ), CDMA is optimal and the optimal point is in  $\mathcal{A}_2$  (or  $\mathcal{A}_4$ ). And if  $\mathcal{B}$  does not overlap with  $\mathcal{A}_3$  and  $\mathcal{A}_2$

<sup>1</sup>Although it is possible that the two boundaries intersect in fewer points, in the following we focus on the more general case where there are four intersection points. Other cases, can be treated similarly, and with less complexity.

Figure 8.5: Rate allocation of CDMA for  $N = 2$ .

(or  $\mathcal{A}_4$ ), TDMA is optimal and the optimal point is in  $\mathcal{A}_1$  (or  $\mathcal{A}_5$ ). These observations are summarized in Table 8.1. In the following, for each case, we find  $r_1^*, r_2^*$ . Having these, one can calculate the maximized sample rate  $S^* = (r_1^* + r_2^*)/\mathcal{H}(X_1, X_2)$  and the optimal SW coding rates  $R_1^* = r_1^*/S^*$  and  $R_2^* = r_2^*/S^*$ .

Table 8.1: The optimal MA scheme for  $N = 2$ 

Optimal MA	$[r_1^*, r_2^*] \in$	The polyhedral cone	The rate ratio
TDMA	$\mathcal{A}_1$	$\mathcal{B} \subset \mathcal{A}_1$	$\frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2 X_1)} < \rho_1$
CDMA	$\mathcal{A}_2$	$\mathcal{B} \cap \mathcal{A}_3 = \emptyset, \mathcal{B} \cap \mathcal{A}_2 \neq \emptyset$	$\rho_1 < \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2 X_1)} < \rho_2$
TDMA	$\mathcal{A}_3$	$\mathcal{B} \cap \mathcal{A}_3 \neq \emptyset$	$\rho_2 < \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2 X_1)}$ and $\frac{\mathcal{H}(X_1 X_2)}{\mathcal{H}(X_2)} < \rho_3$
CDMA	$\mathcal{A}_4$	$\mathcal{B} \cap \mathcal{A}_3 = \emptyset, \mathcal{B} \cap \mathcal{A}_4 \neq \emptyset$	$\rho_3 < \frac{\mathcal{H}(X_1 X_2)}{\mathcal{H}(X_2)} < \rho_4$
TDMA	$\mathcal{A}_5$	$\mathcal{B} \subset \mathcal{A}_5$	$\rho_4 < \frac{\mathcal{H}(X_1 X_2)}{\mathcal{H}(X_2)}$

### 8.5.1 CDMA

Let us denote the corners of the CDMA rate region in  $\mathcal{A}_2$  and  $\mathcal{A}_4$  by

$$\mathcal{C}_2 = \left[ B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N + P_2 - P_{\text{CT}}} \right), B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) \right], \quad (8.34)$$

and

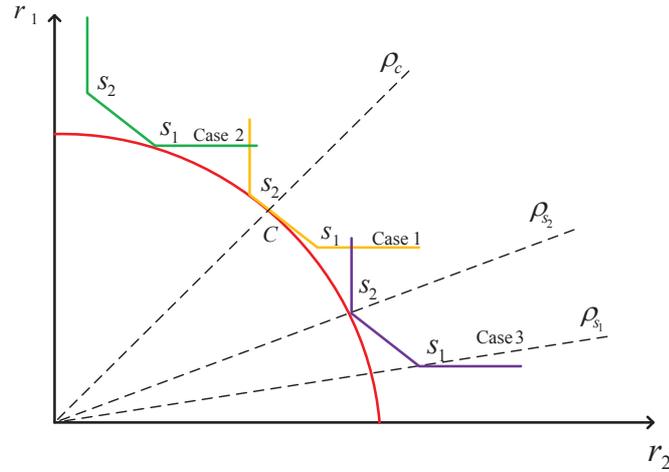
$$\mathcal{C}_4 = \left[ B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} \right), B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N + P_1 - P_{\text{CT}}} \right) \right], \quad (8.35)$$

respectively. Similarly, denote the corners of the SW rate region by  $\mathcal{S}_1 = [S\mathcal{H}(X_1|X_2), S\mathcal{H}(X_2)]$  and  $\mathcal{S}_2 = [S\mathcal{H}(X_1), S\mathcal{H}(X_2|X_1)]$  (see Fig. 8.5). Furthermore, denote the slopes of lines passing through the origin and these corners by

$$\begin{aligned} \rho_{\mathcal{C}_2} &= \frac{\log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N + P_2 - P_{\text{CT}}} \right)}{\log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right)}, & \rho_{\mathcal{C}_4} &= \frac{\log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} \right)}{\log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N + P_1 - P_{\text{CT}}} \right)}, \\ \rho_{\mathcal{S}_1} &= \frac{\mathcal{H}(X_1|X_2)}{\mathcal{H}(X_2)}, & \rho_{\mathcal{S}_2} &= \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2|X_1)}, \end{aligned} \quad (8.36)$$

respectively. Now, suppose the optimal point  $[r_1^*, r_2^*] \in \mathcal{A}_2$ . It is clear that  $\mathcal{C}_2$  is the unconstrained optimal point (without the constraint imposed by the polyhedral cone  $\mathcal{B}$ ). Thus, if  $\mathcal{C}_2 \in \mathcal{B}$ , or  $\rho_{\mathcal{S}_1} < \rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_2}$ , then any point on the line piece connecting  $\mathcal{C}_2$  and  $\mathcal{S}_1$ , including the corner  $\mathcal{C}_2$ , is optimal (case 1 in Fig. 8.5). Otherwise, if  $\mathcal{C}_2$  falls below  $\mathcal{B}$ , or  $\rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_1}$ , then any point on the line piece connecting  $\mathcal{S}_1$  and  $\mathcal{S}_2$ , including the corner  $\mathcal{S}_1$ , is optimal (case 2 in Fig. 8.5). Finally, if  $\mathcal{C}_2$  falls above  $\mathcal{B}$ , or  $\rho_{\mathcal{S}_2} < \rho_{\mathcal{C}_2}$ , then the optimal point is  $\mathcal{S}_2$  (case 3 in Fig. 8.5). To summarize, when  $[r_1^*, r_2^*] \in \mathcal{A}_2$ , we have

$$\begin{aligned} r_1^* &= \begin{cases} \min \left\{ B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} \right), \right. \\ \left. B \frac{\mathcal{H}(X_1|X_2)}{\mathcal{H}(X_1, X_2)} \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) \right\} & \rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_1} \\ B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N + P_2 - P_{\text{CT}}} \right) & \rho_{\mathcal{S}_1} < \rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_2} \\ B \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2|X_1)} \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) & \rho_{\mathcal{S}_2} < \rho_{\mathcal{C}_2} \end{cases} \\ r_2^* &= \begin{cases} r_1^* \frac{\mathcal{H}(X_2)}{\mathcal{H}(X_1|X_2)} & \rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_1} \\ B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) & \rho_{\mathcal{S}_1} < \rho_{\mathcal{C}_2} < \rho_{\mathcal{S}_2} \\ B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) & \rho_{\mathcal{S}_2} < \rho_{\mathcal{C}_2} \end{cases} \end{aligned} \quad (8.37)$$

Figure 8.6: Rate allocation of TDMA for  $N = 2$ .

Similarly, when  $[r_1^*, r_2^*] \in \mathcal{A}_4$ , we have

$$\begin{aligned}
 r_1^* &= \begin{cases} B \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) & \rho_{C_4} < \rho_{S_1} \\ B \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) & \rho_{S_1} < \rho_{C_4} < \rho_{S_2} \\ r_2^* \frac{\mathcal{H}(X_1)}{\mathcal{H}(X_2|X_1)} & \rho_{S_2} < \rho_{C_4} \end{cases} \quad (8.38) \\
 r_2^* &= \begin{cases} B \frac{\mathcal{H}(X_2)}{\mathcal{H}(X_1|X_2)} \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) & \rho_{C_4} < \rho_{S_1} \\ B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N + P_1 - P_{CT}} \right) & \rho_{S_1} < \rho_{C_4} < \rho_{S_2} \\ \min \left\{ B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right), \right. \\ \left. B \frac{\mathcal{H}(X_2|X_1)}{\mathcal{H}(X_1, X_2)} \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \right\} & \rho_{S_2} < \rho_{C_4} \end{cases}
 \end{aligned}$$

### 8.5.2 TDMA

Let  $\mathcal{C} = [r_1^\dagger, r_2^\dagger]$  denote the unconstrained optimal point (without the constraint imposed by the polyhedral cone  $\mathcal{B}$ ) (Fig. 8.6). This point is the solution of (P8.3) for the special case of  $N = 2$ , and the solution for the optimal  $\theta_i$  is given in (8.23). It is clear that if  $\mathcal{C} \in \mathcal{B}$ , or  $\rho_{S_1} < \frac{r_1^\dagger}{r_2^\dagger} < \rho_{S_2}$ , then this point is optimal and  $[r_1^*, r_2^*] = [r_1^\dagger, r_2^\dagger]$  (case 1 in Fig. 8.6). On the other hand, if  $\mathcal{C}$  falls below  $\mathcal{B}$ , or  $\frac{r_1^\dagger}{r_2^\dagger} < \rho_{S_1}$ , the optimal point must

satisfy  $\frac{r_1^*}{r_2^*} = \rho_{S_1}$  (case 2 in Fig. 8.6). We note that this condition is always satisfied in  $\mathcal{A}_5$ . Similarly, if  $\mathcal{C}$  falls above  $\mathcal{B}$ , or  $\frac{r_1^\dagger}{r_2^\dagger} > \rho_{S_2}$ , the optimal point must satisfy  $\frac{r_1^*}{r_2^*} = \rho_{S_2}$  (case 3 in Fig. 8.6). This condition is always satisfied in  $\mathcal{A}_5$ . Thus, in these cases, the optimal point is the solution of

$$\begin{aligned} \max \quad & \theta_1 \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) + \theta_2 \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \\ \text{s.t.} \quad & \theta_1, \theta_2 > 0, \quad \theta_1 + \theta_2 \leq 1, \quad \theta_1 \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) = \rho \theta_2 \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \end{aligned} \quad (8.39)$$

where  $\rho$  is equal to either  $\rho_{S_1}$  or  $\rho_{S_2}$  based on the conditions given above.

### 8.5.3 Overall Rate Allocation Algorithm

Based on the results in Sections 8.5.1 and 8.5.2, we can find the optimal rate allocation using the procedure described in the flow chart in Fig 8.7.

## 8.6 Numerical Results

In this section, we provide numerical results to demonstrate the benefits of the proposed jointly optimal rate allocation algorithm. In our simulations, we assume the circuit power consumption of all source nodes are identically equal to  $P_{CT} = 0.1$  W, the signal bandwidth is 100 KHz, the AWGN has a power spectral density of  $-174$  dBm/Hz, and the path loss exponent is  $n = 4$ .

Figs. 8.8 and 8.9 quantify the increase in the sample rate due to the joint optimization of transmission and SW coding rates, when the optimal MA is TDMA. Similar numerical gains are observed for the case where the optimal MA is CDMA and are omitted for the sake of conciseness. We compare the maximum sample rate  $S^*$  as  $d_1$  varies for five different approaches: (i) Jointly optimal SW coding and transmission rates. We denote this approach by “ $R_i^*, \theta_i^*$ ”. (ii) Optimal transmission rates with equal SW coding rates at each source node. We refer to this approach as “Equal  $R_i, \theta_i^*$ ”. (iii)

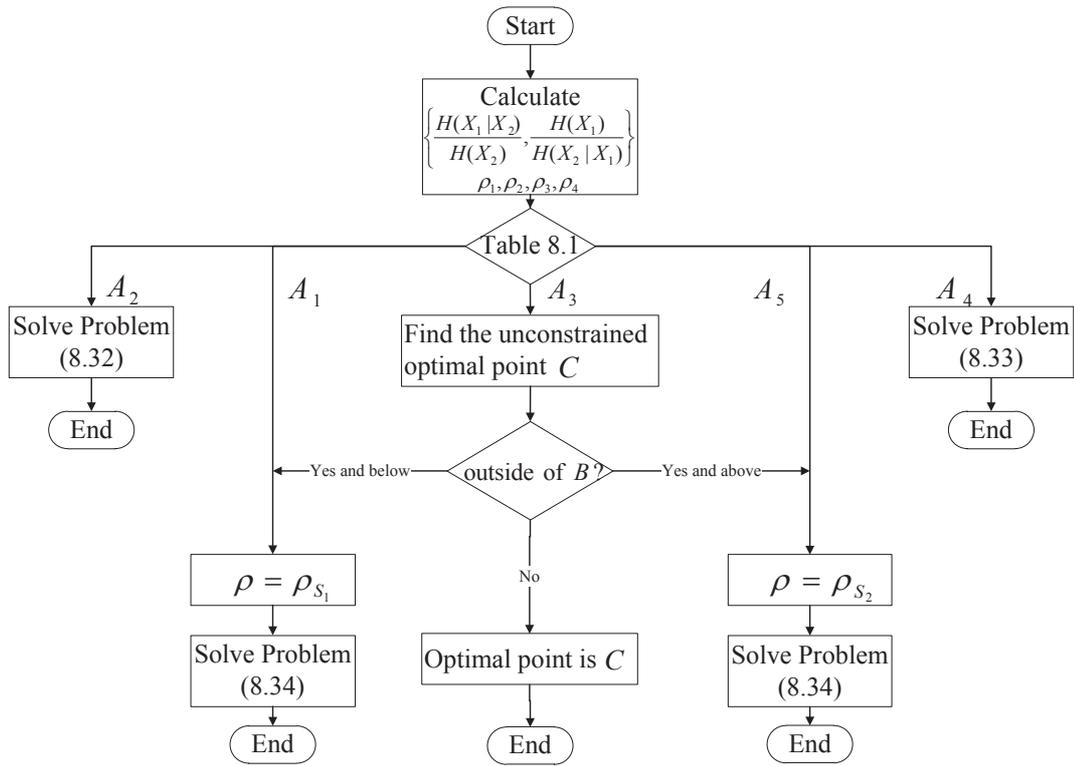


Figure 8.7: The flowchart of the optimal rate allocation for CDMA and TDMA with  $N = 2$ .

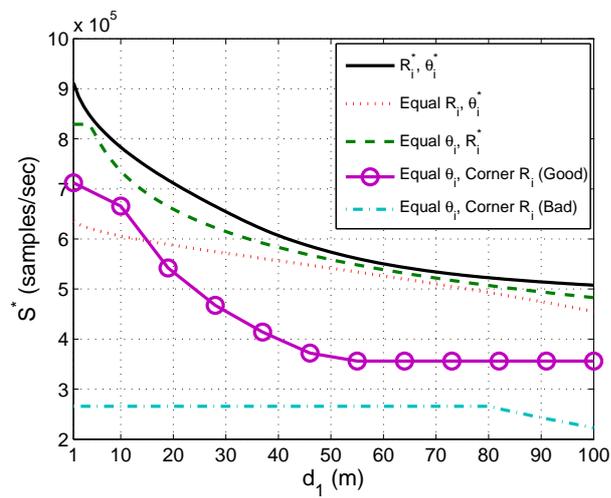


Figure 8.8:  $S^*$  versus  $d_1$  for TDMA and  $N = 3$ .

Optimal SW coding rates with equal TDMA time slots. We refer to this approach as “ $R_i^*$ , Equal  $\theta_i$ ”. (iv) Equal TDMA time slots with a simple SW coding rate allocation strategy which assigns the highest SW coding rate to the source with the highest transmission rate. We refer to this approach as “Equal  $\theta_i$ , Corner  $R_i$  (Good)”, and (v) Equal TDMA time slots with a simple SW coding rate allocation strategy which assigns the lowest SW coding rate to the source with the highest transmission rate. We refer to this approach as “Equal  $\theta_i$ , Corner  $R_i$  (Bad)”.

In Fig. 8.8, we let  $N = 3$  and fix  $d_2 = 50$  m and  $d_3 = 80$  m, while we increase  $d_1$  from 1 m to 100 m. The nodes have equal average power of  $P_i = 0.11$  W. For the sake of simplicity, we assume symmetric entropy where  $\mathcal{H}(X_i) = 1, \forall i \in \{1, 2, 3\}$ ,  $\mathcal{H}(X_i|X_j) = 0.5, \forall i, j \in \{1, 2, 3\}, i \neq j$ , and  $\mathcal{H}(X_i|X_j, X_k) = 0.25, \forall i, j, k \in \{1, 2, 3\}, i \neq j \neq k$ .

Fig. 8.8 shows that the “ $R_i^*, \theta_i^*$ ” approach always performs the best. The performance of the “Equal  $R_i, \theta_i^*$ ” and “ $R_i^*, \text{Equal } \theta_i$ ” approaches is closely related to the experiment setup (e.g. average power, transmission distances and source entropies). If the configuration is symmetric, i.e., nodes have the same average powers and transmission distances and the source entropies are symmetric, the partial optimizations in “Equal  $R_i, \theta_i^*$ ” and “ $R_i^*, \text{Equal } \theta_i$ ” approaches are equivalent to the “ $R_i^*, \theta_i^*$ ” approach. Furthermore,  $S^*$  in “Equal  $R_i, \theta_i^*$ ” and “ $R_i^*, \text{Equal } \theta_i$ ” approaches is similar to that of the “ $R_i^*, \theta_i^*$ ” approach when  $d_1, d_2, d_3$  are also similar, i.e.,  $50 \text{ m} \leq d_1 \leq 70 \text{ m}$ . On the other hand, the performance loss becomes more significant when the differences between the transmission distances increase.

The “Equal  $\theta_i$ , Corner  $R_i$  (Good)” approach is an intuitively sensible SW coding rate allocation strategy with the purpose of coupling the source coding rates with the transmission rates. However, this strategy suffers from thresholding and over compensation. For example, when the difference between the transmission rates with equal  $\theta_i$  is small, an equal SW coding rate allocation is more likely to be a better choice than a corner SW coding rate allocation. This also explains the large performance loss when

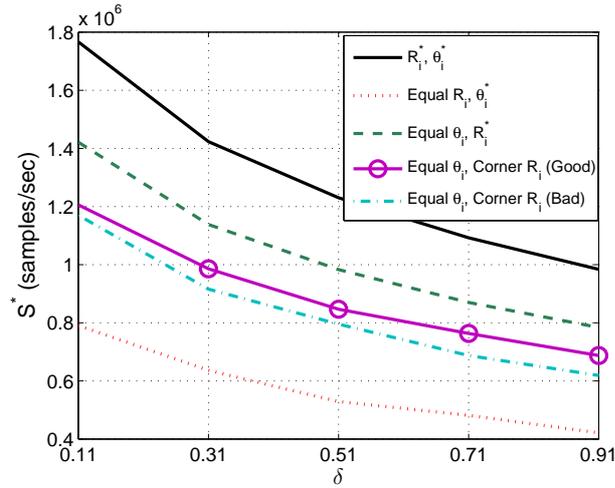


Figure 8.9:  $S^*$  versus  $\delta$  for TDMA and  $N = 2$ .

the transmission distances of the three source nodes are similar, i.e.,  $50 \text{ m} \leq d_1 \leq 70 \text{ m}$ . The “Equal  $\theta_i$ , Corner  $R_i$  (Bad)” approach performs the worst, indicating how much loss can an unwise rate allocation cause.

The benefits of the “ $R_i^*, \theta_i^*$ ” approach are more pronounced in non-symmetric scenarios. In Fig. 8.9, we let  $N = 2$ ,  $P_i = 1.11 - \delta\tau_i \text{ W}$  and  $d_i = 1 + 100\delta\tau_i \text{ m}$ , where  $\tau_i$ s are independent and uniformly distributed between 0 and 1, and the constant  $\delta \in \{0.11, 0.31, 0.51, 0.71, 0.91\}$ . The entropies of both sources are randomly generated with a fixed joint entropy of  $\mathcal{H}(X_1, X_2) = 1.5$ . Each point in Fig. 8.9 is an averaged value of a 1000 trials.

Fig. 8.9 demonstrates that the “ $R_i^*, \theta_i^*$ ” approach provides the largest  $S^*$  among the five approaches. For instance, the value of  $S^*$  provided by the “ $R_i^*, \theta_i^*$ ” approach is up to 2.4 times that of the “Equal  $R_i, \theta_i^*$ ” and “ $R_i^*, \text{Equal } \theta_i$ ” approaches; and is up to 1.5 times that of the “Equal  $\theta_i, \text{Corner } R_i$  (Good)” and “Equal  $\theta_i, \text{Corner } R_i$  (Bad)” approaches.

It is interesting to observe that the “Equal  $R_i, \theta_i^*$ ” approach performs worse than the “Equal  $\theta_i, \text{Corner } R_i$  (Good)” and “Equal  $\theta_i, \text{Corner } R_i$  (Bad)” approaches. This is because, in a non-symmetric scenario, the linear constraint imposed by  $R_1 = R_2$  does

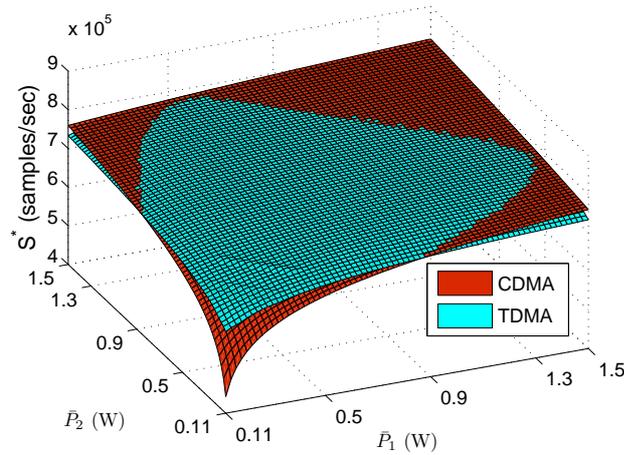


Figure 8.10:  $S^*$  of the “ $R_i^*, \theta_i^*$ ” approach for CDMA and TDMA in terms of  $P_1$  and  $P_2$  for  $N = 3$ ,  $P_3 = 110$  mW.

not necessarily intersect with the SW coding rate region on the boundary specified by  $R_1 + R_2 = \mathcal{H}(X_1, X_2)$ . As a result, we may have  $R_1^* + R_2^* > \mathcal{H}(X_1, X_2)$ , which greatly exacerbates the performance of SW coding. Therefore, the “Equal  $R_i, \theta_i^*$ ” approach is not favorable in practice.

Fig. 8.10 compares the value of  $S^*$  provided by the “ $R_i^*, \theta_i^*$ ” approach for CDMA and TDMA. We let  $N = 3$ ,  $d_1 = d_2 = d_3 = 50$  m,  $P_3 = 110$  mW, while we increase  $P_2$  and  $P_1$  from 110 mW to 1.5 W. We assume symmetric entropy with the same parameters as those used in Fig. 8.8. The value of  $S^*$  is calculated following the procedure described in Section 8.4. As the average powers change, the MA rate regions change as well. For a given rate region defined by (8.11) through source entropies, either CDMA or TDMA can provide the highest sum rate within this rate region as the average powers change. This is evident in Fig. 8.10, which shows that SW coding with TDMA provides larger  $S^*$  for some average power values, while SW coding with CDMA provides larger  $S^*$  for other values.

## 8.7 Summary

In this chapter, we first evaluated the performance of CDMA, TDMA and FDMA in a Gaussian multiple access channel for a sensing system with average power constraint and the consideration of circuit power consumption, in terms of the achievable sample rate when SW source coding is utilized in the sensing device. We concluded that the performance of either CDMA or TDMA can be the best among the three candidate MA schemes depending on the entropies of the random sources, while FDMA is never optimal. Further, we developed an algorithm to obtain the jointly optimal allocation of SW coding rates and CDMA or TDMA transmission rates. Numerical results are provided to demonstrate the performance of the proposed rate allocation algorithm.

## 9 Conclusions and Future Directions

This research dissertation provides an in-depth investigation of cross-layer design to improve the energy efficiency of short range wireless networks. Extending from the traditional physical layer design, I incorporate the energy efficient design of the data link layer, multiple access layer, network layer, and application layer. I propose numerous innovative designs and algorithms that greatly improve the energy efficiency of short-range wireless networks.

### 9.1 Contributions

#### 9.1.1 PHY-layer and Data Link Layer Optimization

First, I start my analysis with an emphasis on the impact of physical layer parameter selection on the energy efficiency of short-range wireless networks. I show that the optimal transmit power, modulation scheme and relay distance are crucial in achieving energy efficiency for a short-range wireless network. The optimal transmit power and modulation scheme are important factors for the physical layer design of short-range wireless networks; while the optimal hop distance has great potential for use in route selection for the network layer design. The analysis is conducted in both AWGN channels and block Rayleigh fading channels.

In the second part of the dissertation, I propose a cross-layer optimization of the PHY and Link layers to maximize the energy efficiency of a short-range wireless network. I investigate the energy consumption minimization problem for a single link for short-range wireless networks over an AWGN channel. Specifically, I propose a generic energy consumption per information bit model considering circuit power, information packetization, communication overhead and duty cycle. Based on this model, I developed an unconstrained, a fixed average power constrained, and a fixed average rate constrained case.

In the unconstrained case, I perform an optimization over both target bit error probability and packet length to minimize the energy consumption per information bit, with the consideration of retransmissions, a detailed packet structure and MQAM modulation schemes. The target bit error probability optimization is a form of the transmit power optimization as they are closely related, and the packet length optimization is to find the balance between the overhead and retransmission probability. For both coded and uncoded MQAM modulations, I find the closed forms for the optimum values of packet length and target bit error probability for a given transmission distance by removing the integer constraint on the packet length. The main results are: when transmission distance is short, a system adopting large packet length, small target bit error probability, and high bandwidth-efficient modulation schemes (*e.g.*, uncoded high-order QAM) is more energy efficient. On the other hand, when transmission distance is large, a system using small packet length, large target bit error probability, and high energy efficient modulation schemes (*e.g.*, coded BPSK) is energy efficient.

Correspondingly, energy minimizations are conducted using the fixed average power and fixed average rate constrained models with the introduction of an additional parameter: duty cycle. It is shown that the minimization of energy consumption per information bit is equivalent to the maximization of information rate for the fixed average power transmission.

In addition, I perform an optimization over packet length, duty cycle, and constella-

tion size to minimize the energy consumption in the fixed average power transmission case. I perform an optimization over packet length, duty cycle, constellation size, and transmit power to minimize energy consumption in the fixed average rate transmission case. I conclude that, fixed average rate transmissions and duty cycling are energy efficient at short transmission distances, while fixed average power transmissions and duty cycling are energy efficient at large transmission distances. The optimization in this part of the dissertation is meaningful in the sense that it provides a guideline to select physical layer parameters (e.g., the transmit power) and data link layer parameters (e.g., the packet length) given a communication environment, such as the transmission distance.

In the third part of the thesis, I investigate the use of cross-layer optimization in IR-UWB networks. I provide the power consumption models of typical transmitter and receiver structures of IR-UWB systems. Moreover I consider a frequency selective time-invariant channel. The optimization parameters used to minimize the energy consumption per information bit are: packet overhead, retransmission, repetition coding and number of RAKE fingers.

I conclude that low-complexity, low-performance transmission schemes are energy efficient at high SNRs, while high-complexity, high-performance schemes are energy efficient at low SNRs. Moreover, I provide the detailed optimum transmission schemes, including packet length, modulations, detection, repetition coding, combining, and number of RAKE fingers, for given transmission distances for a typical IR-UWB link. This information is important for it can serve as a lookup table for the transceiver to choose the optimum transmission schemes. The optimization performed in this part of the dissertation is also on the physical and data link layers.

### **9.1.2 PHY-layer and Network Layer Optimization**

In the fourth part of the thesis, I begin to consider two of the import network layer factors: network topology and multi-hop. In particular, I choose clustering topology as the network topology of interest due to its high potential in improving the energy efficiency of wireless networks.

In this part of the dissertation, I jointly optimize three aspects of a SRWN: the transmit power of the nodes, the cluster head selection and how to choose a route. The contributions of this part are three-fold: first, I propose an iteration-free algorithm for maximizing the amount of data gathered by a cluster; moreover, I developed a low complexity cluster head selection (CHS) algorithm to determine the optimal cluster head; furthermore, I present a low complexity algorithm to decide whether or not multi-hop (only 2-hop route is considered) should be used within a cluster to further improve the performance of a cluster. I show that the performance gain by using the optimal transmit power, the optimal cluster head selection, and the possible use of multi-hop is significant. The proposed algorithms feature low complexity and high performance and they can be easily implemented in the nodes. Thus, the potential of these algorithms in field applications is considerable. The optimization performed in this part of the thesis is on the physical and network layer.

### **9.1.3 PHY-layer and Application Layer Optimization**

In the fifth part of the dissertation, by building upon previous analysis, I further include the concept of distributed source coding.

I propose low-complexity solutions to maximize the amount of samples gathered during a cluster lifetime with Slepian-Wolf (SW) coding for an arbitrary number of sources, with energy constraints and SW coding constraints. The optimization parameters are the transmit time durations (which determine the transmit power) and source coding rates of the source nodes.

I derive closed form solutions for the joint optimal transmit time durations and SW coding rates when there is no active energy constraint at the fusion center. I also derive closed form solutions for the near optimal transmit time durations and source coding rates when there is an energy constraint at the fusion center by resorting to linear approximation. A simple criteria for determining whether the fusion center energy constraint is active or not is also provided. The optimization gain in the number of collected samples is shown to be significant.

The main contributions are the derivation of the closed form solutions that avoid the exceedingly complex iterations in traditional optimization techniques. The proposed algorithms are of significance since a low complexity high performance algorithm is critical for wireless nodes which have very limited energy and computational ability. The optimization performed in this part of the thesis is on the physical layer (transmit power optimization) and application layer (source coding allocation).

#### **9.1.4 Multi-access Layer and Application Layer Optimization**

Finally, I investigate the performance of SW coding over multiple access channels under the considerations of circuit power consumption and average power constraint. Different from our previous study where I assume an infinite number of information samples are available at the transmitters, in this study the information data is delivered over the multiple access channel as the samples are gathered from the environment. This research is of significant practical importance for modern sensing systems built upon short range wireless networks, such as wireless sensor networks, where onboard storage space is rather limited, circuit power consumption cannot be neglected and a stringent average power constraint is often imposed to regulate the functional lifetime of the sensing device.

Through our study, I find that, among CDMA, TDMA, and FDMA, SW coding performs best either with CDMA or TDMA based on the source entropy, and always

performs worst with FDMA. This discovery is, to the best of my knowledge, unprecedented. Also in this work, I propose a joint optimal rate allocation of SW coding rates and CDMA/TDMA rates to maximize the sample rate achievable at the sensing device. The proposed rate allocation features low complexity and is partially analytically solvable.

## **9.2 Future Directions**

While this dissertation has provided many techniques and ideas to improve the energy efficiency of short-range wireless networks, much research still remains.

### **9.2.1 Extension of DSC Optimal Rate Allocation**

One interesting direction is to consider the optimal rate allocation of DSC in multi-hop networks with stringent power and energy constraints. This problem is a natural extension of the work in this dissertation, and it possesses important practical significance. In SRWNs, the data sources may not be able to send the information to the sink through direct communication due to the limited transmit power of the nodes. In this case, multi-hop communication is necessary. Moreover, multi-hop communication introduces many challenging yet interesting constraints on the rate allocation problem of DSC, e.g., the choice of the next hop and the energy storage of the intermediate nodes. This is a promising research topic that deserves substantial research efforts.

### **9.2.2 Implementation of DSC Optimal Rate Allocation**

In this dissertation, optimal rate allocation procedures of DSC are proposed from a theoretical perspective. It is interesting to further consider the implementation issue of the optimal rate allocation of DSC in a practical wireless network. When performing

the optimization, instead of using information theoretical boundaries, we need to consider the actual information rates that are achievable through practical modulation and coding schemes, the realistic source coding rate constraints imposed by implementable DSC techniques, and more complicated yet realistic time varying fading channel models. Although it is highly unlikely that an elegant analytical optimal solution would exist in this case, it is certainly intriguing to study the possible rules of the optimal rate allocation of DSC for a practical wireless network.

# A Intersection Points of CDMA and TDMA Rate Regions Boundaries

## A.1 Finding the points $[r_1^{(1)}, r_2^{(1)}]$ and $[r_1^{(4)}, r_2^{(4)}]$

The point  $[r_1^{(1)}, r_2^{(1)}]$  is the intersection of the TDMA rate region boundary with the following line segment

$$r_2 = B \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right), \quad 0 \leq r_1 \leq B \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N + P_1 - P_{CT}} \right). \quad (\text{A.1})$$

Since the TDMA rate region boundary is specified by (8.20), to find this intersection point we need to solve

$$\begin{aligned} \max \quad & \theta_1 \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) + \theta_2 \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \\ \text{s.t.} \quad & \theta_1 > 0, \quad \theta_2 > 0, \quad \theta_1 + \theta_2 \leq 1 \\ & \theta_2 \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) = \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \\ & 0 \leq \theta_1 \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N} \right) \leq \log_2 \left( 1 + \frac{P_1 - P_{CT}}{d_1^n P_N + P_1 - P_{CT}} \right) \end{aligned} \quad (\text{A.2})$$

It is easy to show that the intersection takes place in the region where  $\theta_1 + \theta_2 = 1$ , since (A.2) is a line segment parallel to (A.1) in the region where  $\theta_1 + \theta_2 < 1$ . Now, let  $0 < \theta_2^\dagger < 1$  be a solution of

$$\theta_2 \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) = \log_2 \left( 1 + \frac{P_2 - P_{CT}}{d_2^n P_N} \right) \quad (\text{A.3})$$

Then  $\theta_1 = 1 - \theta_2^\dagger$  satisfies (A.1). Thus if

$$0 \leq (1 - \theta_2^\dagger) \log_2 \left( 1 + \frac{\frac{P_1}{1-\theta_2^\dagger} - P_{\text{CT}}}{d_1^n P_N} \right) \leq \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N + P_1 - P_{\text{CT}}} \right) \quad (\text{A.4})$$

then

$$[r_1^{(1)}, r_2^{(1)}] = \left( B(1 - \theta_2^\dagger) \log_2 \left( 1 + \frac{\frac{P_1}{1-\theta_2^\dagger} - P_{\text{CT}}}{d_1^n P_N} \right), B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right) \right) \quad (\text{A.5})$$

Otherwise, this point does not exist. The point  $[r_1^{(4)}, r_2^{(4)}]$  can be calculated in a similar manner.

## A.2 Finding the points $[r_1^{(2)}, r_2^{(2)}]$ and $[r_1^{(3)}, r_2^{(3)}]$

We note that these points are the intersection points of the borders of the TDMA and CDMA rate regions, limited to

$$B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N + P_2 - P_{\text{CT}}} \right) \leq r_1 \leq B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} \right), \quad (\text{A.6})$$

$$B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N + P_1 - P_{\text{CT}}} \right) \leq r_2 \leq B \log_2 \left( 1 + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right). \quad (\text{A.7})$$

Thus  $[r_1^{(2)}, r_2^{(2)}]$  and  $[r_1^{(3)}, r_2^{(3)}]$  are the solutions to the following equation set

$$r_1 + r_2 = B \log_2 \left( 1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} \right), \quad (\text{A.8})$$

$$r_1 = B\theta \log_2 \left( 1 + \frac{\frac{P_1 - P_{\text{CT}}}{\theta}}{d_1^n P_N} \right), r_2 = B(1 - \theta) \log_2 \left( 1 + \frac{\frac{P_2 - P_{\text{CT}}}{1-\theta}}{d_2^n P_N} \right). \quad (\text{A.9})$$

Equation (A.6) implies that the intersection points are on the boundary of the CDMA rate region, and (A.7), (A.8), and (A.9) imply that the intersection points are on the boundary of the TDMA rate region where  $\theta_1 + \theta_2 \leq 1$  is active. Alternatively,

$$1 + \frac{P_1 - P_{\text{CT}}}{d_1^n P_N} + \frac{P_2 - P_{\text{CT}}}{d_2^n P_N} = \left( 1 + \frac{\frac{P_1 - P_{\text{CT}}}{\theta}}{d_1^n P_N} \right)^\theta \left( 1 + \frac{\frac{P_2 - P_{\text{CT}}}{1-\theta}}{d_2^n P_N} \right)^{(1-\theta)}. \quad (\text{A.10})$$

Equation (A.10) has only one degree of freedom and is easy to solve numerically. If this equation has two distinct solutions that satisfy (A.7), then  $[r_1^{(2)}, r_2^{(2)}]$  and  $[r_1^{(3)}, r_2^{(3)}]$  are obtained. Otherwise, these points do not exist.

## Bibliography

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, August 2002.
- [2] A. J. Goldsmith and S. B. Wicker, “Design challenges for energy-constrained ad hoc wireless networks,” *Wireless Communications, IEEE*, vol. 9, no. 4, pp. 8–27, Aug 2002.
- [3] S. Cui, A. J. Goldsmith, and A. Bahai, “Energy-constrained modulation optimization,” *IEEE Transactions on Wireless Communications*, vol. 4, no. 8, pp. 2349–2360, September 2005.
- [4] J. Korhonen and Y. Wang, “Effect of packet size on loss rate and delay in wireless links,” in *Proceedings of the IEEE Wireless Communications and Networking Conference*, Mar. 2005, pp. 1608 – 1613.
- [5] Y. Hou, M. Hamamura, and S. Zhang, “Performance tradeoff with adaptive frame length and modulation in wireless network,” in *Proceedings of the IEEE International Conference on Computer and Information Technology*, September 2005, pp. 490– 494.
- [6] C. Chien, M. B. Srivastava, R. Jain, P. Lettieri, V. Aggarwal, and R. Sternowski, “Adaptive radio for multimedia wireless links,” *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 5, pp. 793–813, May 1999.

- [7] A. Goldsmith, *Wireless Communications*, 1st ed. NY: Cambridge University Press, 2005.
- [8] H. Zimmermann, "OSI Reference Model - The ISO Model of Architecture for Open Systems Interconnection," *IEEE Trans. on Communications*, vol. 28, no. 4, pp. 425–432, Apr. 1980.
- [9] R. Madan, S. Cui, S. Lall, and A. J. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," in *Proceedings of the 24th Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM)*, 2005.
- [10] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A Survey on Sensor Networks," *IEEE Communications Magazine*, pp. 102–114, Aug. 2002.
- [11] C. Zhao, M. Perillo, and W. Heinzelman, "General Network Lifetime and Cost Models for Evaluating Sensor Network Deployment Strategies," *IEEE Trans. on Mobile Computing*, pp. 484–497, Apr. 2008.
- [12] C. S. Taek and A. Goldsmith, "Degrees of Freedom in Adaptive Modulation : A Unified View," in *Proceedings of Vehicular Technology Conference, Spring, IEEE (VTS '01)*, 2001.
- [13] Y. Chen, E. Sirer, and S. Wicker, "On selection of optimal transmission power for ad hoc networks," in *Proceedings of the Thirty-Sixth Hawaii International Conference on System Sciences (HICSS-36)*, January 2003, p. 10.
- [14] J. Ammer and J. Rabaey, "The energy-per-useful-bit metric for evaluating and optimizing sensor network physical layers," in *Proceedings of the IEEE International Workshop on Wireless Ad-hoc and Sensor Networks*, September 2006, pp. 695–700.

- [15] A. Wang, S. Chao, C. Sodini, and A. Chandrakasan, "Energy efficient modulation and mac for asymmetric rf microsensor system," in *Proceedings of the IEEE International Symposium on Low Power Electronics and Design*, August 2001, pp. 106–111.
- [16] V. Rodoplu and T. Meng, "Minimum energy mobile wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 8, pp. 1333–1344, August 1999.
- [17] P. Chen, B. O'Dea, and E. Callaway, "Energy efficient system design with optimum transmission range for wireless ad hoc networks," in *Proceedings of the IEEE International Conference on Communications*, May 2002, pp. 945–952.
- [18] S. Panichpapiboon, G. Ferrari, and O. K. Tonguz, "Optimal common transmit power in ad hoc wireless networks," in *Proceedings of the IEEE International Performance, Computing, and Communications Conference*, April 2005, pp. 593–597.
- [19] J. Deng, Y. S. Han, P. N. Chen, and P. K. Varshney, "Optimum transmission range for wireless ad hoc networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, Mar. 2004, pp. 1024–1029.
- [20] E. Modiano, "An adaptive algorithm for optimizing the packet size used in wireless arq protocols," *Wireless Networks*, vol. 5, pp. 279–286, 1999.
- [21] Y. Sankarasubramaniam, I. F. Akyildiz, and S. W. McLaughlin, "Energy efficiency based packet size optimization in wireless sensor networks," in *Proceedings of the IEEE International Workshop on Sensor Network Protocols and Applications*, May 2003, pp. 1–8.
- [22] S. Ci, H. Sharif, and K. Nuli, "Study of an adaptive frame size predictor to enhance energy conservation in wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 283–292, February 2005.

- [23] P. Lettieri and M. B. Srivastava, "Adaptive Frame Length Control for Improving Wireless Link Throughput, Range, and Energy Efficiency," *Proc. INFOCOM '98*, Mar. 1998.
- [24] Y. Sankarasubramaniam, I. Akyildiz, and S. McLaughlin, "Energy Efficiency based Packet Size Optimization in Wireless Sensor Networks," *Proc. of the First IEEE. 2003 IEEE International Workshop on Sensor Network Protocols and Applications*, 2003.
- [25] J. Ammer and J. Rabaey, "The Energy-per-Useful-Bit Metric for Evaluating and Optimizing Sensor Network Physical Layers," *3rd Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON '06)*, Sep. 2006.
- [26] R. K. Reddy, S. De, and H. M. Gupta, "Joint control of transmit power and frame size for energy-optimized data transfer in wireless sensor networks," *Transactions on Communications, IEICE*, vol. E93-B, no. 8, pp. 2043–2052, Aug 2010.
- [27] A. Wang, S. Cho, C. G. Sodini, and A. P. Chandrakasan, "Energy Efficient Modulation and MAC for Asymmetric RF Microsensor Systems," *Proc. of the 2001 international symposium on Low power electronics and design*, Aug. 2001.
- [28] J. Deng, Y. Han, P. Chen, and P. Varshney, "Optimum Transmission Range for Wireless Ad Hoc Networks," *Proc. of Wireless Communications and Networking Conference (WCNC)*, Mar. 2004.
- [29] V. Rodoplu and T. Meng, "Minimum Energy Mobile Wireless Networks," *Journal on Selected Areas in Communications, IEEE*, vol. 17, no. 8, pp. 1333–1344, Aug. 1999.
- [30] C. Chien, M. Srivastava, R. Jain, P. Lettieri, V. Aggarwal, and R. Sternowski, "Adaptive Radio for Multimedia Wireless Links," *Journal on Selected Areas in Communications, IEEE*, vol. 17, no. 5, pp. 793–813, May 1999.

- [31] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-Constrained Modulation Optimization," *IEEE Trans. on Wireless Communications*, vol. 4, no. 8, pp. 2349–2360, Sep. 2005.
- [32] F. C. Commission, "Part 15," *the Commission's Rules Regarding Ultra-Wideband Transmission Systems Federal Communications Commission*, Jul. 2008.
- [33] I. Gresham, A. Jenkins, R. Egri, C. Eswarappa, N. Kinayman, N. Jain, R. Anderson, F. Kolak, R. Wohlert, S. Bawell, J. Bennett, and J. Lanteri, "Ultra-wideband Radar Sensors for Short-range Vehicular Applications," *IEEE Trans. on Microwave Theory and Techniques*, vol. 52, pp. 2105–2122, Sep. 2004.
- [34] S. Gezici, T. Zhi, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization Via Ultra-wideband Radios: A Look at Positioning Aspects for Future Sensor Networks," *IEEE Signal Processing Magazine*, vol. 22, pp. 70–84, Jul. 2005.
- [35] I. Oppermanz, L. Stoica, A. Rabbachin, Z. Shelby, and J. Haapola, "UWB Wireless Sensor Networks: UWEN - a Practical Example," *IEEE Communications Magazine*, vol. 42, pp. S27–S32, Dec. 2004.
- [36] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications Magazine, IEEE*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [37] S. Cui, A. J. Goldsmith, and A. Bahai, "Exact Bit Error Rate Analysis of TH-PPM UWB Systems in The Presence of Multiple-Access Interference," *IEEE Communications Letters*, vol. 7, no. 12, pp. 572–574, Dec. 2003.
- [38] Q. Zhang and J. H. Cho, "On RAKE Receivers for Ultra-Wideband Binary Block-Coded PPM in Dense Multipath Channels," *IEEE Trans. on Vehicular Technology*, vol. 56, no. 4, pp. 1737–1747, Jul. 2007.

- [39] J. D. Choi and W. E. Stark, "Performance of Ultra-wideband Communications with Suboptimal Receivers in Multipath Channels," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 9, pp. 1754–1766, Dec. 2002.
- [40] N. Telzhensky and Y. Leviatan, "Novel Method of UWB Antenna Optimization for Specified Input Signal Forms by Means of Genetic Algorithm," *IEEE Trans. on Antennas And Propagation*, vol. 54, no. 8, pp. 2216–2225, Aug. 2006.
- [41] R. Zhang and X. Dong, "Synchronization and Integration Region Optimization for UWB Signals with Non-coherent Detection and Auto-correlation Detection," *IEEE Trans. on Communications*, vol. 54, no. 5, pp. 790–798, May 2008.
- [42] Y. Shi, Y. T. Hou, and H. D. Sherali, "Cross-Layer Optimization for Data Rate Utility Problem in UWB-based Ad Hoc Networks," *IEEE Trans. on Mobile Computing*, vol. 7, no. 6, pp. 764–777, Jun. 2008.
- [43] U. Onunkwo and Y. Li, "On the Optimum Pulse-Position Modulation Index for Ultra-Wideband Communication," in *Proceedings of the IEEE 6th CAS Symposium on Emerging Technologies: Mobile and Wireless Communication*, Shanghai, China, Jun. 2004, pp. 77–80.
- [44] M. Z. Win and R. A. Scholtz, "On the Energy Capture of Ultrawide Bandwidth Signals in Dense Multipath Environments," *IEEE Communications Letters*, vol. 2, no. 9, pp. 245–247, Sep. 1998.
- [45] D. Cassioli, M. Z. Win, and A. F. Molisch, "The Ultra-Wide Bandwidth Indoor Channel: From Statistical Model to Simulations," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 20, no. 6, Aug. 2002.
- [46] A. F. Molisch, D. Cassioli, C.-C. Chong, S. Emami, A. Fort, B. Kannan, J. Karedal, J. Kunisch, H. G. Schantz, K. Siwiak, and M. Z. Win, "A Comprehensive Standardized Model for Ultrawideband Propagation Channels," *IEEE Trans. on Ant. and Propag.*, vol. 54, pp. 3151–3166, Nov. 2006.

- [47] K. Siwiak and A. Petroff, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal on Selected Areas in Communications*, vol. SAC-5, no. 2, pp. 128–137, Feb. 1987.
- [48] P. Y. Massaad, M. Medard, and L. Zheng, "Impact of Processing Energy on the Capacity of Wireless Channels," *Proc. International Symposium on Information Theory and its Applications (ISITA)*, 2004.
- [49] B. Prabhakar, E. Uysal-Biyikoglu, and A. E. Gamal, "Energy-efficient transmission over a wireless link via lazy packet scheduling," in *Proceedings of the 20th Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM)*, 2001.
- [50] A. E. Gamal, C. Nair, B. Prabhakar, E. Uysal-Biyikoglu, and S. Zahedi, "Energy-efficient scheduling of packet transmissions over wireless networks," in *Proceedings of the 21st Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM)*, 2002.
- [51] E. Uysal-Biyikoglu, B. Prabhakar, and A. E. Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Trans. on Networking*, pp. 487–499, Aug. 2002.
- [52] T. Wang, W. Heinzelman, and A. Seyedi, "Minimization of Transceiver Energy Consumption in Wireless Sensor Networks with AWGN Channels," *Proc. Allerton Conference*, Sep. 2008.
- [53] ———, "Minimizing Energy Consumption in IR-UWB Based Wireless Sensor Networks," *Proc. IEEE International Conference on Communications (ICC)*, Jun. 2009.
- [54] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient mac protocol for wireless sensor networks," *Infocom, IEEE*, vol. 3, no. 4, pp. 1567–1576, Aug 2002.

- [55] I. Demirkol, C. Ersoy, and F. Alagoz, “MAC Protocols for Wireless Sensor Networks: a Survey,” *IEEE Communications Magazine*, 2006.
- [56] S. Cui, A. Goldsmith, and A. Bahai, “Joint modulation and multiple access optimization under energy constraints,” in *Proceedings of Global Telecommunications Conference, IEEE (GLOBECOM '04)*, 2004.
- [57] O. Younis and S. Fahmy, “Distributed clustering in ad-hoc sensor networks: A hybrid, energy-efficient approach,” in *Proceedings of the 23rd Annual Joint Conf. of the IEEE Computer and Communications Societies (INFOCOM)*, 2004.
- [58] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “An application-specific protocol architecture for wireless microsensor networks,” *IEEE Trans. on Wireless Communications*, pp. 660–670, Oct 2002.
- [59] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, “A Performance Comparison of Multi-hop Wireless Ad Hoc Network Routing Protocols,” *Proceedings of the 4th annual ACM/IEEE International Conference on Mobile Computing and Networking*, pp. 85–97, 1998.
- [60] S. Banerjee and A. Misra, “Minimum Energy Paths for Reliable Communication in Multi-hop Wireless Networks,” *Proceedings of the 3rd ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pp. 146–156, 2002.
- [61] M. Perillo, Z. Cheng, and W. Heinzelman, “On the Problem of Unbalanced Load Distribution in Wireless Sensor Networks,” *Proceedings of the IEEE GlobeCom Workshops*, pp. 74–79, 2004.
- [62] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *Information Theory, IEEE Transactions on*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

- [63] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. NJ: Wiley-Interscience, 2006.
- [64] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *Signal Processing Magazine, IEEE*, vol. 21, no. 5, pp. 80–94, Sept. 2004.
- [65] S. Pradhan, J. Kusuma, and K. Ramachandran, "Distributed Compression in a Dense Microsensor Network," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 51–60, Mar. 2002.
- [66] S. Pradhan and K. Ramachandran, "Distributed Source Coding using Syndroms (DISCUS) Design and Construction," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.
- [67] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Compression of Binary Sources with Side Information at the Decoder using LDPC Codes," *IEEE Communications Letters*, vol. 6, no. 10, pp. 440–442, Oct. 2002.
- [68] X. Wang and M. T. Orchard, "Design of Trellis Codes for Source Coding with Side Informantion at the Decoder," *Proceedings of Data Compression Conference*, Mar. 2001.
- [69] J. Bajcsy and P. Mitran, "Coding for the Slepian-Wolf Problem using Turbo Codes," *Proceedings of IEEE Global Telecom. Conference (GLOBECOM)*, pp. 25–29, Nov. 2001.
- [70] P. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms, and Applications*. Elsevier Inc., 2009.
- [71] J. Chou, D. Petrovic, and K. Ramachandran, "A Distributed and Adaptive Signal Processing Approach to Reducing Energy Consumption in Sensor Networks,"

*Proceedings of the IEEE Twenty-Second International Annual Joint Conference (INFOCOM)*, 2003.

- [72] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, “Networked Slepian-Wolf : Theory, Algorithms, and Scaling Laws,” *IEEE Trans. on Information Theory*, vol. 51, no. 12, pp. 4057–4073, Dec. 2005.
- [73] E. Duarte-Melo and M. Liu, “A Computational Approach to the Joint Design of Distributed Data Compression and Data Dissemination in Field-gathering Wireless Sensor Network,” *Proceedings of Fourty One Annual Allerton Conference on Communication, Control, and Computing*, 2003.
- [74] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, “On Network Correlated Data Gathering,” *Proceedings of the IEEE Twenty-Third International Annual Joint Conference (INFOCOM)*, Mar. 2003.
- [75] D. Ganesan, R. Cristescu, and B. Beferull-Lozano, “Power Efficient Sensor Placement and Transmission Structure for Data Gathering under Distortion Constraints,” *Proceedings of the third International Symposium on Information Processing in Sensor Networks (IPSN)*, 2004.
- [76] S. J. Baek, G. Veciana, and X. Su, “Minimizing Energy Consumption in Large-scale Sensor Networks through Distributed Data Compression and Hierarchical Aggression,” *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1130–1140, Aug. 2004.
- [77] L. D. Paolo, C. Fischione, F. Graziosi, F. Santucci, and S. Tennina, “Performance Analysis of Distributed Source Coding and Packet Aggregation in Wireless Sensor Networks,” in *Proceedings of Global Telecommunications Conference, IEEE (GLOBECOM '06)*, 2006.

- [78] E. Fasolo, M. Rossi, J. Widmer, and M. Zorzi, “In-network Aggregation Techniques for Wireless Sensor Networks: a Survey,” *IEEE Wireless Communications*, vol. 14, no. 2, pp. 70–87, Apr. 2007.
- [79] S. Patten, B. Krishnamachari, and R. Govindan, “The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks,” in *Proceedings of the Third International Symposium on Information Processing in Sensor Networks (IPSN)*, 2004.
- [80] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “An application specific protocol architecture for wireless microsensor networks,” *IEEE Transactions on Wireless Communication*, vol. 1, no. 4, pp. 660–670, October 2002.
- [81] S. Shin and H. Schulzrinne, “Experimental measurement of the capacity for voip traffic in ieee 802.11 wlans,” *Proc. of INFOCOM, IEEE*, vol. 5, pp. 2018–2026, May 2007.
- [82] J. Chen, “Carrier recovery in burst-mode 16-qam,” *Thesis for the Degree of Master of Science*, vol. 5, Jun. 2004.
- [83] U. Vilaipornsawai and M. R. Soleymani, “A novel turbo coding scheme for satellite atm using reed-muller codes,” *IEEE Trans. on Communications*, vol. 51, pp. 767–773, May 2003.
- [84] J. Proakis, *Digital Communications*. McGraw-Hill Book Co., 2001.
- [85] H. Y. Cheung, K. S. Cheuizg, and J. Laun, “A low power monolithic agc with automatic dc offset cancellation for direct conversion hybrid cdma transceiver used in telemetering,” *IEEE International Symposium on Circuits and Systems.*, vol. 4, pp. 390–393, May 2001.
- [86] MoteIV, 2007, tmote Sky Data Sheet: [www.moteiv.com](http://www.moteiv.com).
- [87] J. G. Proakis, *Digital Communications*, 4th ed. MA: Addison-Wesley, 1972.

- [88] T. S. Rappaport, *Wireless Communications : Principles and Practice*, 2nd ed. MA: Addison-Wesley, 1972.
- [89] V. Raghunathan, C. Schurgers, P. Sung, and M. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing magazine*, vol. 19, pp. 40–50, Mar. 2002.
- [90] S. Verdú, "Spectral Efficiency in the Wideband Regime," *IEEE Trans. on Information Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
- [91] K. Dan, "A Generalized Logarithm for Exponential-Linear Equations," *The College Mathematics Journal*, 2001.
- [92] S. Cui, A. J. Goldsmith, and A. Bahai, "Power estimation for viterbi decoders," *Wireless Systems Lab, Stanford University.*, 2003.
- [93] M. Z. Win and R. A. Scholtz, "On the Energy Capture of Ultrawide Bandwidth Signal in Dense Multipath Environments," *IEEE Communications Letters*, vol. 2, no. 9, 1998.
- [94] D. Cassioli, M. Z. Win, F. Vatalaro, and A. F. Molisch, "Low Complexity Rake Receivers in Ultra-Wideband Channels," *IEEE Trans. on Wireless Communications*, vol. 4, pp. 1265–1275, Apr. 2007.
- [95] B. Verbruggen, J. Craninckx, M. Kuijk, P. Wambacq, and G. Van der Plas, "A 2.2 mW 5b 1.75 GS/s folding flash ADC in 90 nm digital CMOS," *Proc. ISSCC*, 2007.
- [96] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, July-Oct 1948.
- [97] E. Al-Hussaini and A. Al-Bassiouni, "Performance of MRC Diversity Systems for the Detection of Signals with Nakagami Fading," *IEEE Trans. on Comm.*, vol. 33, pp. 1315–1319, Dec. 1985.

- [98] ———, “MRC Performance for M-ary Modulation in Arbitrarily Correlated Nakagami Fading Channels,” *IEEE Trans. on Comm.*, vol. 47, pp. 44–52, Jan. 1999.
- [99] B. Sklar, *Digital Communications: Fundamentals and Applications*, 2nd ed. Prentice Hall PTR, 2001.
- [100] A. Medi and W. Namgoong, “A High Data-Rate Energy-Efficient Interference-Tolerant Fully Integrated CMOS Frequency Channelized UWB transceiver for Impulse Radio,” *Proc. JSSC*, 2008.
- [101] H. Kao, A. Chin, K. Chang, and S. McAlister, “A Low-power Current-reuse LNA for Ultra-wideband Wireless Receivers from 3.1 to 10.6 GHz,” *Proc. SiRF07*, 2007.
- [102] Y. Gao, K. Cai, Y. Zheng, and B.-L. Ooi, “A Wideband CMOS Multiplier for UWB Application,” *Proc. ICUWB*, 2007.
- [103] C. T. Fu and H. Luong, “A CMOS Linear-in-dB High-linearity Variable-gain Amplifier for UWB Receivers,” *Proc. ASSCC*, 2007.
- [104] V. Lottici, A. D’Andrea, and U. Mengali, “Channel Estimation for Ultra-Wideband Communications,” *Journal on Selected Areas in Communications, IEEE*, vol. 20, no. 9, pp. 1638–1645, Dec. 2002.
- [105] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, 1st ed. Cambridge, UK: Cambridge University Press, 2004.
- [106] R. W. Doerfler, *Dead Reckoning: Calculating without instruments*. Houston: Gulf Publishing Company, 1993.
- [107] E. M. Wright, “Solution of the Equation  $ze^z = a$ ,” *Bull. Amer. Math. Soc.*, vol. 65, pp. 89–93, 1959.

- [108] D. Marco and D. L. Neuhoff, "Reliability vs. efficiency in distributed source coding for field-gathering sensor networks," in *Proceedings of the 3rd international symposium on Information processing in sensor networks*, 2004, pp. 161–168.
- [109] K. Yuen, B. Liang, and B. Li, "A Distributed Framework for Correlated Data Gathering in Sensor Networks," *IEEE Trans. on Vehicular Technology*, vol. 57, no. 1, pp. 578–593, Jan. 2008.
- [110] S. He, J. Chen, D. Yau, and S. Youxian, "Cross-Layer Optimization of Correlated Data Gathering in Wireless Sensor Networks," in *7th Annual IEEE Communications Society on Sensor and Ad Hoc Communications and Networks (SECON '10)*, 2010.
- [111] W. Honggang, P. Dongming, W. Wang, H. Sharif, and H. Chen, "Cross-layer routing optimization in multirate wireless sensor networks for distributed source coding based applications," *IEEE Trans. on Wireless Communications*, vol. 7, no. 10, pp. 3999–4009, Oct. 2008.
- [112] Z. Tang, I. Glover, A. Evans, and J. He, "Energy efficient transmission protocol for distributed source coding in sensor networks," in *Communications, 2007. ICC '07. IEEE International Conference on*, june 2007, pp. 3870–3875.
- [113] W. Wang, D. Peng, H. Wang, H. Sharif, and H. H. Chen, "Cross-layer multirate interaction with distributed source coding in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 8, no. 2, pp. 787–795, feb. 2009.
- [114] A. Roumy and D. Gesbert, "Optimal matching in wireless sensor networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 725–735, dec. 2007.
- [115] S. Li and A. Ramamoorthy, "Rate and power allocation under the pairwise distributed source coding constraint," *Communications, IEEE Transactions on*, vol. 57, no. 12, pp. 3771–3781, dec. 2009.

- [116] S. Agnihotri, P. Nuggehalli, and H. Jamadagni, "On maximizing lifetime of a sensor cluster," in *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, june 2005, pp. 312–317.
- [117] H. Arjmandi and F. Lahouti, "Cost optimized distributed source coding for data gathering single-hop wireless sensor networks," in *Telecommunications (ICT), 2010 IEEE 17th International Conference on*, april 2010, pp. 104–109.
- [118] P. Youssef-Massaad, L. Zheng, and M. Medard, "Bursty transmission and glue pouring: on wireless channels with overhead costs," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 12, pp. 5188–5194, Dec. 2008.
- [119] P. Larranaga and J. A. Lozano, *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation*, 2nd ed. MA: Kluwer Academic Publishers, 2002.
- [120] C. Chow and C. Liu, "Approximating discrete probability distributions with dependence trees," *Information Theory, IEEE Transactions on*, vol. 14, no. 3, pp. 462–467, may 1968.
- [121] J. Beirlant, E. J. Dudewicz, L. Györfi, and E. C. V. D. Meulen, "Nonparametric entropy estimation : an overview," *Intern. J. Math. Stat. Sci.*, vol. 6, no. 1, pp. 1–14, june 2001.
- [122] A. Abbasi and M. M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826–2841, june 2007.
- [123] J. Li, A. Bose, and Y. Q. Zhao, "Rayleigh Flat Fading Channels' Capacity," in *Proc. 3rd Annual Communication Networks and Services Research Conference (CNSR)*, Halifax, Nova Scotia, May 2005.

- [124] F. Chapeau-Blondeau and A. Monir, "Numerical evaluation of the Lambert W function and application to generation of generalized Gaussian noise with exponent  $1/2$ ," *Signal Processing, IEEE Transactions on*, vol. 50, no. 9, pp. 2160–2165, Sep 2002.
- [125] M. Holland, T. Wang, B. Tavli, A. Seyedi, and W. Heinzelman, "Optimizing physical layer parameters for wireless sensor networks," *ACM Trans. on Sensor Networks*, vol. 7, no. 4, pp. 28 – 48, Feb. 2011.
- [126] M. C. M. Thein and T. Thein, "An energy efficient cluster-head selection for wireless sensor networks," in *Intelligent Systems, Modelling and Simulation (ISMS), 2010 International Conference on*, jan. 2010, pp. 287–291.
- [127] S. Soro and W. Heinzelman, "Cluster head election techniques for coverage preservation in wireless sensor networks," *Ad Hoc Networks*, vol. 7, no. 15, pp. 955–972, jul. 2009.
- [128] O. Younis and S. Fahmy, "An experimental study of routing and data aggregation in sensor networks," in *Mobile Adhoc and Sensor Systems Conference, 2005. IEEE International Conference on*, nov. 2005, pp. 49–57.
- [129] M. Sartipi and F. Fekri, "Distributed source coding in wireless sensor networks using ldpc coding: the entire slepian-wolf rate region," in *Wireless Communications and Networking Conference, 2005 IEEE*, vol. 4, march 2005, pp. 1939–1944.
- [130] M. Sartipi, "Ldpc codes for information embedding and lossy distributed source coding," in *Data Compression Conference (DCC), 2010*, march 2010, p. 551.
- [131] T. C. Kuo and A. Willson, "A flexible decoder ic for wimax qc-ldpc codes," in *Custom Integrated Circuits Conference, 2008. CICC 2008. IEEE*, sept. 2008, pp. 527–530.

- [132] S. Pradhan and K. Ramchandran, "Distributed source coding: symmetric rates and applications to sensor networks," in *Data Compression Conference (DCC)*, Mar. 2000, pp. 363–372.
- [133] F. Oldewurtel, J. Riihijarvi, and P. Mahonen, "Efficiency of distributed compression and its dependence on sensor node deployments," in *Vehicular Technology Conference (VTC 2010-Spring), IEEE 71st*, May 2010, pp. 1–5.
- [134] D. Krithivasan and S. Pradhan, "Distributed source coding using abelian group codes: A new achievable rate-distortion region," *Information Theory, IEEE Transactions on*, vol. 57, no. 3, pp. 1495–1519, Mar. 2011.
- [135] H. Luo, Y. Liu, and S. Das, "Routing correlated data in wireless sensor networks: A survey," *Network, IEEE*, vol. 21, no. 6, pp. 40–47, Nov.-Dec. 2007.
- [136] K. Yuen, B. Liang, and B. Li, "A distributed framework for correlated data gathering in sensor networks," *IEEE Transactions on Vehicular*, vol. 57, no. 1, pp. 578–593, Jan. 2008.
- [137] S. He, J. Chen, D. Yau, and Y. Sun, "Cross-layer optimization of correlated data gathering in wireless sensor networks," in *Sensor Mesh and Ad Hoc Communications and Networks (SECON), 2010 7th Annual IEEE Communications Society Conference on*, Jun. 2010, pp. 1–9.
- [138] R. Berry and E. Yeh, "Cross-layer wireless resource allocation," *Signal Processing Magazine, IEEE*, vol. 21, no. 5, pp. 59–68, Sept. 2004.
- [139] D. Tse and S. Hanly, "Multiaccess fading channels. i. polymatroid structure, optimal resource allocation and throughput capacities," *Information Theory, IEEE Transactions on*, vol. 44, no. 7, pp. 2796–2815, Nov. 1998.
- [140] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient mac protocol for wireless sensor networks," in *INFOCOM 2002. Twenty-First Annual Joint Confer-*

- ence of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, Jun. 2002, pp. 1567–1576.
- [141] ———, “Medium access control with coordinated adaptive sleeping for wireless sensor networks,” *Networking, IEEE/ACM Transactions on*, vol. 12, no. 3, pp. 493–506, Jun. 2004.
- [142] V. Rajendran, K. Obraczka, and J. J. Garcia-Luna-Aceves, “Energy-efficient, collision-free medium access control for wireless sensor networks,” *Wirel. Netw.*, vol. 12, pp. 63–78, Feb. 2006. [Online]. Available: <http://dx.doi.org/10.1007/s11276-006-6151-z>
- [143] D. Gunduz and E. Erkip, “Lossless transmission of correlated sources over a multiple access channel with side information,” in *Data Compression Conference (DCC)*, Mar. 2007, pp. 83–92.
- [144] ———, “Correlated sources over an asymmetric multiple access channel with one distortion criterion,” in *Information Sciences and Systems (CISS '07). 41st Annual Conference on*, Mar. 2007, pp. 325–330.
- [145] T. Cover, A. Gamal, and M. Salehi, “Multiple access channels with arbitrarily correlated sources,” *Information Theory, IEEE Transactions on*, vol. 26, no. 6, pp. 648–657, Nov. 1980.
- [146] W. Kang and S. Ulukus, “An outer bound for multiple access channels with correlated sources,” in *Information Sciences and Systems, 40th Annual Conference on*, Mar. 2006, pp. 240–244.
- [147] M. Gastpar and M. Vetterli, “Source-channel communication in sensor networks,” *Information Processing in Sensor Networks*, vol. 2634, no. 553, pp. 648–657, Apr. 2003.

- [148] E. Yeh, “Delay-optimal rate allocation in multiaccess communications: a cross-layer view,” in *Multimedia Signal Processing, IEEE Workshop on*, Dec. 2002, pp. 404–407.
- [149] X. Lin, N. Shroff, and R. Srikant, “A tutorial on cross-layer optimization in wireless networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 8, pp. 1452–1463, Aug. 2006.
- [150] T. Wang, W. Heinzelman, A. Seyedi, and A. Vosoughi, “Maximizing sample rate for distributed source coding over multiple access channels,” in *Communications (ICC), IEEE International Conference on*, Jun. 2011, pp. 1–5.