Theses and Dissertations

8-2014

Optimum Design of Spectral Efficient Green Wireless Communications

Gang Wang
University of Arkansas, Fayetteville

Follow this and additional works at: http://scholarworks.uark.edu/etd

Part of the Systems and Communications Commons

Optimum Design of Spectral Efficient Green Wireless Communications

Optimum Design of Spectral Efficient Green Wireless Communications

A dissertation submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy in Electrical Engineering

By

Gang Wang
Northwestern Polytechnical University
Bachelor of Science in Electrical Engineering, 2005
Beihang University
Master of Science in Electrical Engineering, 2008

August 2014
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

_____

Dr. Jingxian Wu
Dissertation Director

_____   _____

Dr. Roy A. McCann   Dr. Jing Yang
Committee Member   Committee Member

_____

Dr. Qinghua Li
Committee Member

Abstract

This dissertation focuses on the optimum design of spectral efficient green wireless communications. Energy efficiency (EE), which is defined as the inverse of average energy required to successfully deliver one information bit from a source to its destination, and spectral efficiency (SE), which is defined as the average data rate per unit bandwidth, are two fundamental performance metrics of wireless communication systems. We study the optimum designs of a wide range of practical wireless communication systems that can either maximize EE, or SE, or achieve a balanced tradeoff between the two metrics. There are three objectives in this dissertation. First, an accurate frame error rate (FER) expression is developed for practical coded wireless communication systems operating in quasi-static Rayleigh fading channels. The new FER expression enables the accurate modeling of EE and SE for various wireless communication systems. Second, the optimum designs of automatic repeat request (ARQ) and hybrid ARQ (HARQ) systems are performed to by using the EE and SE as design metrics. Specifically, a new metric of normalized EE, which is defined as the EE normalized by the SE, is proposed to achieve a balanced tradeoff between the EE and SE. Third, a robust frequency-domain on-off accumulative transmission (OOAT) scheme has been developed to achieve collision-tolerant media access control (CT-MAC) in a wireless network. The proposed frequency domain OOAT scheme can improve the SE and EE by allowing multiple users to transmit simultaneously over the same frequency bands, and the signal collisions at the receiver can be resolved by using signal processing techniques in the physical layer.

Acknowledgements

Foremost, I would like to thank my advisor Professor Jingxian Wu for his guidance and fully support for me to pursue my interested areas. His unique blend of energy, rigorous, hard-working, and technical knowledge will be an inspiring guidance for my future career. I also thank Professor Roy A. McCann, Professor Jing Yang, and Professor Qinghua Li, for serving as a member of my dissertation committee. Their expertises and visions from several aspects greatly expand my research vision, and help me understand the essence of my research topics.

My thanks extend to all the present and former members of the Wireless Information Network Lab, including Qihao Yang, Ning Sun, Guoqing Zhou, Israel Z. Akingeneye, and Zuoen Wang. I also want to thank the University of Arkansas, including the Electrical Engineering department, the Graduate School, and International Students & Scholars (ISS) office to make my study and life enjoyable here when pursuing the Ph.D. degree.

I express my gratitude to Dr. Genshe Chen at Intelligent Fusion Technology, Inc., Maryland, who provides me consistent help during my internship at the company, also the co-workers, including Dr. Dan Shen, Dr. Zhonghai Wang, and Dr. Bin Jia.

Last, but not the least, I would like to thank my parents, my brother, my sister-in-law, and my parents-in-law for their unconditionally consistent support. I owe it to them. Especially I owe it to my wife, Xue Zhang. Without her sacrifice of time and energy to the family, I will not be able to consistently focus on pursuing the knowledge. Only with her fully support and love, I can devote all my four years to studying. I dedicate this work to my wife.

Contents

List of Papers

- **Chapter 2**, Gang Wang, Jingxian Wu, and Yahong R. Zheng, "An accurate frame error rate approximation of coded diversity systems with non-identical diversity branches," accepted in *IEEE International Conference on Communications, ICC'14*, Jun. 2014.

- **Chapter 3**, Gang Wang, Jingxian Wu, and Yahong R. Zheng, "Cross-layer design of energy efficient coded ARQ systems," *IEEE Global Communications Conference, GLOBECOM*, Dec. 2012.

- **Chapter 4**, Gang Wang, Jingxian Wu, and Yahong R. Zheng, "Optimum energy and spectral efficient transmissions for delay constrained hybrid ARQ systems," accepted in *IEEE Transactions on Vehicular Technology*, May 2014.

- **Chapter 5**, Gang Wang, Jingxian Wu, Guoqing Zhou, and Geoffrey Ye Li, "Collision-tolerant media access control for asynchronous users over frequency selective channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 10, pp. 5162-5171, Sept. 2013.

Chapter 1

Introduction

1.1 Background and Motivation

Energy efficiency (EE) and spectral efficiency (SE) are two most important metrics in wireless communications. EE can be defined as the inverse of average energy required to successfully deliver one information bit from a source to its destination. With the help of EE, the average energy consumption to successfully transmit a information bit can be quantified. SE can be defined as the average data rate per unit bandwidth. It quantifies how efficient the precious spectrum is utilized to transmit information.

Improving SE is always an important goal when developing wireless communication techniques, designing wireless system, and deploying wireless networks. Based on several important factors, in United States (US), the Federal Communications Commission (FCC) and the National Telecommunications and Information Administration (NTIA) administers the spectrum for different types of applications. Because the transmitted signal power in wireless communication degrades at least with square of the transmission distance, the upper spectrum is often limited, currently with 275 GHz [1]. In addition, for wireless communications, the antenna size is directly determined by the inverse of carrier frequency. Therefore, the limitation of antenna size confines the lower spectrum, which is currently 9 KHz [1]. Even the total spectrum resource is limited, they are divided into a lot of much smaller pieces for different types of applications, such as radio channel, television channel, satellite channel, and mobile channel, etc. Nowadays, with growing strong needs of wireless communications,

the spectrum resource is crowded. However, with the limited spectrum resource, high quality-of-service (QoS) is whereas desired to be achieved; for instance, with Wideband Code Division Multiple Access (WCDMA), the channel access is with 5 MHz band, but several users with high quality of voice or video transmissions are desired to be supported. Therefore, SE is important and always a critical metric to designing communication systems.

Besides the SE, EE is gained more and more attention in wireless communications nowadays. One of the reasons is the large global greenhouse gas (GHG) emissions caused by global information and communication technology (ICT). It is estimated that there will be 2.3% of global GHG emitted by global ICT in 2020, and the number will be even increased later [2]. Specifically, we can know how huge of GHG emitted to the air from the respective of cellular towers, which is only a part of global ICT. It shows currently there are more than 5 million towers worldwide to provide cellular commercial network access services [2]. To provide the power to these towers, there are more than 78 Mtons of carbon dioxide emitted to the air [2]. Improving EE is an effective way to reduce the global GHG caused by ICT. In this way, improving EE can be seen as a type of green communications, which is energy-sustainable, resource-saving, and environment-friendly. Another important reason to improve EE is to increase the life time of wireless access applications where power sources are limited, such as the mobile terminals, wireless sensor network (WSN), and satellite communications, etc.

With both needs of large EE and SE; however, EE and SE constructs the most fundamental trade-off in wireless communications. To achieve high SE, it often requires high E_b/N_0, where E_b is the energy per bit and N_0 is noise power spectral density which can be treated as a constant for a designed receiver. Since high E_b means high energy consumption then low EE, therefore, high SE is often achieved with low EE, and vice versa. It constructs the most fundamental tradeoff in wireless

communications. It arises some very interesting questions that how can we quantify this relationship of EE and SE for various communication systems, also how can we achieve the balanced tradeoff between the two metrics. We define the balanced tradeoff point between EE and SE as "With a process of largely increase SE by just sacrificing little EE, SE cannot be increased largely any more by only sacrificing little EE". After the balanced tradeoff point, SE can only be achieved by sacrificing large EE.

Bit error rate (BER) and frame error rate (FER) are two metrics in physical layer to evaluate end-to-end performances of communication systems. It is intuitive that BER and FER can facilitate the analysis of EE and SE; however, FER is more appropriate, since it quantifies the probability that a packet can be successfully transmitted or not, which matches well with the definitions of EE and SE. When formulating EE and SE, it requires the bit is successfully transmitted, which implies the packet is transmitted successfully, since in practical protocols, only the packet with all correct recovered bits is seen as a success transmission. For a packet, even there are plenty of bits are correctly transmitted, the packet is always directly discarded or saved for possible potential use with the retransmissions of the same packet, when there are some bits are wrong, since it is difficult to determine the locations of the error bits. The entire packet rather than the error bits had to be retransmitted again. It then arises a very interesting question that how can the FER be modelled to facilitate the communication system performance analysis and cross-layer design. It is very hard, even not impossible, to model the FER for general coded system, where there are practical channel coding schemes employed in the wireless communications. There are many papers talked about the FER for coded system; however, either the model is only fit for a specialized coding scheme [3] or some parameters need to be re-evaluated each time if any system parameter is changed [4]– [6], such as packet length, modulation scheme, or coding rate, etc. The system parameters adaptation

3

however widely utilized in the cross-layer design to fit for different requirement of QoS with varied wireless channel environments. The general FER model for different combinations of frame length, modulation scheme, and coding scheme remain open but urgently needed.

With a general FER model, the EE and SE can be directly expressed as a function of several system parameters. EE and SE can then be individually optimized. Due to the tradeoff between EE and SE, the optimization scheme of EE and SE could be very different. Therefore, how to effectively obtain the balanced tradeoff point between EE and SE, for practical systems, where expressions of EE and SE could be very complex? This is exactly what we are going to explore in this dissertation.

Besides the analysis of EE and SE, we should also resort to techniques to improve EE and SE simultaneously, since they share some commons. For instance, to improve EE, we want to reduce the energy consumption to successfully transmit a bit. Therefore, if interferences from other sources, such as multi-user interference and co-channel interference, can be reduced, the energy required for a bit can also be reduced, thus EE is increased. Besides, for a packet, if the first transmission is failed, we should utilize the received packet for further detection, rather than directly discard it. By exploiting every received packet, the energy required for retransmission of the packet can be reduced, also the number of retransmission times of the packet. Therefore, less energy will be needed for the total energy consumption to successfully transmit the packet, EE can then be increased. The techniques to reduce interferences and exploiting every received packet can also reduce the transmission time to successfully transmit a bit with unit bandwidth, thus increasing the SE. Therefore, for a wireless network, it would be very interesting to design transmission techniques, powerful digital signal processing (DSP) detections, and multi-user coordination techniques to improve both EE and SE.

1.2 Objectives

The objectives of this dissertation are summarized here.

I. Building the general FER system model for various coded system, considering practical system parameters, such as the frame length, modulation schemes, and coding rate, etc. Besides the single-input single-output (SISO) system, the FER model for the receiver diversity systems are also derived, considering there exists several retransmissions of a packet in practise with maximum ration combining (MRC).

II. With the built FER system model, the EE of automatic repeat-request (ARQ) system and hybrid ARQ system are formulated and analyzed. For ARQ system, the packet length and transmission power are jointly optimized to maximize EE. For HARQ system, the transmission power in each round to retransmit a packet is optimized to maximize EE.

III. Similar with the analysis of EE for ARQ system and HARQ system, the SE are also maximized for the two systems. After analysing both EE and SE for the two systems, the balanced tradeoff between EE and SE are then analyzed.

IV. Besides the performance analysis of practical wireless communication systems, a radio resource management scheme is also proposed to improve both network EE and network SE, by coordinating the frequency resources to each node, also powerful DSP detection techniques.

1.3 Dissertation Outline

The outline of the rest of the dissertation is given as follows.

Chapter 2: In this chapter, a new FER approximation is proposed for wireless communication systems with receiver diversity and independent but non-identically distributed (i.n.d.) diversity

branches. Receiver diversity exists in many practical systems, such as single-input multiple-output (SIMO) systems [21], multi-node relay systems [22], and hybrid ARQ systems [23], where the receiver can coherently combine multiple copies of the transmitted message. Signals at different diversity branches usually experience non-identically distributed fading due to the effects of pathloss and differences in transmission power. The statistical properties of the SNR for receiver diversity systems undergoing i.n.d. quasi-static Rayleigh fading are derived with the help of the moment generating function (MGF). The results are used to develop the FER approximation with the threshold-based method.

Different from existing threshold-based methods that need to calculate a new threshold value for each frame length, the proposed FER analysis models the threshold value as a linear function of the frame length in the log-domain, with the slope and intercept of the linear function determined by the modulation and coding schemes. It is discovered that the threshold value is independent of the actual fading distributions or the number of diversity branches. Enabled by the log-domain linear threshold modelling, the new FER is expressed as an explicit function of various parameters related to modulation, coding, frame length, SNR, number of diversity branches, and power distribution across the diversity branches. Such a parametric FER approximation summarizes the complex operations in the PHY layer into a few adjustable design parameters, and this is especially useful for cross-layer designs where joint optimization can be performed by flexibly adjusting the PHY layer parameters based on the FER requirement of upper layer protocols. It should be noted that none of the existing FER approximations provide such parametric flexibility that is essential for cross-layer designs. The newly proposed FER approximation is general enough to include many previous works as special cases, and simulation results demonstrate that the analytical FER approximation can accurately predict the performance of a wide range of practical wireless communication systems.

Chapter 3: In this chapter and our extended work [19], we studied the optimum energy- and spectral-efficient designs of type-I ARQ systems, where a package is retransmitted if it cannot be recovered at the receiver. Although many varieties of ARQ and HARQ systems exist that offer different tradeoffs between the performance and complexity [12,13,17,18,24], type-I ARQ has a wide range of applications due to its simplicity that is especially important for low cost and low energy communication systems. Three optimization schemes are considered for type-I ARQ in flat fading channels: the first scheme aims to minimize the energy per information bit E_t, which is the energy required to successfully deliver one information bit from a transmitter to a receiver; the second scheme minimizes the normalized energy per information bit $E_m = \frac{E_t}{\eta_s}$, where η_s is the SE in bps/Hz; and the third scheme minimizes E_t under the constraint of a minimum SE, $\eta_s \geq \eta_{th}$, with η_{th} being a constant threshold. The first scheme minimizes the overall energy consumption at the cost of the SE. The second metric E_m can be minimized by either reducing E_t or increasing η_s, thus providing one possible tradeoff between EE and SE. The third optimization scheme allows us to flexibly adjust the EE-SE tradeoff based on specific system requirements, and it includes the E_t or E_m minimization as special cases.

The optimum designs are performed by jointly optimizing the transmission energy in the PHY layer and the frame length in the MAC layer. The system designs incorporate a large number of practical system parameters, such as the efficiency of the power amplifier, the power consumption of digital hardware, data rate, coding and modulation, frame length and protocol overhead, frame error rate (FER), and frame retransmission, etc. To quantify the impacts of transmission energy and frame length, a new log-domain threshold approximation to FER in Rayleigh fading channels is proposed, which enables explicit analytical solutions to the three optimization schemes.

Another important contribution in this chapter is that the fundamental EE-SE tradeoff for type-I ARQ systems are identified through theoretical analysis of the third optimization scheme. The result corroborates that the minimum energy per information bit, which is inversely proportional to the EE, is quasiconvex in the SE, or equivalently, the EE is quasiconcave in the SE. This result agrees with the EE-SE tradeoff relationship presented in [15] for the downlink OFDMA system. Due to the quasiconvexity, the $E_{\rm t}$–$\eta_{\rm s}$ curve can be divided into two regions, one with a negative slope and the other one with a positive slope. The operation parameters corresponding to the negative slope region of the $E_{\rm t}$–$\eta_{\rm s}$ curve are not desirable for practical operations, because there always exist parameters that can outperform any point in this region in terms of both EE and SE. The non-negative slope region of the $E_{\rm t}$–$\eta_{\rm s}$ curve reveals the fundamental EE-SE tradeoff in ARQ systems. It provides a general optimization framework that enables us to obtain optimum system designs by flexibly adjusting the EE-SE tradeoffs based on the various requirements of practical systems.

Chapter 4: In this chapter, we studied the optimum transmission of coded HARQ systems with Chase combining, where the signals from all transmission attempts of the same packet are combined coherently at the receiver during the detection process. The optimum design is performed by identifying the optimum energy distribution, *i.e.*, the sequence of transmission energy that should be used at different transmission attempts, such that the overall energy required to successfully deliver a bit from a source to its destination is minimized. A large number of practical system parameters, such as the efficiency of the energy amplifier, the energy consumption of digital hardware, data rate, modulation and coding schemes, frame length, FER, and the protocol overhead, etc., are considered during the design. With the help of a new backward sequential method, the optimum transmission energy is expressed as closed-form expressions of all the practical system parameters. Both simulation and analytical results demonstrate that the proposed energy efficient system design can achieve

significant energy savings over conventional systems.

Chapter 5: In this chapter, we proposed a new frequency-domain on-off accumulative transmission (OOAT) scheme that achieves collision-tolerant MAC with MUD in the PHY layer. The proposed scheme is extended from a time-domain OOAT scheme in our previous work [25], which can only operate in frequency-flat fading. In the frequency-domain OOAT, the channel is divided into multiple orthogonal sub-channels with the help of OFDM. Different from conventional OFDM, each symbol is transmitted over multiple sub-channels in our scheme. Consequently, the proposed scheme can not only deal with frequency-selectivity of wideband wireless channels as OFDM, but also exploit frequency diversity since each symbol is spread to several sub-channels.

The frequency-domain OOAT scheme converts the relative transmission delays among the users in the time domain into phase shifts in the frequency domain, such that the sub-channels from different users are perfectly aligned. This allows us to carefully plan the on-off patterns employed by different users to minimize the number of users colliding on each sub-channel. Based on the OOAT signal structure, optimum and sub-optimum MUDs are proposed, and an analytical matched filter bound is derived to quantify the performance of the proposed scheme.

Chapter 6: Conclusion remarks are drawn in this chapter. The major contributions of this dissertation are summarized and some future research topics are listed.

1.4 References

[1] *Radio Spectrum Allocation*, [Online]. Available: http://www.fcc.gov/encyclopedia/radio-spectrum-allocation.

[2] Available: http://www.greentouch.org.

[3] H. Xiao and A. H. Banihashemi, "Estimation of bit and frame error rates of finite-length low-density parity-check codes on binary symmetric channels," *IEEE Transactions on Communications*, vol. 55, no. 12, pp. 2234 - 2239, Dec. 2007.

[4] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," *42nd Annual Conference on Information Sciences and Systems, CISS 2008*, pp. 577 - 581, 2008.

[5] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "Threshold-based frame error rate analysis of MIMO systems over quasistatic fading channels," *Electronics Letters*, vol. 45, no. 4, Feb. 2009.

[6] T. Liu, L. Song, and B. Jiao, "Threshold-based frame error rate analysis of incremental hybrid relay selection scheme," in *Proc. IEEE Global Telecommun. Conf. Globecom'10*, Dec. 2010.

[7] G. Y. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorials, survey, and open issues," *IEEE Trans. Wireless Communications*, vol. 18, pp. 28 - 35, 2011.

[8] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Communications*, vol. 4, pp. 2349 - 2360, Sept. 2005.

[9] F. M. Costa and H. Ochiai, "Energy-efficient physical layer design for wireless sensor network links," *IEEE Int. Conf. Communications (ICC)*, Kyoto, Japan, Jun. 2011. pp. 1-5.

[10] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Trans. Networking*, vol. 12, pp.493 - 506, June 2004.

[11] A. E. Gamal, C. Nair, B. Prabhakar, E. Uysal-Biyikoglu, and S. Zahedi, "Energy-efficiency scheduling of packet transmissions over wireless networks," *IEEE INFOCOM 2002*, vol. 3, June 2002. pp. 1773 - 1782.

[12] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "Optimal power assignment for minimizing the average total transmission power in hybrid-ARQ Rayleigh fading links," *IEEE Trans. Communications*, vol. 59, no. 7, pp.1867 - 1877, July 2011.

[13] T. Chaitanya and E. Larsson, "Outage-optimal power allocation for hybrid ARQ with incremental redundancy," *IEEE Trans. Wireless Communications*, vol. 10, pp.2069 - 2074, Jul. 2011.

[14] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Information Theory*, vol. 48, pp. 1319 - 1343, June 2002.

[15] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Trans. Wireless Communications*, vol. 10, pp.3874 - 3886, Nov. 2011.

[16] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental tradeoffs on green wireless networks," *IEEE Communications Magazine*, vol. 49, pp.30-37, June 2011.

[17] D. Tuninetti, "On the benefits of partial channel state information for repetition protocols in block fading channels," *IEEE Trans. Information Theory*, vol. 57, pp.5036 - 5053, Aug. 2011.

[18] G. Wang, J. Wu, and Y. R. Zheng, "Cross-layer design of energy efficient coded ARQ systems," in *Proc. IEEE Global Telecommun. Conf. Globecom'12*, Anaheim, CA, Dec. 2012.

[19] J. Wu, G. Wang, and Y. R. Zheng, "Energy efficiency and spectral efficiency tradeoff in type-I ARQ systems," IEEE Journal on Selected Areas in Communications, vol. 32, no. 2, pp. 356 - 366, Feb. 2014

[20] J. Wu, G. Wang, and Y. R. Zheng, "Energy and spectral efficient transmissions of coded ARQ systems," in *Proc. IEEE International Conference on Communications, ICC'13*, June 2013.

[21] J. Wu and C. Xiao, "Optimal diversity combining based on linear estimation of Rician fading channels," *IEEE Transactions on Communications*, vol. 56, pp. 1612-1615, Oct. 2008.

[22] J. Wu and Y. R. Zheng, "Optimum multi-hop transmission strategies for energy constrained wireless sensor networks," *IEEE International Conf. on Communications, ICC'12*, 2012.

[23] G. Wang, J. Wu, and Y. R. Zheng, "Optimum energy efficient communications for hybrid ARQ systems," accepted in *IEEE Global Telecommun. Conf. Globecom'13*, 2013.

[24] K. Nguyen, L. Rasmussen, A. Guillen i Fabregas, and N. Letzepis, "MIMO ARQ with multibit feedback: outage analysis," *IEEE Trans. Information Theory*, vol. 58, pp.765 - 779, Feb. 2012.

[25] J. Wu and Y. Li, "Collision-tolerant media access control: on-off accumulative transmission," submitted to *IEEE Trans. Wireless Commun.*, July 2011.

[26] C. Comaniciu, N. B. Mandayam, and V. H. Poor, *Wireless Networks Multiuser Detection in Cross-Layer Design*, Springer, April 2005.

[27] B. Lu, X. Wang, and J. Zhang, "Throughput of CDMA data networks with multiuser detection, ARQ, and packet combining," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1576 - 1589, Sept. 2004.

[28] R. Merz, J. Widmer, J.-Y. Le Boudec, and B. Radunovic, "A joint PHY/MAC architecture for low-radiated power TH-UWB wireless ad hoc networks," *Wireless Commun. Mobile Computing*, vol. 5, pp. 567-580, 2005.

[29] J. Tao, J. Wu, and Y. Zheng, "Reliability-based turbo detection," *IEEE Trans. Wireless Commun.*, vol. 10, pp. 2352-2361, July 2011.

[30] P. Casari, M. Levorato, and M. Zorzi, "On the implications of layered space-time multiuser detection on the design of MAC protocols for ad hoc networks," in *Proc. Intern. Symp. Personal, Indoor Mobile Radio Commun. PIMRC'05*, vol. 2, pp. 1354 - 1360, 2005.

[31] P. Casari, M. Levorato, and M. Zorzi, "MAC/PHY cross-layer design of MIMO ad hoc networks with layered multiuser detection," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4596 - 4607, Nov. 2008.

[32] G. Mergen and L. Tong, "Receiver controlled medium access in multihop ad hoc networks with multipacket reception," in *Proc. IEEE Military Commun. Conf. MILCOM 2001*, vol. 2, pp. 1014 - 1018, 2001.

[33] L. Tong, V. Naware, and P. Venkitasubramaniam, "Signal processing in random access," *IEEE Sig. processing Mag.*, vol. 21, pp. 29-39, Sept. 2004.

[34] Q. Zhao and L. Tong, "A dynamic queue protocol for multiaccess wireless networks with multipacket reception," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 2221 - 2231, Nov. 2004.

[35] R. Samano-Robles, M. Ghogho, and D. C. Mclernon, "An infinite user model for random access protocols assisted by multipacket reception and retransmission diversity," in *Proc. IEEE Sig. Processing Advances Wireless Commun. SPAWC 2008*, pp. 111 - 115, 2008.

<center>**Chapter 2**</center>

<center>**An Accurate Frame Error Rate Approximation of Coded Diversity Systems with**</center>

<center>**Non-identical Diversity Branches**</center>

<center>Gang Wang, Jingxian Wu, and Yahong Rosa Zheng</center>

2.1 Abstract

This paper presents an accurate approximation of the frame error rate (FER) of coded wireless communication systems with receiver diversity, such as single-input multiple-output (SIMO) systems with maximum ratio combining (MRC) or hybrid automatic repeat request (HARQ) systems with Chase combining. The signals at different diversity branches experience independent but non-identically distributed Rayleigh fading. The FER approximation is obtained with a threshold-based method. Specifically, the threshold value, which is critical to the FER approximation accuracy, is modeled as a linear function of the frame length in the log-domain, with the slope and intercept of the linear function determined by the underlying modulation and channel coding schemes. The analytical FER approximation is expressed as an explicit function of parameters related to modulation, coding, frame length, number of diversity branches, and the power distribution across the diversity branches. Such an FER approximation summarizes the complex physical layer operations into a few parameters, and it provides the parametric flexibility that is not available in most existing FER approximations. Simulation results show that the proposed FER approximation can accurately predict the FER performance of a wide range of receiver diversity systems.

2.2 Introduction

Bit error rate (BER) and frame error rate (FER) are two important metrics used to measure the quality of service (QoS) of communication systems. BER provides a measurement of the end-to-end communication quality in the physical (PHY) layer. FER is also a PHY layer metric, but it can be used to facilitate the design across multiple protocol layers in a wireless network. For example, a frame error after PHY layer channel decoding can be detected with cyclic redundancy check (CRC) or parity check in the media access control layer (MAC), and frames with unrecoverable errors will be retransmitted if automatic repeat request (ARQ) is employed. The knowledge of the FER at the upper layers can thus greatly facilitate the design of wireless networks, especially for cross-layer designs that aim to improve the system performance by jointly optimizing the system parameters across two or more protocol layers [1] and [13].

Recently there are growing interests in FER analysis for wireless communication systems due to the increasing importance of cross-layer designs. It is extremely difficult, if not impossible, to obtain the exact analytical FER expression, especially for coded systems. Therefore almost all the works in the literature focus on the analytical approximations of the actual FER. In [3], an FER approximation is provided for systems with low-density parity-check (LDPC) codes over binary symmetric channel (BSC) by enumerating a large number of error patterns. A threshold-based FER approximation for a system with quasi-static Rayleigh fading is proposed in [4]. The threshold-based method is developed by using an assumption that the FER in an additive white Gaussian noise (AWGN) channel can be approximated by either 1 or 0 depending on whether the signal-to-noise ratio (SNR) is less than or greater than a waterfall threshold. The results in [4] are extended to diversity systems with independent and identically distributed (i.i.d.) Rayleigh fading channels in [5] and to 2-hop relay networks with incremental hybrid relay selection in [6]. The FER approximations in [4]– [6]

are semi-analytical in that the threshold values of coded systems need to be evaluated through FER simulations in AWGN channels. None of the above FER approximations can *explicitly* quantify many important design parameters, such as frame length, modulation, coding, and the diversity orders, etc.

In this paper, a new FER approximation is proposed for wireless communication systems with receiver diversity and independent but non-identically distributed (i.n.d.) diversity branches. Receiver diversity exists in many practical systems, such as single-input multiple-output (SIMO) systems [7], multi-node relay systems [8], and hybrid ARQ systems [9], where the receiver can coherently combine multiple copies of the transmitted message. Signals at different diversity branches usually experience non-identically distributed fading due to the effects of pathloss and differences in transmission power. The statistical properties of the SNR for receiver diversity systems undergoing i.n.d. quasi-static Rayleigh fading are derived with the help of the moment generating function (MGF). The results are used to develop the FER approximation with the threshold-based method.

Different from existing threshold-based methods that need to calculate a new threshold value for each frame length, the proposed FER analysis models the threshold value as a linear function of the frame length in the log-domain, with the slope and intercept of the linear function determined by the modulation and coding schemes. It is discovered that the threshold value is independent of the actual fading distributions or the number of diversity branches. Enabled by the log-domain linear threshold modelling, the new FER is expressed as an explicit function of various parameters related to modulation, coding, frame length, SNR, number of diversity branches, and power distribution across the diversity branches. Such a parametric FER approximation summarizes the complex operations in the PHY layer into a few adjustable design parameters, and this is especially useful for cross-layer designs where joint optimization can be performed by flexibly adjusting the PHY layer parameters

based on the FER requirement of upper layer protocols. It should be noted that none of the existing FER approximations provide such parametric flexibility that is essential for cross-layer designs. The newly proposed FER approximation is general enough to include many previous works as special cases, and simulation results demonstrate that the analytical FER approximation can accurately predict the performance of a wide range of practical wireless communication systems.

2.3   System Model

Consider a wireless communication system with receiver diversity. The receiver has N copies of the transmitted signal, each goes through i.n.d. Rayleigh fading.

At the transmitter, the binary information bits are divided into frames with L bits per frame. The bits are encoded with a rate-r channel code, and the coded bits are modulated with M-ary modulation. The number of symbols per frames is thus $K_s = \frac{L}{r \log_2 M}$, where L is chosen in a way such that K_s is an integer. Denote a frame as $\mathbf{x} = [x_1, \cdots, x_{K_s}]^T \in \mathcal{S}^{K_s \times 1}$, where $\mathcal{S}$ is the modulation alphabet set, and $\mathbf{A}^T$ denote the transpose of the matrix $\mathbf{A}$.

The signals are transmitted through an N-branch diversity channel, with each branch experiencing i.n.d. Rayleigh fading. The N-dimensional equivalent discrete-time signals observed at the receiver is

$$\mathbf{y}_k = \mathbf{h}x_k + \mathbf{z}_k, \text{ for } k = 1, \cdots, K_s \tag{2.1}$$

where $\mathbf{y}_k = [y_k(1), \cdots, y_k(N)]^T \in \mathcal{C}^{N \times 1}$ is the received signal vector corresponding to the k-th transmitted symbol in a frame, $\mathbf{z}_k = [z_k(1), \cdots, z_k(N)]^T \in \mathcal{C}^{N \times 1}$ is the additive white Gaussian noise vector with zero mean and correlation matrix $N_0 \mathbf{I}_N$, with $\mathbf{I}_N$ being a size-N identity matrix, and $\mathbf{h} = [h(1), \cdots, h(N)] \in \mathcal{C}^{N \times 1}$ are the fading coefficient vector. The channel fading is assumed to be quasi-static, *i.e.*, it is constant within one frame, but changes from frame to frame.

For a system experiencing Rayleigh fading, $\mathbf{h}$ is a zero-mean complex Gaussian distributed random vector with the auto-covariance matrix being a diagonal matrix as $\mathbf{R}_h = \text{diag}([\sigma_1^2, \cdots, \sigma_N^2]^T) \in \mathcal{R}^{N \times N}$. The fading coefficients are normalized to unit energy with $\text{trace}(\mathbf{R}_h) = \sum_{n=1}^{N} \sigma_n^2 = 1$.

The receiver performs maximal ratio combining (MRC) over the signals. The decision variable for the k-th symbol at the output of the MRC is

$$\alpha_k = \mathbf{h}^H \mathbf{y}_k = \mathbf{h}^H \mathbf{h} x_k + \mathbf{h}^H \mathbf{z_k}. \tag{2.2}$$

The equivalent signal-to-noise ratio (SNR) at the output of the MRC detector is thus

$$\gamma = r \log_2 M \gamma_0 \cdot \mathbf{h}^H \mathbf{h} = \gamma_0 \mu_M \cdot \sum_{n=1}^{N} |h(n)|^2, \tag{2.3}$$

where $\mu_M = r \log_2 M$ is related to the coding rate and modulation level, $\gamma_0 = \frac{E_b}{N_0}$ with E_b being the energy per uncoded information bit. Since $\{h(n)\}_{n=1}^{N}$ are not identically distributed, γ cannot be modelled as a χ^2 distributed random variable. The statistical properties of the SNR will be discussed in the next Section.

2.4 FER Analysis

In this section, an accurate threshold-based approximation of the FER for a system with receiver diversity and i.n.d Rayleigh fading channel is derived with the help of the statistical properties of the SNR at the receiver.

2.4.1 Statistical Properties of the Receiver SNR

The probability density function (pdf) of the SNR, γ, is derived in this section with the help of its MGF.

Define $\gamma_n = \gamma_0 \mu_M |h(n)|^2$ as the SNR on the n-th branch. Since $h(n)$ is a zero-mean complex

Gaussian random variable with variance σ_n^2, the MGF of γ_n can be written as

$$\Phi_n(s) = \mathbb{E}\left[\exp(-s\gamma_n)\right] = \frac{1}{1+s\lambda_n}, \tag{2.4}$$

where $\lambda_n = \mu_M \gamma_0 \sigma_n^2$.

Based on the assumption of independence among γ_n, the MGF of γ at the output of the MRC detector, $\Phi(s) = \mathbb{E}\left[\exp(-s\gamma)\right]$, can be written as

$$\Phi(s) = \prod_{n=1}^{N} \frac{1}{1+s\lambda_n}. \tag{2.5}$$

Even though the channel experiences i.n.d. fading, a subset of the fading coefficients might share the same variance. Without loss of generality, assume there are $V \leq N$ distinct values of λ_n. Divide the index set, $\mathcal{N} = \{1, \cdots, N\}$, into V subsets, $\mathcal{N}_v \subseteq \mathcal{N}$, for $v = 1, \cdots, V$, such that $\lambda_m = \lambda_n$ if and only if m, n are in the same subset. Denote the cardinality of the v-th subset as n_v, and define $\bar{\lambda}_v = \lambda_m, \forall m \in \mathcal{N}_v$.

Based on the subset notation given above, the MGF can be alternatively represented by

$$\Phi(s) = \prod_{v=1}^{V} \frac{1}{\left(1+s\bar{\lambda}_v\right)^{n_v}}. \tag{2.6}$$

Performing partial fraction expansion over $\Phi(s)$ given in (2.6) yields

$$\Phi(s) = \sum_{v=1}^{V} \sum_{q=0}^{n_v-1} \frac{\beta_q^{(v)}}{q!} \frac{\bar{\lambda}_v^{-q}}{(1+s\bar{\lambda}_v)^{n_v-q}}, \tag{2.7}$$

where $\beta_q^{(v)}$ is

$$\beta_q^{(v)} = \frac{\partial^q}{\partial s^q}\left[\prod_{k=1,k\neq v}^{V} \frac{1}{(1+s\bar{\lambda}_k)^{n_k}}\right]\Bigg|_{s=-\bar{\lambda}_v^{-1}}. \tag{2.8}$$

The pdf of γ can then be obtained by performing the inverse Laplace transform over (2.7), and the result is

$$f_\gamma(x) = \sum_{v=1}^{V} \sum_{q=0}^{n_v-1} \frac{\beta_q^{(v)}}{q!} \frac{x^{n_v-q-1}}{\bar{\lambda}_v^{n_v}} \frac{\exp\left(-\frac{x}{\bar{\lambda}_v}\right)}{\Gamma(n_v-q)}, \tag{2.9}$$

19

where $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$ is the Gamma function.

The cumulative distribution function (CDF) of the SNR, $F_\gamma(x) = \int_0^x f_\gamma(u)du$, is then

$$F_\gamma(x) = \sum_{v=1}^{V} \sum_{q=0}^{n_v-1} \frac{\beta_q^{(v)}}{q!} \frac{\bar{\lambda}_v^{-q}}{\Gamma(n_v-q)} \ \psi\left(n_v - q, \frac{x}{\bar{\lambda}_v}\right), \tag{2.10}$$

where $\psi(s,x)$ is low incomplete gamma function with $\psi(s,x) = \int_0^x t^{s-1}e^{-t}dt$.

2.4.2 Threshold-based FER Approximation

The exact FER of a receiver diversity system experiencing i.n.d quasi-static Rayleigh fading channel can be expressed as

$$\mathrm{FER} = \int_0^{+\infty} \mathrm{FER}_G(\gamma) \times f_\gamma(x)dx, \tag{2.11}$$

where $\mathrm{FER}_G(\gamma)$ is the FER for a system operating in an AWGN channel with SNR γ. The exact evaluation of FER requires the knowledge of $\mathrm{FER}_G(\gamma)$, which depends on a number of factors, such as modulation, coding, and frame length, etc. For most practical systems, $\mathrm{FER}_G(\gamma)$ eludes an analytical expression because of the complex channel coding employed in the system.

An accurate approximation of the FER can be obtained by using a threshold-based method [4], which was originally designed for a single-input and single-output system with quasi-static Rayleigh fading. The basic idea of the threshold-based method is that there exists a threshold γ_ω, such that $\mathrm{FER}_G(\gamma) \simeq 1$ when $\gamma \leq \gamma_\omega$ and $\mathrm{FER}_G(\gamma) \simeq 0$ when $\gamma > \gamma_\omega$. With the threshold γ_ω, the FER of a diversity receiver with i.n.d. quasi-static Rayleigh fading can be approximated by

$$\mathrm{FER} = \int_0^{\gamma_\omega} \mathrm{FER}_G(\gamma) \times f_\gamma(x)dx + \int_{\gamma_\omega}^{+\infty} \mathrm{FER}_G(\gamma) \times f_\gamma(x)dx$$
$$\simeq \int_0^{\gamma_\omega} f_\gamma(x)dx. \tag{2.12}$$

Combining (2.10) and (2.12) yields

$$\mathrm{FER} \simeq \sum_{v=1}^{V} \sum_{q=0}^{n_v-1} \frac{\beta_q^{(v)}}{q!} \frac{\bar{\lambda}_v^{-q}}{\Gamma(n_v-q)} \ \psi\left(n_v - q, \frac{\gamma_\omega}{\bar{\lambda}_v}\right), \tag{2.13}$$

20

which is the CDF of the instantaneous SNR γ evaluated at the threshold γ_ω. The value of γ_ω is critical to the evaluation of the FER. The calculation of γ_ω for different modulation and coding schemes is discussed in the next subsection.

We next consider three special cases: 1) $\sigma_i^2 \neq \sigma_j^2$, $\forall i \neq j$; 2) $\sigma_1^2 = \cdots = \sigma_N^2$; and 3) $N = 1$.

2.4.2.1 $\sigma_i^2 \neq \sigma_j^2$, $\forall i \neq j$

This corresponds to the special case that no two channels share the same power. This case is quite common in practical systems. For example, in an HARQ system with Chase combining, the energy used at each retransmission round is usually different from the others in order to minimize the overall energy consumption [9]– [10]. In a SIMO system, the distances between the transmit antenna to the various receive antennas are usually different, which introduces different SNR at different channels. In this case, $V = N$, $n_v = 1$, and $\bar{\lambda}_n = \lambda_n$, for $v = 1, \cdots, N$. With unique SNR on each diversity branch, the MGF in (2.7) can be simplified to

$$\Phi(s) = \sum_{n=1}^{N} \left(\prod_{k=1,k\neq n}^{N} \frac{\lambda_n}{\lambda_n - \lambda_k} \right) \frac{1}{1 + s\lambda_n}. \tag{2.14}$$

The pdf of γ can then be obtained by performing the inverse Laplace transform of $\Phi(s)$ over (2.14), and the result is

$$f_\gamma(x) = \sum_{n=1}^{N} \left(\prod_{k=1,k\neq n}^{N} \frac{\lambda_n}{\lambda_n - \lambda_k} \right) \frac{1}{\lambda_n} \exp\left(-\frac{x}{\lambda_n} \right). \tag{2.15}$$

The threshold-based FER approximation can then be obtained by combining (2.12) and (2.15), and the result is

$$\text{FER} \simeq \sum_{n=1}^{N} \left(\prod_{k=1,k\neq n}^{N} \frac{\lambda_n}{\lambda_n - \lambda_k} \right) \left[1 - \exp\left(-\frac{\gamma_\omega}{\lambda_n} \right) \right]. \tag{2.16}$$

2.4.2.2 $\sigma_1^2 = \cdots = \sigma_N^2 = \frac{1}{N}$

This corresponds to the case that the N diversity branches are i.i.d., which is a common assumption used in many SIMO systems, such as in [5] and [11]. In this case, $V = 1$, $n_v = N$, and denote $\lambda = \lambda_n$ for $n = 1, 2, \cdots, N$. The pdf can then be written as

$$f_\gamma(x) = \frac{x^{N-1}}{\Gamma(N)\lambda^N} \exp\left(-\frac{x}{\lambda}\right). \tag{2.17}$$

The SNR follows a χ^2 distribution with $2N$-degree of freedom.

The threshold-based FER approximation can then be obtained by combining (2.12) and (2.17),

$$\mathrm{FER} \simeq \int_0^{\gamma_\omega} \frac{x^{N-1}}{\Gamma(N)\lambda^N} \exp\left(-\frac{x}{\lambda}\right) dx, \tag{2.18}$$

which simplifies to the result in [5].

2.4.2.3 $N = 1$

This corresponds to a non-diversity system. The FER approximation can be directly obtained by setting $N = 1$ in (2.18), and the result is

$$\mathrm{FER} \simeq \int_0^{\gamma_\omega} \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right) dx = 1 - \exp\left(-\frac{\gamma_\omega}{\lambda}\right), \tag{2.19}$$

which matches the result in [4].

2.4.3 A Log-Domain Linear Approximation of the Threshold

The threshold, γ_ω, depends on a number of factors, such as the frame length, the modulation scheme, and the channel coding employed at the receiver. In [4], the value of γ_ω is evaluated by integrating over the difference between the exact and approximated FERs, and then force the integral to be zero. The result is expressed as [4, eqn. (20)]

$$\gamma_\omega = \left(\int_0^\infty \frac{\mathrm{FER}_G(\gamma)}{\gamma^2} d\gamma\right)^{-1}. \tag{2.20}$$

22

For practical systems with channel code, the result in (2.20) usually needs to be evaluated through simulations because there is no analytical expression of $\text{FER}_G(\gamma)$ for such systems. Therefore the change in any system parameter other than the average SNR will require a revaluation of γ_ω. Therefore, it does not build an explicit relationship between γ_ω and the various system design parameters, and such relationship is usually critical for the design and optimization of practical systems.

In this subsection, a log-domain linear approximation of the threshold is proposed to build an explicit connection between FER and the various system parameters. To gain insights about the value of γ_ω, simulations are performed to obtain the simulated FER under different combinations of modulation, channel coding, frame length, diversity order, and channel power distributions. The value of γ_ω is then optimized by matching the approximate FER in (2.13) to the simulated FER with the least squares (LS) method.

Fig. 2.1 shows the values of γ_ω as a function of the log-domain frame length, $\log L$, for systems with different combinations of modulations and channel codings. There are $N = 4$ i.n.d. Rayleigh fading branches, with the ratio of the channel variance being $\sigma_1^2 : \sigma_2^2 : \sigma_3^2 : \sigma_4^2 = 1 : 3 : 5 : 7$, or, $\sigma_n^2 = \frac{2n-1}{N^2}$, which satisfies $\sum_{n=1}^{N} \sigma_n^2 = 1$. The modulation schemes are QPSK, 16QAM, and 64QAM, respectively. Code 1 is a popular convolutional code with coding rate $r = \frac{1}{2}$, generator polynomial $[171, 133]_8$, and a constraint length 7, commonly utilized in space communications; code 2 is a convolutional code with coding rate $r = \frac{1}{3}$, generator polynomial $[557, 663, 711]_8$, and a constraint length 9, which is adopted in wideband code division multiple access (W-CDMA) standard. For both codes, soft Viterbi decoding are used. From the figure, it is interesting to note that γ_ω can be accurately approximated as a linear function of $\log L$, with the slope and intercept determined by the different modulation and coding scheme.
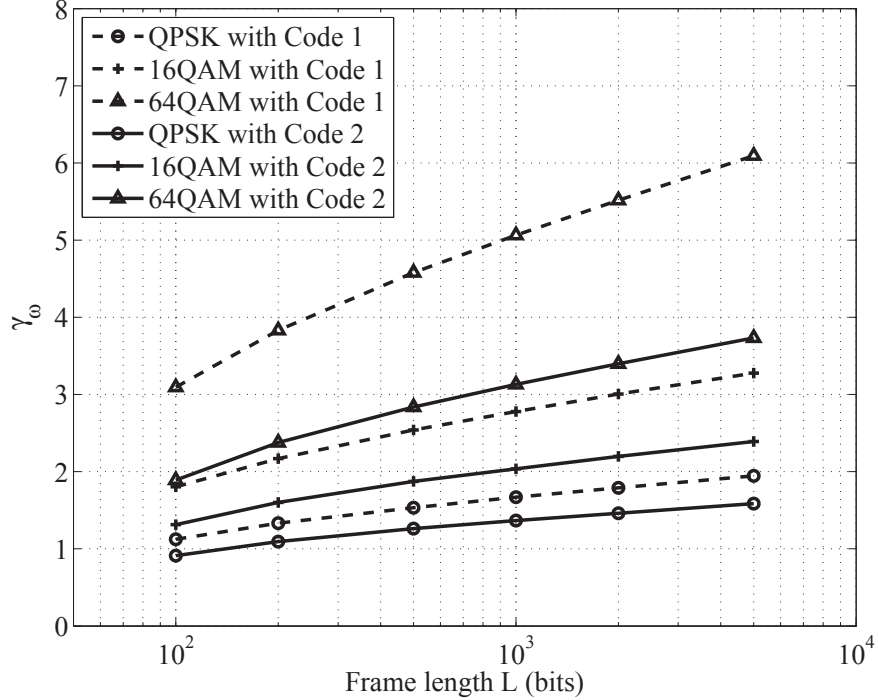
23

Figure 2.1: γ_ω for different combination of modulation scheme, coding scheme, and frame length.

Therefore, we propose to model γ_ω as

$$\gamma_\omega \simeq k_M \log L + b_M, \tag{2.21}$$

where k_M and b_M are the slope and intercept determined by the modulation scheme and the actual channel code. The values of k_M and b_M can be obtained by applying the LS curve fitting to the results in Fig. 2.1. In additional to the convolutional codes shown in Fig. 2.1, we also evaluated the FER of systems with turbo codes and low-density parity check (LDPC) codes, and observed a similar trend. The k_M and b_M for various coding and modulation combinations are given in Table 2.1. The turbo code is a rate $r = \frac{1}{3}$ code with generator polynomials $[1, 5/7, 5/7]_8$ and a block interleaver. The turbo decode is performed with six iterations [13]. The LDPC code is a variable coding rate irregular code with maximum iteration decoding number set to 20 [17]. The frame length of the LDPC code is less than 3819 bits during the evaluation.

The results in Fig. 2.1 are obtained for systems with $N = 4$ i.n.d. fading branches. More

24

Table 2.1: k_M and b_M in (2.21)

	QPSK		16QAM		64QAM	
	k_M	b_M	k_M	b_M	k_M	b_M
Code 1	0.209	0.207	0.372	0.170	0.667	-0.131
Code 2	0.169	0.180	0.270	0.140	0.462	-0.122
Turbo Code	0.053	0.836	0.087	1.191	0.164	1.590
LDPC Code	0.331	0.025	0.849	-0.800	2.414	-4.168

simulations are performed for systems with different number of diversity branches and different fading

power distributions. The simulation results indicate that the number of fading branch, N, and the

power distribution have negligible effects on γ_ω. That is, the slope and intercept of the log-domain

linear approximation of γ_ω is independent of N and the power distributions among the branches,

and they depend only on the modulation and channel coding. As an example, the values of γ_ω are

calculated for six different systems and the results are given in Table 2.2. The modulation is QPSK

and the channel code is code 1. The number of diversity branches are $N = 1, 3, 6$, respectively. There

are two different power distributions of the fading branches. The fading variances in distributions 1

and 2 satisfy $\sigma_n^2 = \frac{1}{N}$ and $\sigma_n^2 = \frac{2n-1}{N^2}$, respectively, for $n = 1, \cdots, N$. The values of γ_ω in Table 2.2 are

obtained through LS curve fitting between the analytical FER approximation and the simulated FER

under the various system configurations. A frame length of $L = 1,000$ is used for the calculations.

It is clear that there is only a slight variation (less than 0.18%) in the values of γ_ω across the six

different system configurations. The slight change in value is mainly caused by the limited numerical

precision used in the calculations.

The independence of γ_ω on N and power distribution can be intuitively explained by the mecha-

nism behind the threshold-based FER approximation. The value of γ_ω is chosen such that $\text{FER}_G(\gamma) \simeq$

1 when $\gamma \leq \gamma_\omega$ and $\text{FER}_G(\gamma) \simeq 0$ otherwise. Therefore, the value of γ_ω mainly depends on the shape

of $\text{FER}_G(\gamma)$ obtained in the AWGN channel, and it is independent of the actual SNR distribution

Table 2.2: The value of γ_ω in a QPSK modulated system with code 1.

	$N = 1$	$N = 3$	$N = 6$
Distribution 1	1.670	1.669	1.670
Distribution 2	1.670	1.667	1.667

$f_\gamma(x)$. As a result, the values of γ_ω is independent of the underlying fading distribution as shown in Table 2.2.

Combining (2.13) and (2.21) yields an accurate approximation of the FER in i.n.d. Rayleigh fading channel as

$$\text{FER} \simeq \sum_{v=1}^{V} \sum_{q=0}^{n_v - 1} \frac{\beta_q^{(v)}}{q!} \frac{\bar{\lambda}_v^{-q}}{\Gamma(n_v - q)} \; \psi\left(n_v - q, \frac{k_M \log L + b_M}{\bar{\lambda}_v}\right). \tag{2.22}$$

In (2.22), the FER is expressed as an explicit function of various system parameters, such as the number of branches $N = \sum_{v=1}^{V} n_v$, the value of $\lambda_n = r \log_2(M)\gamma_0 \sigma_n^2$ that is related to coding rate, modulation level, and fading power distribution, the frame length L, and the values of k_M and b_M, which are determined by the combination of modulation and coding used in the system.

2.5 Numerical Results

Simulation results are provided in this section to validate the newly proposed FER approximation, and to study the impacts of various system parameters on the FER performance.

Fig. 2.2 studies the impacts of fading power distribution on the FER for systems with $N = 4$. The channel code is code 1. The frame length is $L = 4,000$ bits for all systems. Power distributions 1 and 2 are the same as those in Table 2.2. The fading variances in power distribution 3 satisfy $\sigma_1^2 : \sigma_2^2 : \sigma_3^2 : \sigma_4^2 = 1 : 5 : 10 : 20$. The analytical FER approximations from (2.22) are compared to the FER obtained from Monte Carlo simulations, and they agree very well for all system configurations. Equal power distribution results in the best performance, and the FER degrades as the energy difference becomes larger. The diversity gain in a communication system can be measured by $\prod_{n=1}^{N} \sigma_n^2$. Based
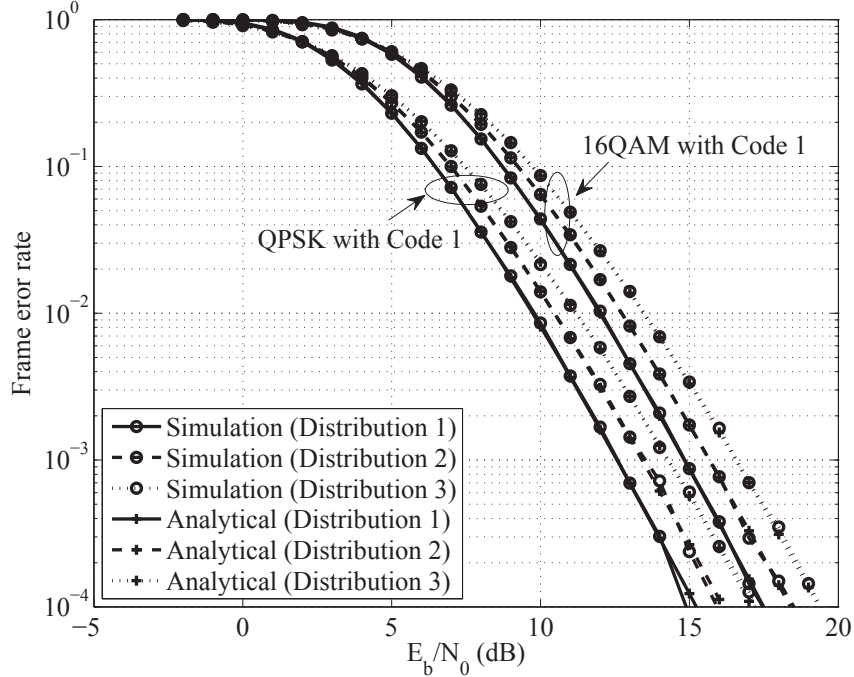
Figure 2.2: FER comparison of 3 distributions with $N = 4$.

on the geometric mean-arithmetic mean inequality, we have $\prod_{n=1}^{N} \sigma_n^2 \leq \prod_{n=1}^{N} \frac{1}{N}$ under the constraint $\sum_{n=1}^{N} \sigma_n^2 = 1$, and the equality is achieved when $\sigma_n^2 = \frac{1}{N}$. Therefore, the system with equal power on all branches has the highest diversity gain, thus the best FER performance.

Next, the FER performance for systems with different numbers of diversity branches is compared, and the results are shown in Fig. 2.3. The modulation scheme is QPSK, channel code is code 1, and the channel variances satisfy $\sigma_n^2 = \frac{2n}{N(N+1)}$, for $n = 1, \cdots, N$. All other parameters are the same as in Fig. 2.2. Excellent agreement is observed between the analytical and simulation results for all cases. As expected, the FER improves as N increases due to higher diversity orders. At FER $= 10^{-2}$, increasing N from 1 to 2 or from 2 to 4 yields a performance gain of 8 dB and 4.5 dB, respectively.

In Fig. 2.4, the analytical FER is shown as a function of L in the log domain. There are $N = 4$ branches with power distribution 2. The channel code is code 1. The average E_b/N_0 is 10 dB. It is shown that the FER increases monotonically with L, for all three modulation schemes. Given
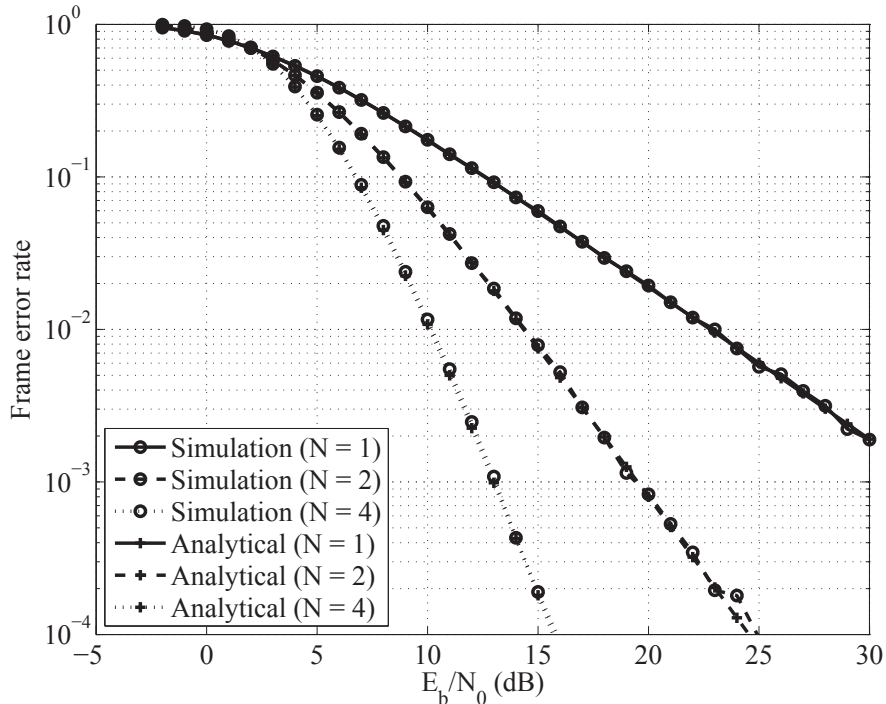
Figure 2.3: FER comparison of 3 transmission schemes for different N.

constraints on FER and L, we can easily choose the modulation scheme that will meet the design requirements by using the results in Fig. 2.4.

2.6 Conclusions

An accurate FER approximation was proposed for wireless communication systems with receiver diversity and i.n.d quasi-static Rayleigh fading. The FER approximation was obtained with a threshold-based method, with the threshold value modelled as a linear function of the frame length in the log domain. The analytical FER approximation was expressed as an explicit function of a large number of system parameters related to modulation, channel code, frame length, the number of diversity branches, and power distribution of the diversity branches. The parametric FER representation allows the flexible design and optimization of a wide range of practical communication systems, such as SIMO systems with MRC and HARQ systems with Chase combining. Simulation
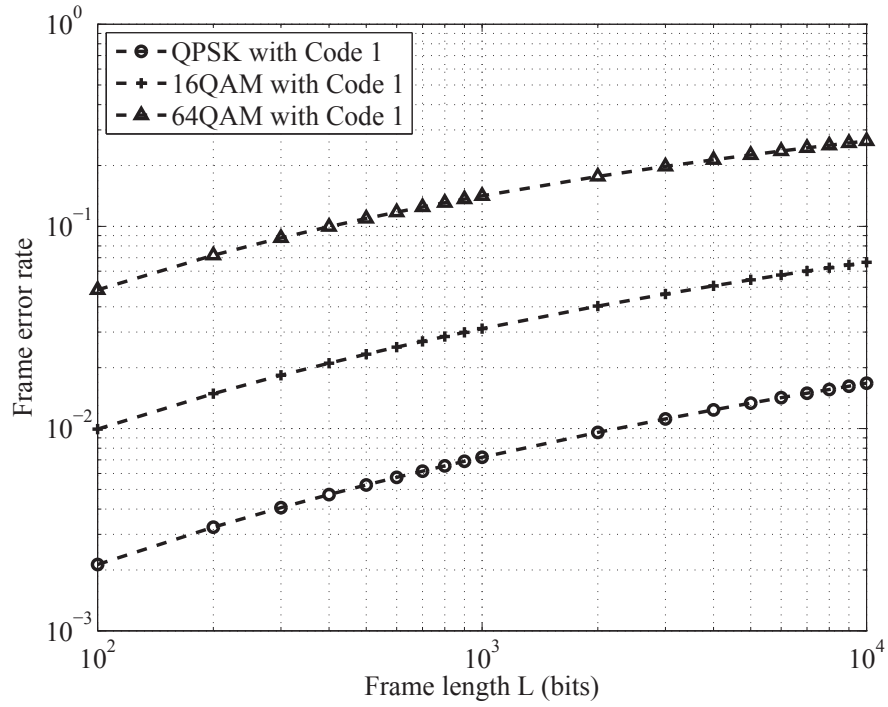
Figure 2.4: FER comparison of MQAM and Code 1 with $N = 4$ at $E_b/N_0 = 10$ dB.

results demonstrated that the proposed FER approximation can accurately predict the actual FER performance.

2.7 Documentation of multi-authored chapter

UNIVERSITY OF
ARKANSAS.

College of Engineering
Department of Electrical Engineering

June 11, 2014

To Whom It May Concern

This letter is to certify that Mr. Gang Wang, a Ph.D. candidate under my supervision at the Department of Electrical Engineering, has contributed more than 51% of the work for the following paper,

Gang Wang, Jingxian Wu, and Yahong R. Zheng, "An accurate frame error rate approximation of coded diversity systems with non-identical diversity branches," accepted in IEEE International Conference on Communications, ICC'14, Jun. 2014.

It should be noted that in my research group, the first author of a paper is usually the person who comes up with the original idea. For this paper, I came up with the original idea, and Mr. Gang Wang improved upon the idea and contributed more than 51% of the work to the paper.

Sincerely Yours,

Jingxian Wu
Director, Wireless Information Network Lab
Associate Professor of Electrical Engineering
University of Arkansas
Email: wuj@uark.edu
Tel: (479) 575-6584

2.8 References

[1] G. Wang, J. Wu, and Y. R. Zheng, "Cross-layer design of energy efficient coded ARQ systems," in *Proc. IEEE Global Telecommun. Conf. Globecom'12*, Anaheim, CA, Dec. 2012.

[2] J. Wu, G. Wang, and Y. R. Zheng, "Energy efficiency and spectral efficiency tradeoff in type-I ARQ systems," in *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 356 - 366, Feb. 2014.

[3] H. Xiao and A. H. Banihashemi, "Estimation of bit and frame error rates of finite-length low-density parity-check codes on binary symmetric channels," *IEEE Transactions on Communications*, vol. 55, no. 12, pp. 2234 - 2239, Dec. 2007.

[4] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," *42nd Annual Conference on Information Sciences and Systems, CISS 2008*, pp. 577 - 581, 2008.

[5] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "Threshold-based frame error rate analysis of MIMO systems over quasistatic fading channels," *Electronics Letters*, vol. 45, no. 4, Feb. 2009.

[6] T. Liu, L. Song, and B. Jiao, "Threshold-based frame error rate analysis of incremental hybrid relay selection scheme," in *Proc. IEEE Global Telecommun. Conf. Globecom'10*, Dec. 2010.

[7] J. Wu and C. Xiao, "Optimal diversity combining based on linear estimation of Rician fading channels," *IEEE Transactions on Communications*, vol. 56, pp. 1612-1615, Oct. 2008.

[8] J. Wu and Y. R. Zheng, "Optimum multi-hop transmission strategies for energy constrained wireless sensor networks," *IEEE International Conf. on Communications, ICC'12*, 2012.

[9] G. Wang, J. Wu, and Y. R. Zheng, "Optimum energy efficient communications for hybrid ARQ systems," accepted in *IEEE Global Telecommun. Conf. Globecom'13*, 2013.

[10] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "Optimum power assignment for minimizing the average total transmission power in hybrid-ARQ Rayleigh fading links," *IEEE Transactions on Communications*, vol. 59, no. 7, pp. 1867 - 1877, July 2011.

[11] M.-S. Alouini and A. J. Goldsmith, "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 4, pp. 1165 - 1181, Jul. 1999.

[12] J. Wu, N. B. Mehta, A. F. Molisch, and J. Zhang, "Unified spectral efficiency analysis of cellular systems with channel-aware schedulers," *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3463 - 3474, Dec. 2011.

[13] J. Wu, G. Wang, and Y. R. Zheng, "Energy and spectral efficient transmissions of coded ARQ systems," in *Proc. IEEE International Conference on Communications, ICC'13*, June 2013.

[14] Y. Geng, Y. Wan, J. He and K. Pahlavan, "An empirical channel model for the effect of human body on ray tracing," *IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[15] Y. Wang, S. Cui, R. Sankar, and S. Morgera, "Delay-throughput tradeoff with opportunistic relaying in wireless networks", in *Proc. IEEE Global Telecommun. Conf. Globecom'11*, Houston, Dec. 2011.

[16] K. Xu, D. Tipper, Y. Qian, P. Krishnamurthy, and S. Tipmongkonsilp, "Time-Varying performance analysis of multihop wireless networks with CBR traffic", accepted in *IEEE Transactions on Vehicular Technology*, 2014.

[17] C. Prasartkaew and S. Choomchuay, "A design of parity check matrix for irregular LDPC codes," *IEEE 9th International Symposium on Communications and Information Technology, ISCIT 2009*, pp. 239 - 242, 2009.

[18] Y. Geng, J. Chen and K. Pahlavan, "Motion detection using RF signals for the first responder in emergency operations," *Proc. IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[19] Y. Wang, G. Yue, S. Rangarajan, R. Sankar and S. Morgera, "Buffer-Aware adaptive scheduling for downlink multiuser systems", in *Proc. IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[20] K. Xu, D. Tipper, P. Krishnamurthy, and Y. Qian, "An efficient hybrid model and dynamic performance analysis for multihop wireless networks", in *Proc. IEEE International Conference on Computing, Networking and Communications ICNC*, San Diego, CA, Jan., 2013.

[21] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power gird by sparse optimization", to appear *IEEE Transactions on Smart Grid*, 2014.

[22] L. Liu, M. Esmalifalak, and Z. Han, "Detection of false data injection in power grid exploiting low rank and sparsity", *IEEE International Conference on Communication, ICC'13*, Hungary, Jun. 2013.

[23] H. E. Gamal and A. R. Hammons, Jr., "Analyzing the turbo decoder using the Gaussian approximation," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 671 - 686, Feb. 2001.

[24] M. K. Simon and M. Alouini, *Digital communication over fading channels: a unified approach to performance analysis*, John Wiley & Sons, 2000.

<center>**Chapter 3**</center>

<center>**Cross-Layer Design of Energy Efficient Coded ARQ Systems**</center>

<center>Gang Wang, Jingxian Wu, and Yahong Rosa Zheng</center>

3.1 Abstract

The energy efficient design of coded automatic-repeat-request (ARQ) systems is studied in this paper. The optimization aims to minimize the energy required for the successfully delivery of one information bit from a transmitter to a receiver. The design is performed by incorporating a wide range of practical system parameters and metrics, such as hardware power consumption, modulation, channel coding, and frame error rate (FER) in the physical layer, and frame length and protocol overhead in the media access control layer. A new log-domain threshold approximation method is proposed to analytically quantify the impacts of the various system parameters on the FER, and the results are used to facilitate the system design. The optimum transmission energy and frame length that minimize the energy per information bit are identified in closed-form expressions as functions of the various practical system parameters. The analytical and simulation results demonstrate that the total energy consumption in a coded ARQ system can be *reduced* by *increasing* the transmission energy during one transmission attempt, and significant energy saving as high as 9.5 dB is achieved with the optimum system.

3.2 Introduction

Energy efficient communication can extend the battery life of communication terminals, reduce the energy cost, and make the communication process more environmental friendly.

A large number of energy efficient communication techniques have been developed in the physical

(PHY) layer [1] and [2] and the media access control (MAC) layer [3]– [7]. Most PHY layer energy efficient communication techniques are developed by exploiting the trade-off between power efficiency and spectral efficiency through various coding, modulation, and signal processing techniques [1] and [2]. In the MAC layer, the energy consumption can be reduced in a number of ways, such as decreasing the transmission duty cycle [3] and [4], carefully scheduling the transmissions to reduce or avoid collisions [5] and [6], or power controls [7], etc.

Most schemes are developed by following the traditional layered-protocol design approach, and they do not directly take advantage of the interactions among the protocol layers that might be critical to energy efficient communications [8]. A cross-layer power-rate-distortion framework is proposed in [9] by considering the trade-off among source distortion, data rate, and hardware complexity, but with an assumption of error-free channel. A PHY/MAC cross-layer design is considered in [10], where the optimum power assignment for the hybrid automatic-repeat-request (H-ARQ) technique in fading channel is studied to reduce the total average power consumption. The optimization in [10] is performed under the constraint of a targeted outage probability, and it does not consider the effects of practical system parameters such as overhead, modulation, data rate, and bit error rate (BER), etc.

In this paper, we propose a new optimum design of practical ARQ systems to minimize the energy required to successfully deliver an information bit from a transmitter to a receiver through a Rayleigh fading channel. The optimization incorporates a large number of practical system parameters that cover the operations in the hardware, the PHY layer, and the MAC layer, such as the efficiency of the power amplifier, the power consumption of digital hardware, data rate, modulation, frame length, frame error rate (FER), and the protocol overhead, etc.

The system design is performed by jointly optimizing the transmission energy in the PHY layer

and the frame length in the MAC layer. For a system employing ARQ, a lower transmission energy does not necessarily mean less total energy consumption, because it might increase the number of retransmissions, thus the total energy required to *successfully* deliver a frame. On the other hand, increasing the transmission energy beyond its optimum operation point will result in a waste of the energy resource. Similarly, a longer frame usually means a higher FER, yet a shorter frame has poor overhead efficiency. To quantify the impacts of transmission energy and frame length, a new log-domain threshold approximation is proposed to build an explicit analytical relationship between the FER and the design parameters. The optimum transmission energy and frame length are expressed as closed-form expressions of the various practical system parameters. The analytical and simulation results demonstrate that significant energy savings are achieved through the optimization.

3.3 System Model

Consider a transmitter and a receiver separately by a distance d. The information bits at the transmitter are divided into frames. Each frame has L uncoded information bits and L_0 overhead bits. The information bits and overhead bits from the transmitter are encoded with a channel encoder with code rate r. For a system employing M-QAM, the number of symbols in each frame is $L_s = \frac{L+L_0}{r \log_2 M}$, where L is chosen in a way such that L_s is an integer.

The m-th symbol observed at the receiver is

$$y_m = \sqrt{E_r} h_m x_m + z_m, \text{ for } m = 1, 2, \cdots, L_s, \tag{3.1}$$

where E_r is the average energy of a symbol at the receiver, $x_m \in \mathcal{S}$ is the m-th modulated symbol transmitted, $\mathcal{S}$ is the modulation constellation set with the cardinality $M = |\mathcal{S}|$, y_m, h_m, and z_m are the received sample, the fading coefficient between the transmitter and the receiver, and additive white Gaussian noise (AWGN) with single-sided power spectral density N_0, respectively. It is assumed

that the system undergoes quasi-static Rayleigh fading, such that the fading coefficient is constant within one frame, and changes from frame to frame.

Define the average E_b/N_0 of an uncoded information bit at the receiver as

$$\gamma_b \triangleq \frac{E_b}{N_0} = \frac{E_r}{r N_0 \log_2 M}. \tag{3.2}$$

For a transmitter and receiver pair separated by a distance d, the average transmission energy for each symbol at the transmitter can be modelled as [1]

$$E_s = E_r G_1 d^\kappa M_l, \tag{3.3}$$

where κ is the path-loss exponent, G_1 is the gain factor (including path-loss and antenna gain) at a unit distance, and M_l is the link margin compensating the hardware process variations and other additive background noise or interference.

In addition to the actual transmission energy, we also need to consider the circuit energy per symbol that can be modelled as [1],

$$E_c = \left(\frac{\xi_M}{\eta} - 1 \right) E_s + \frac{\beta}{R_s}, \tag{3.4}$$

where $R_s = \frac{1}{T_s}$ is the gross symbol rate, η is the drain efficiency of the power amplifier, ξ_M is the peak-to-average power ratio (PAPR) of an M-ary modulation signal, β incorporates the effects of baseband processing, such as signal processing, encoding and modulation. For M-ary quadrature amplitude modulated (MQAM) systems with square constellations, $\xi_M \simeq 3(\sqrt{M} - \frac{1}{\sqrt{M}} + 1)$ for $M \geq 4$ [11].

From (3.2), (3.3), and (3.4), the energy required to transmit one information bit during one transmission attempt is

$$E_0 = \frac{L_s}{L}(E_s + E_c) = \frac{L + L_0}{L} \frac{\gamma_b \xi_M N_0 G_d}{\eta} + \frac{\beta}{R_b}, \tag{3.5}$$

where $G_d = G_1 d^\kappa M_l$, and $R_b = \frac{L}{L_s} R_s$ is the net bit rate of the uncoded information bit.

Due to the effects of channel fading and noise, the receiver might not be able to successfully recover the transmitted signal. The probability that a transmitted frame cannot be recovered equals to FER, which is a function of the γ_b at the receiver, the frame length L_s, the modulation level M, and the channel code. The packet will be retransmitted if the transmitter receive a negative acknowledgement (NACK). Since the retransmissions are independent, the number of retransmissions is a geometric random variable with the parameter FER. The average number of retransmissions is thus

$$\Lambda = \frac{1}{1 - \text{FER}}. \tag{3.6}$$

The total energy required to successfully deliver an information bit from the transmitter to the receiver can then be calculated by $E_t = \Lambda E_0$, which can be expanded by combining (3.5) and (3.6) as

$$E_t = \frac{1}{1 - \text{FER}} \left[\frac{L + L_0}{L} \frac{\gamma_b \xi_M N_0 G_d}{\eta} + \frac{\beta}{R_b} \right]. \tag{3.7}$$

The total energy per information bit E_t relies on a number of system parameters, including E_b/N_0 at the receiver γ_b, the number of information bits L and the number of overhead bits L_0 per frame, the modulation level M, the net data rate R_b, and the FER that inherently depends on all the above parameters and the code rate r, etc.

The value of γ_b has two opposite effects on E_t. On one hand, FER is a decreasing function in γ_b. Therefore, increasing γ_b will decrease the average number of retransmissions Λ, thus reduce E_t. On the other hand, E_0 is a strictly increasing function in γ_b, thus it translates a positive relationship between γ_b and E_t.

A similar observation can also be obtained for the relationship between E_t and L. Λ translates a positive relationship between E_t and L because FER is an increasing function in L for a given channel code and modulation scheme, whereas E_0 is a decreasing function in L.

Therefore, it is critical to identify the optimum values of γ_b and L that can achieve minimal energy per information bit.

3.4 Optimum System Design

The optimum system design that can minimize E_t under the constraints of fixed M, R_b and L_0 are studied in this section.

3.4.1 FER with a Log-Domain Linear Threshold Approximation

In this subsection, an accurate approximation of the FER of coded systems in quasi-static Rayleigh fading is obtained with the threshold-based method originally presented in [12]. Furthermore, we propose a new log-domain linear approximation method for the calculation of the threshold value required for the FER approximation. The threshold-based method with the newly proposed log-domain linear approximation explicitly build a connection between the FER and the various system parameters.

With the threshold-based method [12], the FER of a coded system in a quasi-static Rayleigh fading channel can be accurately approximated by

$$\text{FER} \simeq 1 - \exp\left(-\frac{\gamma_\omega}{\gamma_b}\right), \tag{3.8}$$

where γ_ω is a threshold value that can be calculated as

$$\gamma_\omega = \left[\int_0^\infty \frac{1 - \text{FER}_G(\gamma)}{\gamma^2} d\gamma\right]^{-1}, \tag{3.9}$$

where $\text{FER}_G(\gamma)$ is the FER in an AWGN channel.

Fig. 3.1 shows γ_ω as a function of $L + L_0$ under various modulation schemes. The channel code is a rate $r = \frac{1}{2}$ convolutional code with the generator polynomial $[5, 7]_8$ and constraint length 3. It is

observed from the figure that γ_ω can be modelled as a linear function of $\log(L + L_0)$, with the slope and intercept determined by the different modulation schemes. Similar linear relationships are also observed for other channel codes. Therefore, we propose to model γ_ω as

$$\gamma_\omega \simeq k_M \log(L + L_0) + b_M, \tag{3.10}$$

where k_M and b_M are the slope and intercept determined by the modulation scheme and the actual channel code. The value of k_M and b_M can be estimated by performing the least squares (LS) method on the results in Fig. 3.1. For the $M = 4$, we have $k_4 = 0.3744$ and $b_4 = -0.31$.

Combining (3.8) and (3.10) leads to a new FER approximation

$$\text{FER} \simeq 1 - (L + L_0)^{-\frac{k_M}{\gamma_b}} \exp\left(-\frac{b_M}{\gamma_b}\right). \tag{3.11}$$

Fig. 3.2 compares the actual FER obtained through simulation with the corresponding analytical approximation by using (3.11), under different values of $L + L_0$, for systems with $M = 4$. The convolutional code is the same as the one used in Fig. 3.1. Excellent agreements are observed between the actual simulation results and their analytical approximations. Therefore, the analytical expressions in (3.8) and (3.10) give a very accurate approximation of the actual FER.

3.4.2 Optimum γ_b

The optimum value of γ_b at the receiver that minimizes E_t is studied in this subsection.

Before proceeding to the actual optimization, we present the following theorem about convexity, which will be used in identifying the optimum system parameters.

Theorem 1: Consider a decreasing function $f(x)$ and an increasing function $g(x)$. If both $f(x)$ and $g(x)$ are convex, then $f(x)g(x)$ is convex.
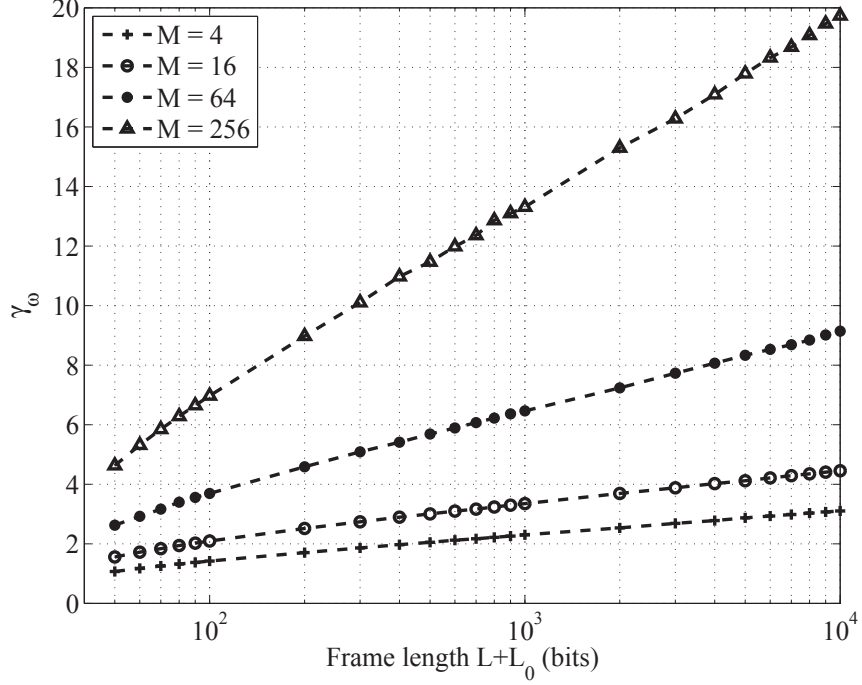
Figure 3.1: γ_ω as a function of $L + L_0$

Proof: Consider $0 < x_1 < x_2$ and $\alpha \in [0,1]$. Define $\theta_1 = \alpha f(x_1)g(x_1) + (1-\alpha)f(x_2)g(x_2)$, and $\theta_2 = f(\alpha x_1 + (1-\alpha)x_2)g(\alpha x_1 + (1-\alpha)x_2)$. Since $f(x)$ and $g(x)$ are convex, we have $\theta_2 \geq \theta_3$ with θ_3 defined as

$$\theta_3 = [\alpha f(x_1) + (1-\alpha)f(x_2)][\alpha g(x_1) + (1-\alpha)g(x_2)] \tag{3.12}$$

Since $\theta_1 = \theta_1(1-\alpha+\alpha)$, the term θ_1 can be alternatively represented as

$$\theta_1 = \alpha^2 f(x_1)g(x_1) + (1-\alpha)^2 f(x_2)g(x_2) + \alpha(1-\alpha)\left[f(x_1)g(x_1) + f(x_2)g(x_2)\right] \tag{3.13}$$

From (3.12) and (3.13), we have

$$\frac{\theta_3 - \theta_1}{\alpha(1-\alpha)} = [f(x_1) - f(x_2)][g(x_2) - g(x_1)] \geq 0. \tag{3.14}$$

Therefore $\theta_2 \geq \theta_3 \geq \theta_1$, and this completes the proof. ∎

We can prove that Λ in (3.6) is a decreasing function in γ_b, and it is convex in γ_b by showing that $\frac{\partial^2 \Lambda}{\partial \gamma_b^2} \geq 0$, and details are omitted here for brevity. It is straightforward to show that E_0 in (3.5) is
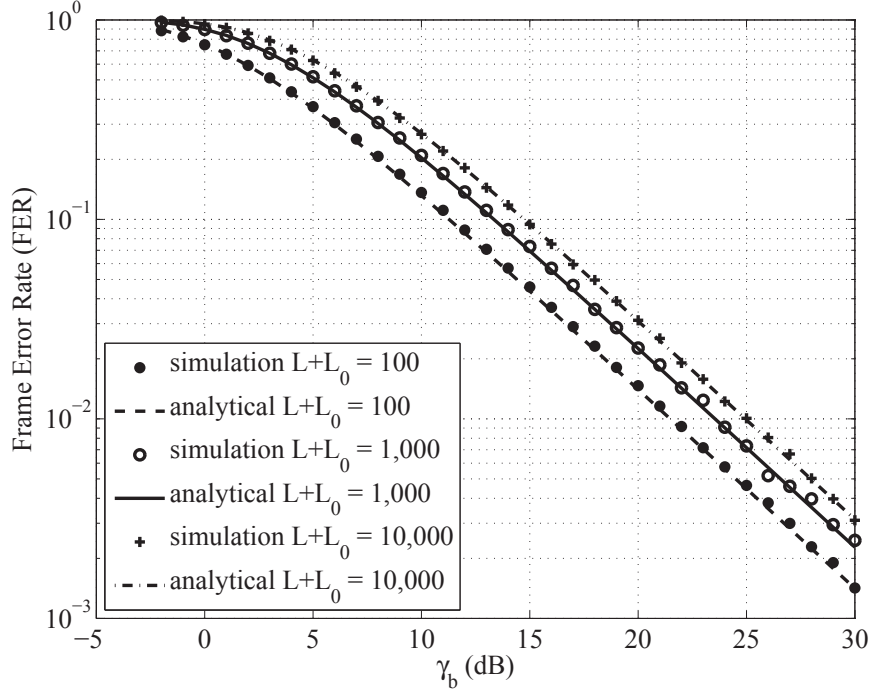
41

Figure 3.2: Comparison of the simulation FER with the analytical approximation in (3.11).

an increasing and convex function in γ_b. Therefore, based on the results in Theorem 1, we have the following corollary about the convexity of $E_t = \Lambda E_0$.

Corollary 1: For the FER given in (3.11), the total energy per information bit, E_t, in (3.7) is convex in γ_b. ∎

Once we establish the convexity of E_t in γ_b, the optimum γ_b can be solved as stated in the following corollary.

Corollary 2: In a quasi-static Rayleigh fading channel, if the FER is given in (3.11), then the optimum γ_b that minimizes E_t is

$$\hat{\gamma}_b = \frac{1}{2}\left(\gamma_\omega + \sqrt{\gamma_\omega^2 + 4\gamma_\omega \frac{B}{A}\frac{L}{L+L_0}}\right) \tag{3.15}$$

where $A = \frac{\xi_M N_0 G_d}{\eta}$, and $B = \frac{\beta}{R_b}$.

42

Proof: Since E_t is convex in γ_b, the optimum γ_b that minimize E_t can be obtained by solving $\frac{\partial E_n}{\partial \gamma_b} = 0$, which yields

$$\gamma_b^2 - \gamma_\omega \gamma_b - \gamma_\omega \frac{B}{A} \frac{L}{L + L_0} = 0 \tag{3.16}$$

The result in (3.15) can be obtained by solving (3.16). ∎

It should be mentioned here that the optimum γ_b is the average E_b/N_0 at the receiver. Correspondingly, the optimum energy per symbol required at the transmitter is

$$\hat{E}_s = \hat{\gamma}_b \times N_0 \times r \times \log_2 M \times G_d, \tag{3.17}$$

where $\hat{\gamma}_b$ is the optimum value calculated from (3.15).

3.4.3   Optimum L

The optimum number of information bits L that minimizes E_t is studied in this subsection.

Similar to the results in Corollary 2, the optimum solution of L relies on the convexity of E_t. However, the direct proof of the convexity of E_t with respect to L is quite tedious. To simplify analysis, we can show that E_t is convex in $\xi = \log(L + L_0)$.

We can prove that 1) Λ in (3.6) is an increasing and convex function in ξ; and 2) E_0 in (3.5) is a decreasing and convex function in ξ, and details are omitted here for brevity. Therefore, based on Theorem 1, we have the following corollary regarding the convexity of E_t with respect to ξ.

Corollary 3: For the FER given in (3.11), the total energy per information bit E_t in (3.7) is convex in $\xi = \log(L + L_0)$. ∎

Based on the convexity of E_t in L, the optimum L is stated as follows.

Corollary 4: In a quasi-static Rayleigh fading channel, if the FER is given in (3.11), then the optimum L that minimize E_t satisfies the following equality

$$\hat{L} = \frac{\sqrt{A^2(k_M + \gamma_b)^2 + 4Ak_M B} - A(k_M - \gamma_b)}{2k_M(A\gamma_b + B)}\gamma_b L_0 \tag{3.18}$$

where $A = \frac{\xi_M N_0 G_d}{\eta}$, and $B = \frac{\beta}{R_b}$.

Proof: The optimum L is obtained by solving $\frac{\partial E_n}{\partial \xi} = 0$, which yields

$$k_M(A\gamma_b + B)L^2 + A\gamma_b L_0 (k_M - \gamma_b) L - A\gamma_b^2 L_0^2 = 0 \tag{3.19}$$

The result in (3.18) can be obtained by solving (3.19). ∎

It is worth pointing out that even though the result in Corollary 4 is obtained through $\frac{\partial E_n}{\partial \xi} = 0$, it is exactly the same as solving $\frac{\partial E_n}{\partial L} = 0$ because $\frac{\partial \xi}{\partial L} = \frac{1}{L+L_0} \neq 0$.

3.4.4 Joint Optimum γ_b and L

In (3.15) and (3.18), the optimum value of γ_b is expressed as a function of L and vice versa. The global optimum operation point can be achieved by jointly optimizing γ_b and L.

Since E_t is convex in both γ_b and L, the joint optimum values can be obtained by treating (3.15) and (3.18) as a system of two equations with two variables in γ_b and L. The analytical results are very tedious and are omitted here for brevity.

Alternatively, the joint optimum values of γ_b and L can be efficiently calculated by iteratively invoking (3.15) and (3.18). Given an initial value L, we can calculate the optimum γ_b by using (3.15), the output of which is then used to update the value of L with (3.18). This procedure can be performed iteratively, and it will converge to the joint optimum value of γ_b and L that achieves the global minimum energy consumption.

Table 3.1: Simulation Parameters

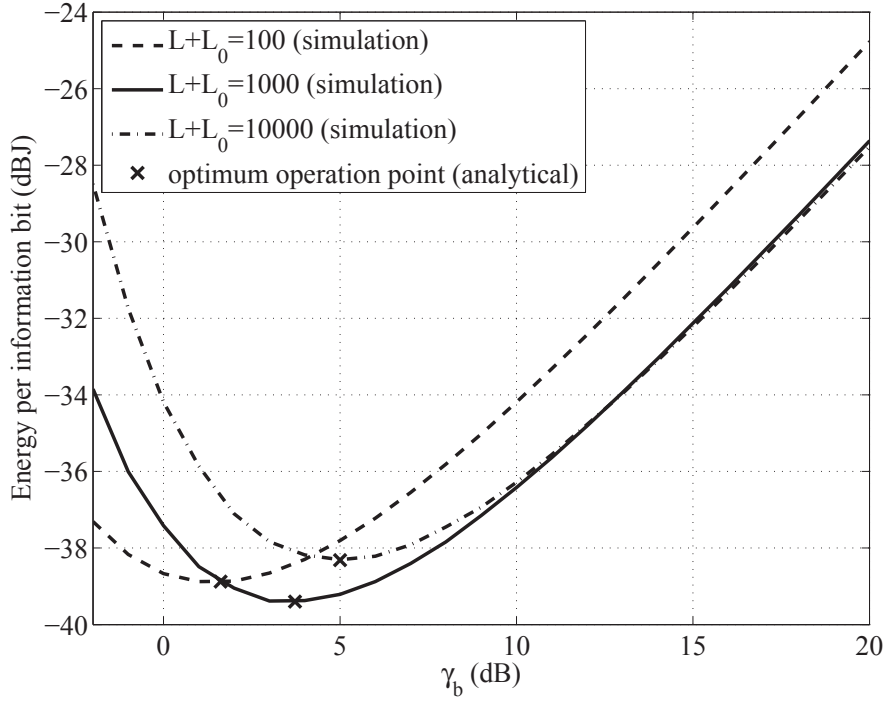| L_0	48 bits
Bit Rate	300 kbps
η	0.35
β	310.014 mw
$N_0/2$	-174 dBm/Hz
G_1	30 dB
κ	3.5
M_l	40 dB

Figure 3.3: Energy per information bit E_t v.s. γ_b at the receiver.

3.5 Numerical Results

Numerical results are presented in this section. The simulation parameters are summarized in Table 3.1.

Fig. 3.3 shows E_t as a function of γ_b, with various values of $L + L_0$. The distance is $d = 100$ m. The optimum values of $\hat{\gamma}_b$ for different L calculated from (3.15) are marked on the figure as the optimum operation points. It can be seen from the figure that E_t is a convex function in γ_b. The
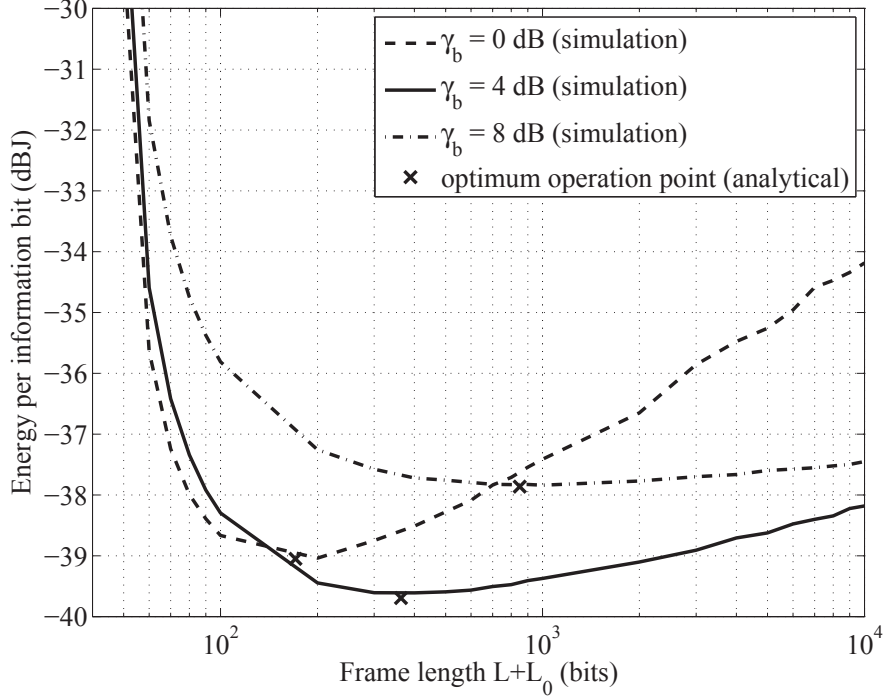
Figure 3.4: Energy per information bit E_t v.s. number of bits per frame $L + L_0$.

optimum operation points obtained from the analytical results match perfectly with the simulation results. If $\gamma_b < \hat{\gamma}_b$, the FER is so high such that the total energy consumption is dominated by the effect of the retransmissions. In this case, we can *reduce* the total energy consumption by *increasing* γ_b. For example, for $L + L_0 = 10,000$, increasing γ_b from -2 to 5 dB will result in an energy saving of 9.5 dB. When $\gamma_b > \hat{\gamma}_b$, E_t increases almost linearly with γ_b because the FER is low enough such that the effect of retransmission is negligible. The result demonstrates that a higher E_b/N_0 does not necessarily mean a better performance. Significant energy saving can be achieved with carefully choosing the operation point.

In Fig. 3.4, E_t is plotted as a function of $L + L_0$ under various values of γ_b. The distance is $d = 100$ m. The optimum values of $\hat{L}$ for different γ_b are calculated from (3.18), and are marked on the figure. As expected, E_t is convex in $\log(L + L_0)$. Again, excellent agreement is observed between the analytical optimum operation points and the simulation results. When $L < \hat{L}$, the
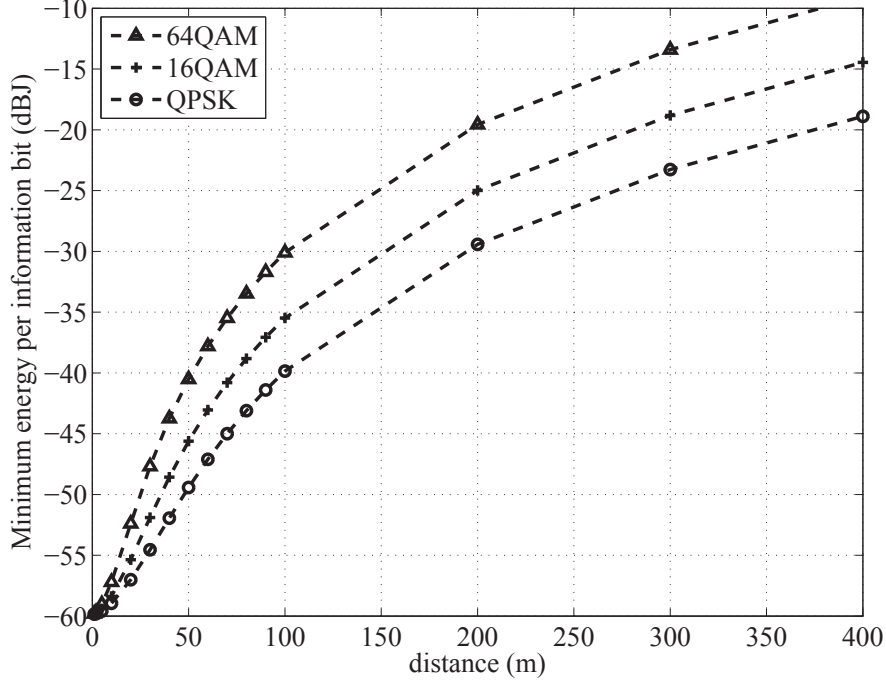
46

Figure 3.5: Minimum energy per information bit as a function of distance.

energy consumption is dominated by the overhead. Thus significant energy saving can be achieved by slightly increasing L. When $L > \hat{L}$, the slope of E_t with respect to $\log(L + L_0)$ decreases as γ_b increases. This is because the impact of increasing L on FER becomes smaller at higher γ_b. Therefore, system operates at lower γ_b is more sensitive to the frame length.

In the last example, the global optimum E_t is shown as a function of the transmitter-receiver distance d, for systems employing different modulation schemes. The joint optimum $(\hat{\gamma}_b, \hat{L})$ are obtained by iteratively invoking (3.15) and (3.18), and the results are then used to calculate the optimum E_t. For example, at $d = 100$ m, the optimum values are (2.66 dB, 230 bits), (4.29 dB, 243 bits), and (6.77 dB, 191 bits) for QPSK, 16-QAM, and 64-QAM, respectively. Lower level modulation has better energy performance at the cost of worse spectral efficiency. Increasing M from 2 to 4, or from 4 to 6, results in approximately 5 dB energy loss, when $d \geq 100$ m.

3.6   Conclusions

The energy efficient design of coded ARQ systems operating in a quasi-static Rayleigh fading channel has been studied in this paper. A new log-domain threshold approximation method has been proposed to analytically quantify the impacts of receiver E_b/N_0 and frame length on the FER, and the results have been used to facilitate the system optimization. The optimum transmission energy and frame length that minimize the energy per information bit have been obtained in closed-form expressions, and they incorporate the effects of a large number of practical system operation parameters in hardware, the PHY layer, and the MAC layer. From the analytical and simulation results, we have the following observations: 1) The total energy consumption in ARQ can be *reduced* by *increasing* the transmission energy in one transmission attempt; 2) systems operating at higher E_b/N_0 are less sensitive to the frame length; 3) increasing the modulation level by a factor of 4 leads to approximately 5 dB energy loss; 4) significant energy savings (as high as 9.5 dB) can be achieved through the proposed optimum system design.

3.7 Documentation of multi-authored chapter

UNIVERSITY OF ARKANSAS

College of Engineering
Department of Electrical Engineering

June 11, 2014

To Whom It May Concern

This letter is to certify that Mr. Gang Wang, a Ph.D. candidate under my supervision at the Department of Electrical Engineering, has contributed more than 51% of the work for the following paper,

Gang Wang, Jingxian Wu, and Yahong R. Zheng, "Cross-layer design of energy efficient coded ARQ systems," IEEE Global Communications Conference, GLOBECOM, Dec. 2012.

It should be noted that in my research group, the first author of a paper is usually the person who comes up with the original idea. For this paper, I came up with the original idea, and Mr. Gang Wang improved upon the idea and contributed more than 51% of the work to the paper.

Sincerely Yours,

Jingxian Wu
Director, Wireless Information Network Lab
Associate Professor of Electrical Engineering
University of Arkansas
Email: wuj@uark.edu
Tel: (479) 575-6584

3.8   References

[1] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transaction on Wireless Communications*, vol. 4, pp. 2349 - 2360, Sept. 2005.

[2] F. M. Costa and H. Ochiai, "Energy-efficient physical layer design for wireless sensor network links," *IEEE International Conference on Communications, ICC*, June 2011.

[3] L. L. Dai and P. Basu, "Energy and delivery capacity of wireless sensor networks with random duty-cycles," *IEEE International Conference on Communications, ICC,* vol. 8, pp. 3503 - 3510, June 2006.

[4] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Transaction on Networking*, vol. 12, pp. 493 - 506, June 2004.

[5] A. E. Gamal, C. Nair, B. Prabhakar, E. Uysal-Biyikoglu, and S. Zahedi, "Energy-efficienty scheduling of packet transmissions over wireless networks," *IEEE INFOCOM 2002*, vol. 3, pp. 1773 - 1782, June 2002.

[6] C. Schurgers and M. B. Srivastava, "Energy optimal scheduling under average throughput constraint communications," *IEEE International Conference on Communications, ICC,* vol. 3, pp. 1648 - 1652, May 2003.

[7] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless networks," *IEEE Transaction on Communications,* vol. 5, pp. 3306 - 3315, Nov. 2006.

[8] G. Miao, N. Himayat, Y. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wiley J. Wireless Communication and Mobile Computing,* vol. 9, pp. 529 - 542, Apr. 2009.

[9] Zhihai He, Wenye Chen, and Xi Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 18, pp. 596 - 608, May 2008.

[10] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "Optimal power assignment for minimizing the average total transmission power in hybrid-ARQ Rayleigh fading links, *IEEE Transactions on Communications*, vol. 59, no. 7, pp.1867-1877, July 2011.

[11] J. Cioffi, *Digital Communications,* Stanford Univ. Press, Fall 2001.

[12] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," *42nd Annual Conference on Information Sciences and Systems, CISS 2008*, pp. 577 - 581, Mar. 2008.

3.9 Copyright Clearance

Optimum Energy and Spectral Efficient Transmissions for Delay Constrained Hybrid ARQ Systems

Gang Wang, Jingxian Wu, and Yahong Rosa Zheng

4.1 Abstract

This paper discusses the energy and spectral efficient transmissions for delay constrained wireless communication systems utilizing hybrid automatic repeat request (HARQ). Three design metrics are considered: 1) the energy efficiency (EE) measured by the average energy required to successfully deliver one information bit from a source to its destination; 2) the spectral efficiency (SE) defined as the average data rate per unit bandwidth; and 3) the average energy per bit normalized by the SE. Optimum system designs with respect to different metrics are developed by analytically identifying the feasible energy distributions, i.e., the sequences of transmission energy that should be employed at different retransmissions. The optimum energy distributions are calculated with a new backward sequential calculation method, where the energy at each retransmission round is expressed as a closed-form expression of a wide range of practical system parameters, such as circuit power and coding in the physical layer, and frame length and protocol overhead in the media access control layer. Reducing the SE-normalized energy per bit yields a balanced tradeoff between the EE and SE. Numerical results demonstrate that reducing the SE-normalized energy per bit achieves a 26.7% SE gain over the system that minimizes the average energy per bit, yet the EE of the two systems are almost identical.

4.2 Introduction

Energy efficiency (EE) and spectral efficiency (SE) are two essential design metrics for wireless communication systems. Energy efficient communications reduce energy consumption, extend the battery life of wireless terminals, and reduce environment pollution [1]– [2]. Spectral efficient communications can support more simultaneous users or a higher data rate in a unit bandwidth, thus achieve better utilizations of the scarce spectrum resources [3]. EE and SE feature one of the most fundamental tradeoffs in communications [4]. Many communication techniques improve EE at the cost of SE, or vice versa. With the increased demands for high rate multimedia communications, and the growing needs of green communications, it is imperative to develop communication technologies that can balance the EE-SE tradeoff.

There are a large number of works in the literature devoted to the development of energy efficient communication systems, and many are designed across multiple protocol layers [3], [5]– [12]. Cross-layer design is especially beneficial to multimedia transmissions [7]– [8], because the transmission parameters in the lower protocol layers can be directly optimized by upper layers based on specific quality of service (QoS) requirements, such as data rate and transmission delay. In [5], an energy efficient video transmission system is designed with an energy-rate-distortion framework that considers the tradeoffs among source distortion, data rate, and hardware complexity, yet with an assumption of error-free channel. In [6], the average transmission energy is minimized for automatic repeat request (ARQ) systems by gradually increasing the transmission energy after each unsuccessful retransmission attempt, in both single-hop and multi-hop communications under a delay constraint. In [9] and [10], the optimum transmission energy distribution, i.e., the sequence of energy employed at different retransmission rounds of hybrid ARQ (HARQ) systems, is identified by minimizing the total average energy consumption, under the assumption that the retransmissions of the same packet

undergo the same fading. The quasi-static fading assumption is reasonable for multimedia communications given the stringent delay constraint. It is assumed in [9] and [10] that a retransmission occurs if the signal-to-noise ratio (SNR) is below a threshold, yet in practical systems a packet is retransmitted if there are unrecoverable errors in the received packets. The results in [6]– [10] do not consider the circuit energy consumption, which is shown to be non-negligible especially for short range communications [11]. In [3] and [12], the energy per bit in a type-I ARQ system is minimized by considering many practical system parameters including packet length, circuit energy, protocol overhead, modulation and coding schemes, and frame error rate (FER).

The EE is usually achieved at the cost of SE. Recently there are growing interests in system designs that consider the EE-SE tradeoff [3], [4], [13]– [15]. In [3], it is shown through analytical study that the EE is a quasi-concave function of SE in type-I ARQ systems. A similar observation is obtained in [13] for a downlink orthogonal frequency division multiple access (OFDMA) system, and in [14] for a band unlimited system under a rate constraint. The system configurations corresponding to the positive EE-SE slope are non-desirable because the EE and SE can always be improved simultaneously in this region by adjusting the operation parameters. Consequently, all practical systems should operate at the region with the negative EE-SE slope, which means any improvement in SE will be achieved at the cost of EE. Both [13] and [14] use the Shannon channel capacity to measure data rate without considering practical operations such as coding, hardware energy consumption, and protocol overhead. A new metric, the SE-normalized energy per bit, which is defined as the average energy per bit normalized by the SE, is introduced in [3] and [15] to balance the EE-SE tradeoff. It is shown that minimizing the SE-normalized energy per bit can achieve significant gain in SE with only a small cost in EE. The results in [3] and [15] are only applicable to type-I ARQ systems.

In this paper, we study the energy and spectral efficient transmissions for coded HARQ systems

55

with delay constraints. In HARQ systems with Chase combining [16], the signals from all transmission attempts of the same packet are combined coherently at the receiver during the detection process. A strict delay constraint can thus be enforced by setting the maximum number of retransmissions allowed by the system. The optimum designs are performed by identifying the optimum energy distributions that can minimize or maximize certain design metrics. Three metrics are considered for different EE-SE tradeoffs. The first metric is the average energy required to successfully deliver one information bit from a source to its destination; the second metric is the SE defined as the effective data rate per unit bandwidth; and the third metric is the SE-normalized energy per bit as defined in [3] and [15]. The EE can be maximized by minimizing the average energy per bit at the cost of SE, and vice versa. The minimization of the SE-normalized energy per bit provides a balanced tradeoff between the EE and SE. The optimum energy distributions with respect to EE and SE are analytically identified by considering a large number of practical system parameters, such as the efficiency of the energy amplifier, the energy consumption of digital hardware, data rate, modulation and coding schemes, frame length, FER, and the protocol overhead, etc. With the help of a new backward sequential method, the optimum transmission energy distributions with respect to the EE and SE metrics are expressed as closed-form expressions of all practical system parameters. In addition, a balanced tradeoff between EE and SE is also provided by reducing the SE-normalized energy per bit. Both simulation and analytical results demonstrate that the proposed energy efficient system design can achieve significant energy savings over conventional systems.

The remainder of the paper is organized as follows. Section II presents the system model by considering the operations of the hardware, the physical layer, and the media access control layer. The optimum energy distributions with respect to EE and SE are studied in Sections III and IV, respectively. In Section V, the balanced tradeoff between EE and SE is studied. Numerical results

56

are given in Section VI, and Section VII concludes the paper.

4.3   System Model

4.3.1   Energy Consumption Model

Consider a transmitter and a receiver separated by a distance d. The information bits are divided into packets. There are L_b uncoded information bits and L_0 overhead bits in each packet. The information bits and overhead bits are encoded with a rate-r channel encoder. For a system with modulation level M, the number of symbols in each packet is $L_s = \frac{L_b + L_0}{r \log_2 M}$, where L_b is chosen in a way such that L_s is an integer.

The signal is transmitted through a flat Rayleigh fading channel. Due to the effects of channel fading and noise, the receiver might not be able to successfully recover the transmitted signal. In case of a detection failure in a system with HARQ, the receiver will send back to the transmitter a negative acknowledgement (NACK), upon receiving which the transmitter retransmits the original packet. The receiver detects the packet by performing maximum ratio combining (MRC) over all the retransmissions of the same packet. The detection is successful if the recovered packet passes a cyclic redundancy check (CRC), and the receiver sends an acknowledgement (ACK) packet to the transmitter. Otherwise, either an NACK will be sent back to the transmitter, or the packet will be discarded if the maximum number of retransmissions, K, is reached. The value of the maximum retransmission number K is determined by the delay constraint of the application. There are usually stringent delay constraints in real time multimedia communications, and packets exceeding the delay constraint are directly discarded.

The signals received by the receiver in the k-th transmission attempt can be represented as

$$y_{mk} = \sqrt{E_{rk}} h_{mk} x_m + z_{mk}, \ m = 1, 2, \cdots, L_s, \tag{4.1}$$

57

where E_{rk} is the average energy of a symbol observed at the receiver, $x_m \in \mathcal{S}$ is the m-th modulated symbol in a length-L_s packet, with $\mathcal{S}$ being the modulation constellation set with the cardinality $M = |\mathcal{S}|$, y_{mk}, h_{mk}, and z_{mk} are the received sample, the fading coefficient, and additive white Gaussian noise (AWGN) with single-sided power spectral density N_0, respectively. Similar to [9] and [10], it is assumed that the retransmissions of the same packet undergoes the same fading, *i.e.*, the fading coefficient is constant within the K transmissions of one packet, and changes independently from packet to packet. Such an assumption is reasonable for multimedia communications with stringent delay constraints.

The average transmission energy for each symbol at the transmitter in the k-th transmission is [1]

$$E_{sk} = E_{rk}G_d, \tag{4.2}$$

where $G_d = G_1 d^\kappa M_c$, d is the transmitter-receiver distance, κ is the path-loss exponent, G_1 is the gain factor (including path-loss and antenna gain) at a unit distance, and M_c is the link margin compensating the hardware process variations and other additive background noise or interference.

In addition to the actual transmission energy, we also need to consider the hardware energy per symbol consumed at both the transmitter and receiver, and it can be modeled as [1],

$$E_{ck} = \left(\frac{\xi_M}{\mu} - 1\right) E_{sk} + \frac{\beta}{R_s}, \tag{4.3}$$

where $R_s = \frac{1}{T_s}$ is the gross symbol rate, μ is the drain efficiency of the power amplifier, ξ_M is the peak-to-average power ratio (PAPR) of an M-ary modulation signal, β incorporates the effects of baseband processing, such as signal processing, encoding in the transmitter and decoding in the receiver, and it can be treated as a constant determined by the particular transceiver structure. For MQAM systems, $\xi_M \simeq 3(\sqrt{M} - \frac{1}{\sqrt{M}} + 1)$ for $M \geq 4$ [27].

From (4.2) and (4.3), the energy required to transmit an uncoded information bit during the k-th

transmission is

$$E_{tk} = \frac{L_s}{L_b}(E_{sk} + E_{ck}) = \gamma_k \eta_b + B, \tag{4.4}$$

where $\eta_b = \frac{L_b + L_0}{L_b} \frac{\xi_M}{\eta} \frac{N_0 G_d}{\eta}$ with $G_d = G_1 d^\kappa M_c$, $B = \frac{\beta}{R_b}$, $R_b = \frac{L_b}{L_s} R_s$ is the net bit rate of the uncoded

information bit, and $\gamma_k = \frac{E_{rk}}{r N_0 \log_2 M}$ is the average E_b/N_0 at the receiver for an uncoded bit (including

both information bit and overhead bit) in the k-th transmission.

4.3.2 Frame Error Rate

In a system with HARQ, a retransmission occurs when the received frame has an unrecoverable error.

Therefore the performance of an HARQ system is directly related to the FER. Denote the FER at

the k-th transmission as ε_i. With HARQ and Chase combining, ε_i is a function of $\{\gamma_k\}_{k=1}^i$. Let

p_k and e_k represent, respectively, the probability that a packet is successfully recovered at the k-th

transmission, and the probability that a packet is not correctly detected in the first k transmissions.

The probabilities can be represented by

$$p_k = (1 - \varepsilon_k) \prod_{i=1}^{k-1} \varepsilon_i, \tag{4.5}$$

$$e_k = \prod_{i=1}^{k} \varepsilon_i. \tag{4.6}$$

Combining (4.5) and (4.6) yields

$$p_k = e_{k-1} - e_k, \tag{4.7}$$

with $e_0 = 1$.

In a quasi-static Rayleigh fading channel, the FER of a coded system can be accurately approxi-

mated by [26]

$$\text{FER}(\gamma_b) \simeq 1 - \exp\left(-\frac{\gamma_\omega}{\gamma_b}\right), \tag{4.8}$$

where γ_b is the average E_b/N_0 at the receiver, and γ_ω is a threshold value that can be calculated with a linear approximation to frame length in the log-domain, as [3]

$$\gamma_\omega \simeq k_M \log(L_b + L_0) + b_M, \tag{4.9}$$

where k_M and b_M are determined by the modulation scheme and the actual channel code. The values of k_M and b_M for various modulation and coding combinations are given in [3, Table 1]. It is shown in [3] that the combination of (4.8) and (4.9) yields a very accurate approximation of the FER over a large range of modulation and coding schemes.

In a system with HARQ, each packet is detected by performing MRC over all the current and previous retransmissions of the same packet. Therefore, the E_b/N_0 at the receiver after the k-th retransmission is $\sum_{i=1}^{k} \gamma_i$. From (4.8), the FER, ε_k, can be expressed as

$$\varepsilon_k = 1 - \exp\left(-\frac{\gamma_\omega}{\sum_{i=1}^{k} \gamma_i}\right). \tag{4.10}$$

To ensure the QoS of the system, we impose an upper bound on the FER after the K-th retransmission as $\varepsilon_K \leq \delta$.

4.4 Optimum Energy Distribution to Maximize Energy Efficiency

The optimum energy distribution, i.e., the sequence of energy at different retransmissions, that can minimize the average energy per uncoded information bit is discussed in this section.

4.4.1 Optimum Energy Distribution

With the notations in (4.4)–(4.7), the average energy required to transmit one uncoded information bit in HARQ can be denoted as

$$E_t = \sum_{k=1}^{K} p_k \sum_{i=1}^{k} E_{ti} + e_K \sum_{k=1}^{K} E_{tk}. \tag{4.11}$$

Substituting (4.7) into (4.11), and switching the order of summation in the first term of (4.11), we have

$$E_t = \sum_{i=1}^{K-1} E_{ti} \sum_{k=i}^{K-1} (e_{k-1} - e_k) + e_{K-1} \sum_{k=1}^{K} E_{tk}. \tag{4.12}$$

The above equation can then be simplified to

$$E_t = E_{t1} + \sum_{k=2}^{K} E_{tk} e_{k-1}. \tag{4.13}$$

Substituting (4.4) and (4.6) into (4.13), we can get the average energy per information bit as

$$E_t = \eta_b \gamma_1 + B + \sum_{k=2}^{K} [\eta_b \gamma_k + B] \times \prod_{i=1}^{k-1} \varepsilon_i. \tag{4.14}$$

From (4.10) and (4.14), the average energy per information bit can be represented as an explicit function of γ_j as

$$E_t = \eta_b \gamma_1 + B + \sum_{k=2}^{K} (\eta_b \gamma_k + B) \prod_{i=1}^{k-1} \left[1 - \exp\left(-\frac{\gamma_\omega}{\sum_{j=1}^{i} \gamma_j} \right) \right]. \tag{4.15}$$

The constraint, $\varepsilon_K \leq \delta$, can be alternatively expressed as

$$\sum_{k=1}^{K} \gamma_k \geq -\frac{\gamma_\omega}{\log(1-\delta)}. \tag{4.16}$$

Therefore, the optimization problem can be represented as

$$\begin{aligned} \text{minimize} \quad & E_t \quad \text{with respect to} \quad \gamma_1, \gamma_2, \cdots, \gamma_K \geq 0, \\ \text{subject to} \quad & \sum_{k=1}^{K} \gamma_k \geq -\frac{\gamma_\omega}{\log(1-\delta)} \end{aligned} \tag{4.17}$$

with E_t specified in (4.15).

The optimization problem in (4.17) is non-linear and non-covex, due to the fact that E_t is not a convex function in $\boldsymbol{\gamma} = [\gamma_1, \cdots, \gamma_K]^T \in \mathcal{Y} = \{\boldsymbol{\gamma} | \sum_{k=1}^{K} \gamma_k \geq -\frac{\gamma_\omega}{\log(1-\delta)}, \gamma_k \geq 0, k = 1, \cdots, K\}$. However, it can be easily shown that E_t is continuously differentiable in $\mathcal{Y}$.

To solve the non-linear optimization problem, we first find the values of $\boldsymbol{\gamma}$ that satisfy the Karush–Kuhn–Tucker (KKT) conditions, which are necessary conditions for the solution of the optimization problem. The global optimality of the solutions that satisfy the KKT conditions will be shown in the next subsection.

For the optimization problem in (4.17) for the minimum E_t, the stationary objective function of the KKT conditions can be expressed by,

$$\psi(\boldsymbol{\gamma}, \lambda) = \eta_b \gamma_1 + B + \sum_{k=2}^{K} (\eta_b \gamma_k + B) \prod_{i=1}^{k-1} \varepsilon_i - \lambda \left[\sum_{k=1}^{K} \gamma_k + \frac{\gamma_\omega}{\log(1-\delta)} \right], \tag{4.18}$$

where λ is the Lagrangian multiplier, and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_K]$.

The first derivatives of $\psi(\boldsymbol{\gamma}, \lambda)$ with respect to γ_n, for $n = 1, \cdots, K$, are

$$\frac{\partial \psi(\boldsymbol{\gamma}, \lambda)}{\partial \gamma_n} = \eta_b \varphi_n - \sum_{k=n+1}^{K} (\eta_b \gamma_k + B) \sum_{l=n}^{k-1} \vartheta_l \prod_{\substack{i=1 \\ i \neq l}}^{k-1} \varepsilon_i - \lambda, \text{ for } n = 1, \cdots, K-1, \tag{4.19a}$$

$$\frac{\partial \psi(\boldsymbol{\gamma}, \lambda)}{\partial \gamma_K} = \eta_b \varphi_K - \lambda, \tag{4.19b}$$

where

$$\varphi_n = \begin{cases} 1, & n = 1 \\ \prod_{i=1}^{n-1} \varepsilon_i, & \text{for } n = 2, \cdots, K. \end{cases} \tag{4.20}$$

and

$$\vartheta_l = \frac{(1 - \varepsilon_l)\gamma_\omega}{(\sum_{j=1}^{l} \gamma_j)^2}. \tag{4.21}$$

Setting (4.19b) to 0 yields $\lambda = \eta_b \varphi_K > 0$, which satisfies the KKT dual feasibility condition.

Based on the KKT complementary slackness condition, $\lambda \left(\sum_{k=1}^{K} \gamma_k + \frac{\gamma_\omega}{\log(1-\delta)} \right) = 0$, since $\lambda > 0$, we have

$$\gamma_0 \triangleq \sum_{k=1}^{K} \gamma_k = -\frac{\gamma_\omega}{\log(1-\delta)}. \tag{4.22}$$

With (4.22), the KKT primal feasibility condition is also satisfied. Therefore, the solution to the equation system described in (4.19) and (4.22) satisfies the KKT conditions. Consequently, the

global optimum value $\boldsymbol{\gamma}^*$ must be one of the solutions to the equation system in (4.19) and (4.22). We will first solve the equation system that satisfies the KKT conditions, and then show the solution is actually the global optimum one.

Even though the equation system in (4.19) and (4.22) can be solved numerically, the complexity is quite high given that the equations are non-linear. In addition, numerical solutions do not provide information regarding the global optimality of the results. We provide an analytical solution to the problme in the next subsection.

4.4.2   Optimization via Backward Sequential Calculations

A backward sequential optimization method is proposed to analytically identify the solutions that satisfy the KKT conditions. The method calculates γ_n by using the values of γ_i, for $i = n+1, \cdots, K$. This enables the backward calculations of γ_n, i.e., γ_K is calculated first, and γ_n is calculated by using $\{\gamma_k\}_{k=n+1}^K$, for $n = 2, \cdots, K-1$.

Since $\partial\psi(\boldsymbol{\gamma}, \lambda)/\partial\gamma_n = 0$, we have $\partial\psi(\boldsymbol{\gamma}, \lambda)/\partial\gamma_{n-1} - \partial\psi(\boldsymbol{\gamma}, \lambda)/\partial\gamma_n = 0$. From (4.19), the expression of $\frac{\partial\psi(\boldsymbol{\gamma}, \lambda)}{\partial\gamma_{n-1}} - \frac{\partial\psi(\boldsymbol{\gamma}, \lambda)}{\partial\gamma_n}$ for different values of n are

$$\frac{\partial E_t}{\partial\gamma_n} - \frac{\partial E_t}{\partial\gamma_{n+1}} = \eta_b(1 - \varepsilon_n)\varphi_n - (\eta_b\gamma_{n+1} + B)\vartheta_n\varphi_n - \sum_{k=n+2}^K (\eta_b\gamma_k + B)\vartheta_n \prod_{i=1, i\neq n}^{k-1} \varepsilon_i,$$

$$\text{for } n = 1, \cdots, K-2, \tag{4.23a}$$

$$\frac{\partial E_t}{\partial\gamma_{K-1}} - \frac{\partial E_t}{\partial\gamma_K} = \eta_b(1 - \sigma_{K-1})\varphi_{K-1} - (\eta_b\gamma_K + B)\vartheta_{K-1}\varphi_{K-1}. \tag{4.23b}$$

4.4.2.1   Calculation of γ_K

To initiate the backward sequential calculation, we would need the value of γ_K, which can be calculated by setting (4.23b) to 0, and the result is

$$\sum_{k=1}^{K-1} \gamma_k = \sqrt{\gamma_\omega\left(\gamma_K + \frac{B}{\eta_b}\right)}. \tag{4.24}$$

Combining (4.22) with (4.24) leads to

$$\sqrt{\gamma_\omega \left(\gamma_K + \frac{B}{\eta_b}\right)} + \gamma_K = -\frac{\gamma_\omega}{\log(1-\delta)}. \tag{4.25}$$

Solving the above equation yields

$$\gamma_K = \gamma_0 + \frac{\gamma_\omega}{2} - \sqrt{\gamma_0\gamma_\omega + \frac{\gamma_\omega^2}{4} + \gamma_\omega\frac{B}{\eta_b}}. \tag{4.26}$$

It should be noted that the value of γ_K in (4.26) is unique, because the other root of the second-order linear equation of (4.25) is $\gamma_0 + \frac{\gamma_\omega}{2} + \sqrt{\gamma_0\gamma_\omega + \frac{\gamma_\omega^2}{4} + \gamma_\omega\frac{B}{\eta_b}}$, which violates (4.22) with $\gamma_k \geq 0$ for $k = 1, 2, \cdots, K$.

4.4.2.2 Calculation of γ_n, for $n = 1, \cdots, K - 1$

Setting (4.23a) to 0 leads to

$$\sum_{k=1}^{n} \gamma_k = \sqrt{\gamma_\omega} \times \sqrt{\gamma_{n+1} + \frac{B}{\eta_b} + \sum_{l=n+2}^{K}\left(\gamma_l + \frac{B}{\eta_b}\right)\prod_{i=n+1}^{l-1}\varepsilon_i}, \text{ for } n = 1, 2, \cdots, K-2. \tag{4.27}$$

Eqns. (4.24) and (4.27) can be written in a unified form as

$$\sum_{k=1}^{n} \gamma_k = \sqrt{\gamma_\omega} \times \sqrt{\gamma_{n+1} + C_n}, \text{ for } n = 1, 2, \cdots, K-1, \tag{4.28}$$

where

$$C_n = \begin{cases} \frac{B}{\eta_b} + \sum_{l=n+2}^{K}\left(\gamma_l + \frac{B}{\eta_b}\right)\prod_{i=n+1}^{l-1}\varepsilon_i, & n = 1, \cdots, K-2 \\[2mm] \frac{B}{\eta_b}, & n = K-1 \end{cases} \tag{4.29}$$

From (4.28), and based on the fact that $\gamma_n = \sum_{k=1}^{n}\gamma_k - \sum_{k=1}^{n-1}\gamma_k$, we have

$$\gamma_n = \sqrt{\gamma_\omega} \times \left(\sqrt{\gamma_{n+1} + C_n} - \sqrt{\gamma_n + C_{n-1}}\right), \text{ for } n = 2, \cdots, K-1. \tag{4.30}$$

In the above equation, the calculation of γ_n requires the knowledge of C_{n-1} and C_n, which in turn depend on σ_k and $\sum_{i=1}^{k}\gamma_i$, for $k = n, \cdots, K$, as in (4.10). To enable the backward sequential calculation, we obtain an alternative expression of the FER by using (4.22) and the fact that

64

$\sum_{i=1}^{k} \gamma_i = \sum_{i=1}^{K} \gamma_i - \sum_{i=k+1}^{K} \gamma_i$, and the result is

$$\varepsilon_k = 1 - \exp \left[\frac{\gamma_\omega}{\frac{\gamma_\omega}{\log(1-\delta)} + \sum_{i=k+1}^{K} \gamma_i} \right]. \tag{4.31}$$

From (4.29) and (4.31), it can be seen that C_n and C_{n-1} depend only on $\{\gamma_k\}_{k=n+1}^{K}$. Therefore, (4.30) is a second-order linear equation in γ_n, which can be expressed as

$$\gamma_n^2 - \left[\gamma_\omega + 2\sqrt{\gamma_\omega(\gamma_{n+1} + C_n)} \right] \gamma_n + \gamma_\omega(\gamma_{n+1} + C_n - C_{n-1}) = 0. \tag{4.32}$$

The solution of (4.32) can be expressed in a closed-form as

$$\gamma_n = \frac{1}{2}\gamma_w - \sqrt{\frac{\gamma_w^2}{4} + \gamma_w\sqrt{\gamma_w(\gamma_{n+1} + C_n)} + \gamma_w C_{n-1}} + \sqrt{\gamma_w(\gamma_{n+1} + C_n)}, \text{ for } n = 2, \cdots, K - 1. \tag{4.33}$$

Appendix 4.9.1 shows that (4.33) is the unique solution to (4.30).

In (4.33), γ_n is expressed as an explicit function of $\{\gamma_k\}_{k=n+1}^{K}$, with C_n calculated from (4.29) and (4.31). Therefore, γ_n can be calculated by using $\{\gamma_k\}_{k=n+1}^{K}$, for $n = 2, \cdots, K - 1$. Finally, γ_1 can be calculated from (4.22) as

$$\gamma_1 = -\frac{\gamma_w}{\log(1 - \delta)} - \sum_{k=2}^{K} \gamma_k. \tag{4.34}$$

With the above analytical results, the solution for γ_i, $i = 1, 2, \cdots, K$ in (4.26), (4.33), and (4.34) is unique. Since the solutions are the necessary conditions for the optimality and they are unique, the solution yields the global optimum operation point to the optimization problem.

The backward sequential calculation method is summarized in Algorithm 1.

Once the optimum values of $\{\gamma_n\}_{n=1}^{K}$, which are the average E_b/N_0 at the receiver, are identified, then the optimum energy per symbol at the transmitter can be calculated as (c.f. (4.2))

$$E_{sk} = \gamma_k \times N_0 \times r \times \log_2 M \times G_d. \tag{4.35}$$

65

Algorithm 1 Backward sequential calculation

1: Input: K and δ.
2: Calculate γ_K with (4.22) and (4.26).
3: **for** $n = K - 1, K - 2, \cdots, 1$ **do**
4:    Calculate ε_n from (4.31),
5:    Calculate C_n and C_{n-1} from (4.29),
6:    Calculate γ_n from (4.33).
7: **end for**
8: Calculate γ_1 from (4.34).
9: Output: $\boldsymbol{\gamma}$.

4.5 Optimum Energy Distribution to Maximize Spectral Efficiency

In this section, the optimum energy distribution that can maximize the average spectral efficiency, η_s, is identified.

If the transmission is successful during the k-th transmission attempt, then the net data rate would be

$$R_{bk} = \frac{L_b}{kT_0} = \frac{L_b}{k\frac{L_b + L_0}{rR_s \log_2 M}} = \frac{L_b}{L_b + L_0}\frac{rR_s \log_2 M}{k}, \tag{4.36}$$

where T_0 is the duration of one frame, and R_s is the gross symbol rate. If the transmission is successful during the k-th attempt, then the corresponding spectral efficiency is

$$\eta_k = \frac{R_{bk}}{(1 + \alpha)R_s} = \frac{L_b}{L_b + L_0}\frac{r \log_2 M}{k(1 + \alpha)}, \tag{4.37}$$

where α is the roll-off factor of the pulse shaping filter, and it expands the bandwidth of each symbol from R_s to $(1 + \alpha)R_s$.

The average spectral efficiency of an HARQ system can then be calculated by

$$\eta_s = \sum_{k=1}^{K} \eta_k p_k = \sum_{k=1}^{K} \eta_k (e_{k-1} - e_k). \tag{4.38}$$

Substituting (4.6), (4.10) and (4.37) into the above equation and simplifying yield

$$\eta_s = \frac{L_b}{L_b + L_0}\frac{r \log_2 M}{1 + \alpha} \times \sum_{k=1}^{K} \frac{1}{k} \times \exp\left(-\frac{\gamma_\omega}{\sum_{j=1}^{k} \gamma_j}\right) \prod_{i=1}^{k-1}\left[1 - \exp\left(-\frac{\gamma_\omega}{\sum_{j=1}^{k} \gamma_j}\right)\right]. \tag{4.39}$$

The first derivatives of η_s with respect to γ_n, for $n = 1, \cdots, K$, is

$$\frac{\partial \eta_s}{\partial \gamma_n} = \frac{\Lambda \vartheta_n \varphi_n}{n} + \Lambda \sum_{k=n+1}^{K} \left(\frac{\vartheta_k}{k} \prod_{i=1}^{k-1} \varepsilon_i - \frac{1 - \varepsilon_k}{k} \sum_{l=n}^{k-1} \vartheta_l \prod_{i=1, i \neq l}^{k-1} \varepsilon_i \right), \quad \text{for} \quad n = 1, 2, \cdots, K-1, \tag{4.40a}$$

$$\frac{\partial \eta_s}{\partial \gamma_K} = \frac{\Lambda \vartheta_K \varphi_K}{K}, \tag{4.40b}$$

where ε_i, φ_n and ϑ_l are defined in (4.10), (4.20), and (4.21), respectively, and $\Lambda = \frac{L_b}{L_b + L_0} \frac{r \log_2 M}{1 + \alpha}$. We have the following proposition regarding the values of the first derivatives of η_s.

Lemma 1: For an HARQ system that employs the MRC receiver, if the FER is given in (4.10), then $\frac{\partial \eta_s}{\partial \gamma_1} > \frac{\partial \eta_s}{\partial \gamma_2} > \cdots > \frac{\partial \eta_s}{\partial \gamma_K} > 0$.

Proof: The proof is in Appendix 4.9.2. ■

From Lemma 1, it can be seen that η_s is an increasing function in γ_n, for $n = 1, \cdots, K$, and the slope of η_s with respect to γ_n decreases as n increases. That is, allocating additional energy to the n-th transmission round yields a larger SE gain compared to adding the same amount of energy to the k-th round with $k > n$. Therefore, we have the following corollary regarding the optimum energy distribution that can maximize the SE.

Proposition 1: Consider an HARQ system that employs MRC at the receiver. Under the constraint, $\sum_{k=1}^{K} \gamma_k = \gamma_0 = -\frac{\gamma_\omega}{\log(1-\delta)}$, the average SE can be maximized by employing the energy distribution $\boldsymbol{\gamma} = [\gamma_0, 0, \cdots, 0]^T$. On the other hand, the average SE can be minimized by the energy distribution $\boldsymbol{\gamma} = [0, \cdots, 0, \gamma_0]^T$. The maximum and minimum achievable SE is given as follows

$$\max(\eta_s) = \frac{L_b}{L_b + L_0} \frac{r \log_2 M}{1 + \alpha} (1 - \delta), \tag{4.41}$$

$$\min(\eta_s) = \frac{L_b}{L_b + L_0} \frac{r \log_2 M}{1 + \alpha} \frac{1 - \delta}{K}. \tag{4.42}$$

Proof: The proof is in Appendix 4.9.3. ■

4.6 Energy Distribution to Achieve a Balanced Tradeoff between EE and SE

The transmission energy distribution that can reduce the SE-normalized energy per bit, $E_m = \frac{E_t}{\eta_s}$, is discussed in this section. The metric E_m can be reduced by either decreasing E_t or increasing η_s. Therefore, reducing E_m can achieve a balanced tradeoff between the EE and SE.

To facilitate the analysis, the solution is performed in the log-domain by considering the metric $\log E_m = \log E_t - \log \eta_s$. Since $\log x$ is a monotonic increasing function in x for $x > 0$, reducing $\log E_m$ yields the same results as reducing E_m. Since E_t is lower bounded, and η_s are bounded both above and below, $\log E_m$ is lower bounded.

We propose to find the optimum energy distribution with respect to E_m by solving the equations, $\frac{\partial \log E_m}{\partial \gamma_n} - \frac{\partial \log E_m}{\partial \gamma_{n+1}} = 0$, for $n = 1, \cdots, K-1$, which can be alternatively written as

$$\frac{1}{E_t}\left(\frac{\partial E_t}{\partial \gamma_n} - \frac{\partial E_t}{\partial \gamma_{n+1}}\right) = \frac{1}{\eta_s}\left(\frac{\partial \eta_s}{\partial \gamma_n} - \frac{\partial \eta_s}{\partial \gamma_{n+1}}\right). \tag{4.43}$$

It should be noted that the solutions to (4.43) might lead to local optimum points. However, our extensive numerical evaluations indicate that the solutions to (4.43) always coincide with the global optimum point obtained through exhaustive search for all system parameters considered in this paper.

Substituting (4.23) and (4.40) into the above equation and simplifying, we have

$$\sum_{k=1}^{n}\gamma_i = \sqrt{\gamma_\omega}\sqrt{\frac{\Lambda E_t}{\eta_s \eta_b}\left(\frac{1}{n} - \frac{1-\sigma_{n+1}}{n+1} - D_n\right) + \gamma_{n+1} + C_n}, \text{ for } n = 1, 2, \cdots, K-1, \tag{4.44}$$

where C_n is defined in (4.29), and

$$D_n = \begin{cases} \sum_{l=n+2}^{K}\frac{1-\varepsilon_l}{l}\prod_{i=n+1}^{l-1}\varepsilon_i, & n = 1, \cdots, K-2, \\ 0, & n = K-1. \end{cases} \tag{4.45}$$

From (4.22), (4.44) and the identity, $\gamma_n = \sum_{k=1}^{n}\gamma_k - \sum_{k=1}^{n-1}\gamma_k$, we have

$$\gamma_K = \gamma_0 - \sqrt{\gamma_\omega}\sqrt{\frac{\Lambda E_t}{\eta_s \eta_b}\left(\frac{1}{K-1} - \frac{1-\sigma_K}{K}\right) + \gamma_K + \frac{B}{\eta_b}}, \tag{4.46}$$

68

and

$$\gamma_n = \left[\sqrt{ \frac{\Lambda E_t}{\eta_s \eta_b} \left(\frac{1}{n} - \frac{1-\sigma_{n+1}}{n+1} - D_n \right) + \gamma_{n+1} + C_n } - \sqrt{ \frac{\Lambda E_t}{\eta_s \eta_b} \left(\frac{1}{n-1} - \frac{1-\sigma_n}{n} - D_{n-1} \right) + \gamma_n + C_{n-1} } \right] \sqrt{\gamma_\omega},$$

$$\text{for} \quad n = 2, 3, \cdots, K-1, \quad (4.47)$$

Eqns. (4.46) and (4.47) are linear second order equations in γ_K or γ_n. They can be solved and expressed as closed-form functions of E_t, η_s, γ_{n+1}, and $\{\varepsilon_k\}_{k=n}^K$, which in turn depend on $\{\gamma_k\}_{k=n+1}^K$ as in (4.31). In order to implement the backward sequential calculation as in Section 4.4, we need to know E_t and η_s, the values of which depend on the optimum distribution $\{\gamma_k\}_{k=1}^K$ that are to be solved.

To address this problem, we propose to identify the feasible values of $\{\gamma_k\}_{k=1}^K$ by combining the backward sequential calculation with an iterative algorithm. In the iterative algorithm, the values of E_t and η_s are initialized by assuming equal energy distribution, i.e., $\gamma_1 = \cdots = \gamma_K = \frac{\gamma_0}{N}$. Then the backward sequential calculation can be used to calculate γ_n with (4.46) for $n = K$, with (4.47) for $n = 2, \cdots, K-1$, and with (4.34) for $n = 1$, where γ_K is calculated first and γ_1 is calculated last. During the backward sequential calculation, once the value of γ_m is obtained, we update the values of E_t with (4.15) and η_s with (4.39) by setting $\gamma_n = \frac{1}{m-1}(\gamma_0 - \sum_{k=m}^K \gamma_m)$, for $n = 1, \cdots, m-1$. The updated values of E_t and η_s are then used for the calculation of γ_{m-1}. At the end of the backward sequential calculation, E_t and η_s are calculated by using the values of $\{\gamma_n\}_{n=1}^K$ obtained in this iteration. The new values of E_t and η_s are then used as the initial values for the calculation of γ_K at the beginning of the next iteration. The above procedure is performed iteratively until stopping criteria are satisfied.

The stopping criteria of the algorithm are determined by comparing the values of E_t and η_s at the beginning and end of one iteration. If we let E_{t0} and η_{s0} denote the values at the beginning of an iteration, and E_t and η_s denote the values at the end of the same iteration, then the algorithm

69

stops if $|E_t - E_{t0}| < \epsilon_E$ and $|\eta_s - \eta_{s0}| < \epsilon_\eta$, with ϵ_E and ϵ_η being very small positive numbers. In the numerical results, we use $\epsilon_E = \epsilon_\eta = 10^{-6}$. The numerical results demonstrated that the feasible solution can usually be obtained in three iterations.

The iterative backward sequential calculation algorithm is summarized as follows in algorithm 2.

Algorithm 2 Iterative backward sequential calculation algorithm

---

1: Input: K, δ, ϵ_E, and ϵ_η.
2: Set $\gamma_n = \frac{\gamma_0}{K} = -\frac{1}{K} \frac{\gamma_\omega}{\log(1-\delta)}$ for $n = 1, 2, \cdots, K$.
3: Calculate E_t with (4.15) and η_s with (4.39).
4: Set $E_{t0} = +\infty$ and $\eta_{s0} = +\infty$.
5: **while** $|E_t - E_{t0}| < \epsilon_E$ and $|\eta_s - \eta_{s0}| < \epsilon_\eta$ **do**
6:     Set $E_{t0} \leftarrow E_t$, $\eta_{s0} \leftarrow \eta_s$,
7:     **for** $n = K, K-1, \cdots, 1$ **do**
8:         Calculate γ_n from (4.46) for $n = K$, or (4.47) for $2 \leq n \leq K-1$, or (4.34) for $n = 1$,
9:         Update $\gamma_m = -\frac{1}{n-1} \frac{\gamma_\omega}{\log(1-\delta)} - \sum_{k=n}^{K} \gamma_k$ for $1 \leq m \leq n-1$,
10:       Update E_t with (4.15) and η_s with (4.39).
11:     **end for**
12: **end while**
13: Output: $\boldsymbol{\gamma}$.

Once the values of $\{\gamma_n\}_{n=1}^K$ are obtained, we can calculate the transmission energy for each transmission attempt as in (4.35).

4.7   Numerical Results

Numerical results are presented in this section to demonstrate the proposed energy and/or spectral efficient transmissions for various HARQ systems. Most of the simulation parameters follow [1]: $\mu = 3.5$, $\beta = 310.014$ mW, $N_0/2 =$ -174 dBm/Hz, $G_1 =$30 dB, $\kappa = 3.5$, and $M_c = 40$ dB. The rest of the parameters are $L_0 = 48$ bits, $L_b = 2048$ bits, $R_b = 2$ Mbps, and $d = 100$ m. The modulation scheme is QPSK. The channel code is a rate $r = \frac{1}{2}$ convolutional code with the generator polynomial $[171, 133]_8$ and a constraint length 7. The parameters $k_M = 0.2066$ and $b_M = 0.14$ are used in the

linear threshold approximation for the FER as in (4.9) [3].

The performance of the systems with the various optimization criteria will be compared to those with equal energy used for all retransmissions. For fairness of comparison, the same FER upper bound δ is enforced for all systems. For systems with equal energy retransmissions, from (4.16), the average $\frac{E_b}{N_0}$ observed by the receiver during each retransmission is

$$\gamma'_k = -\frac{1}{K}\frac{\gamma_\omega}{\log(1-\delta)}, \quad \text{for} \quad k = 1, 2, \cdots, K. \tag{4.48}$$

Fig. 4.1 shows the optimum energy distribution, $\{\gamma_k\}_{k=1}^{K}$, that minimizes the average energy per bit, E_t, when the maximum number of retransmissions is set to $K = 6$. As expected, the optimum values of γ_k varies with respect to the retransmission round index k. For systems considered in this example, the second round uses the least amount of energy and the last round uses the most. The optimum systems use significantly less energy compared to their equal energy counterparts during the first five transmission attempts, but use more energy during the last transmission attempt. In this example, the probability of using the sixth transmission attempt is $e_5 = 2 \times 10^{-4}$ for $\delta = 10^{-3}$ and $e_5 = 2 \times 10^{-5}$ for $\delta = 10^{-4}$. Therefore overall the optimum system requires much less energy than its equal energy counterpart, as evident in the results in Fig. 4.2.

Fig. 4.2 shows the total average energy per information bit, E_t, as a function of the maximum number of retransmissions, K, under various values of δ. The optimum energy distributions are obtained by minimizing E_t. Systems with optimum energy distributions achieve significant energy savings compared to their equal energy counterparts. When $K = 6$, systems with optimum energy distribution outperform their equal energy counterparts by 16.9 dB, 26.8 dB and 36.8 dB, for $\delta = 10^{-3}$, 10^{-4}, and 10^{-5}, respectively. The performance of systems with optimum energy distributions and different outage probabilities converge when $K > 6$, i.e., systems with $\delta = 10^{-3}$ and $\delta = 10^{-5}$ have similar E_t. Therefore, the optimum energy distribution scheme can greatly improve the outage
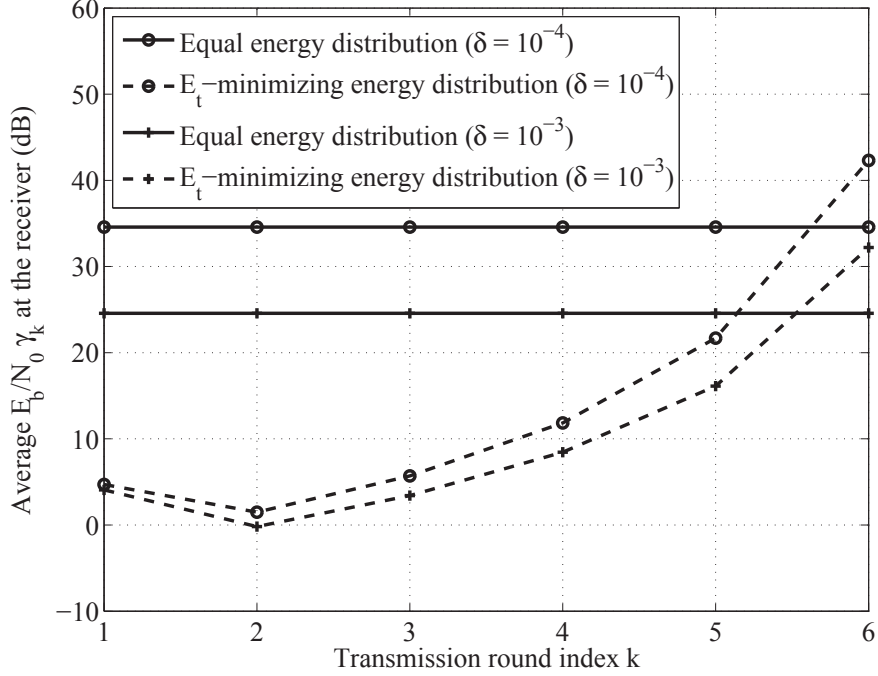
Figure 4.1: Optimum energy distribution of γ_k as a function of the transmission round index ($K = 6$).

performance without requiring extra energy when K is sufficiently large. On the other hand, for systems with equal energy retransmissions, 10 dB extra energy per information bit is needed to improve the system outage by a factor of 10, regardless of K.

Fig. 4.3 shows the impacts of limiting δ on average energy per information bit E_t with different values of K. The optimum energy distributions are obtained by minimizing E_t. For both equal energy and optimum energy systems, E_t decreases with δ, but with different slopes. The performance difference between the optimum and equal energy systems increases as δ decreases. For the optimum system with $K = 6$, improving the system outage from 10^{-1} to 10^{-5} only requires 3.6 dB additional energy per information bit; yet an extra 38.1 dB is required to achieve the same performance improvement for its equal energy counterpart. In addition, for the optimum system, the absolute value of the slope of the E_t-δ curve decreases as K increases, i.e., less extra energy is needed to improve the system outage when K is larger. The performance gain is achieved at the cost of a longer delay
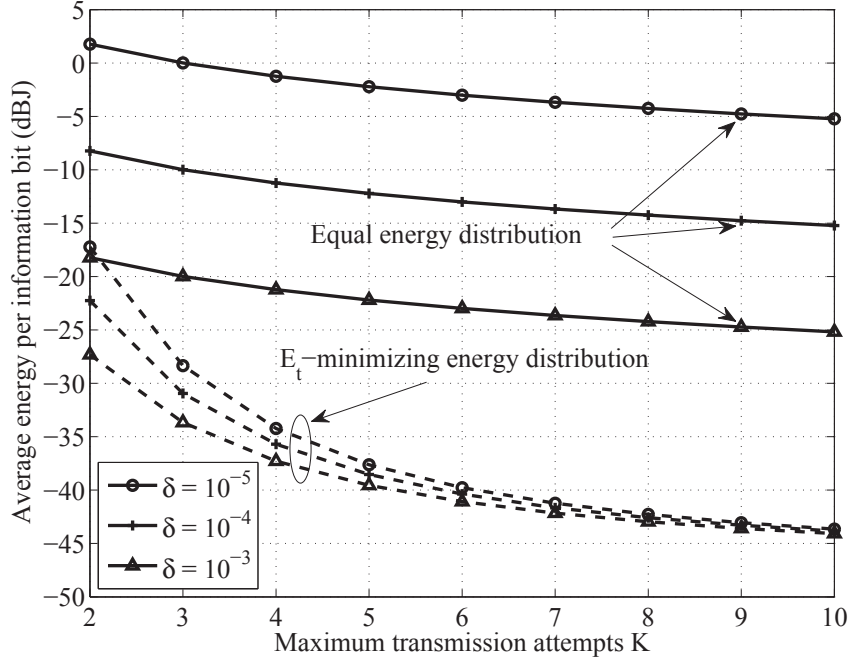
Figure 4.2: Average energy per information bit versus maximum number of transmission attempts K.

with a larger K. Therefore, for a specific δ, the value of K should be determined by considering both the energy and delay requirements.

Next we compare the performance of systems employing different optimization metrics, i.e., minimizing the average energy per information bit E_t, maximizing the average SE η_s, or reducing the SE-normalized energy per information bit E_m. Figs. 4.4 and 4.5 show, respectively, the average SE, η_s, and average energy per information bit, E_t, as functions of K, under various system configurations. The outage probability is set at $\delta = 10^{-3}$. The system that maximizes η_s allocates all the energy to the first transmission attempt, therefore its performance is independent of K. For the equal energy distribution, it achieves the second best SE (0.8 bps/Hz) as in Fig. 4.4, but at the cost of significant EE loss as shown in Fig. 4.5. For example, when $K = 4$, the SE of the equal energy distribution is 0.1 bps/Hz better than that of the E_m-reducing scheme as in Fig. 4, but at the cost of an energy loss of 16 dB as shown in Fig. 5. In Fig. 4.4, the η_s-maximizing energy distribution
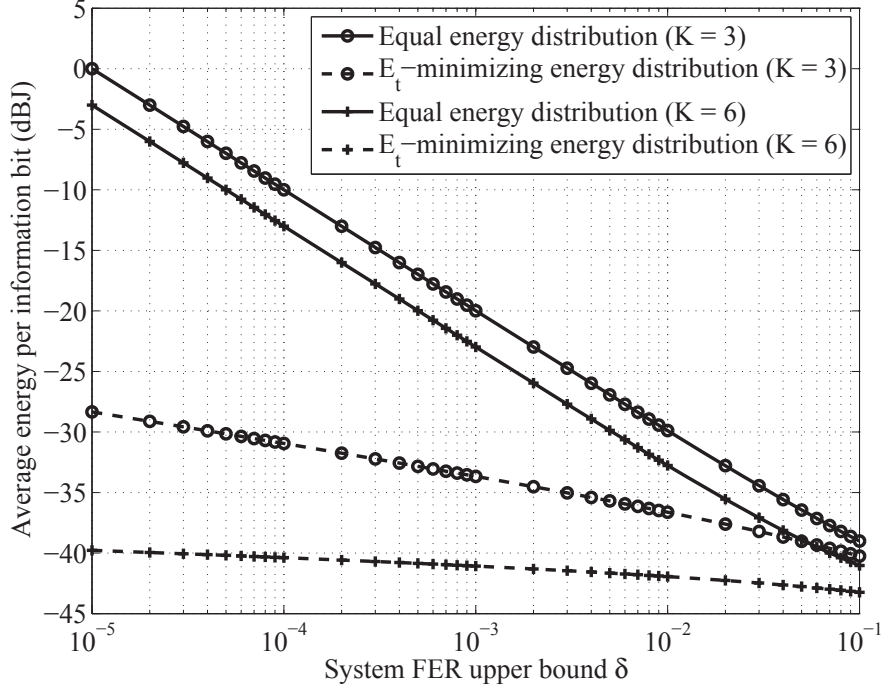
Figure 4.3: Average energy per information bit versus δ.

achieves the best SE as expected, followed by the equal energy system, the E_m-reducing system, and the E_t-minimizing system. On the other hand, the order of the four systems is reversed in terms of EE in Fig. 4.5, where the E_t-minimizing system has the best EE, followed by the E_m-reducing system, the equal energy system, and the η_s-maximizing system.

The E_m-reducing system achieves a balanced tradeoff between EE and SE. The SE of the E_m-reducing system is consistently better than that of the E_t-minimizing system, yet the two systems maintain a similar EE. At $K = 10$, the E_m-reducing system achieves an additional 0.11 bps/Hz, or 26.7%, SE compared to the E_t-minimizing system with only negligible cost in EE. For E_m-reducing or E_t-minimizing systems, the SE decreases and the EE increases as K increases. The SE gap between the E_t-minimizing and E_m-reducing systems grows as K increases.
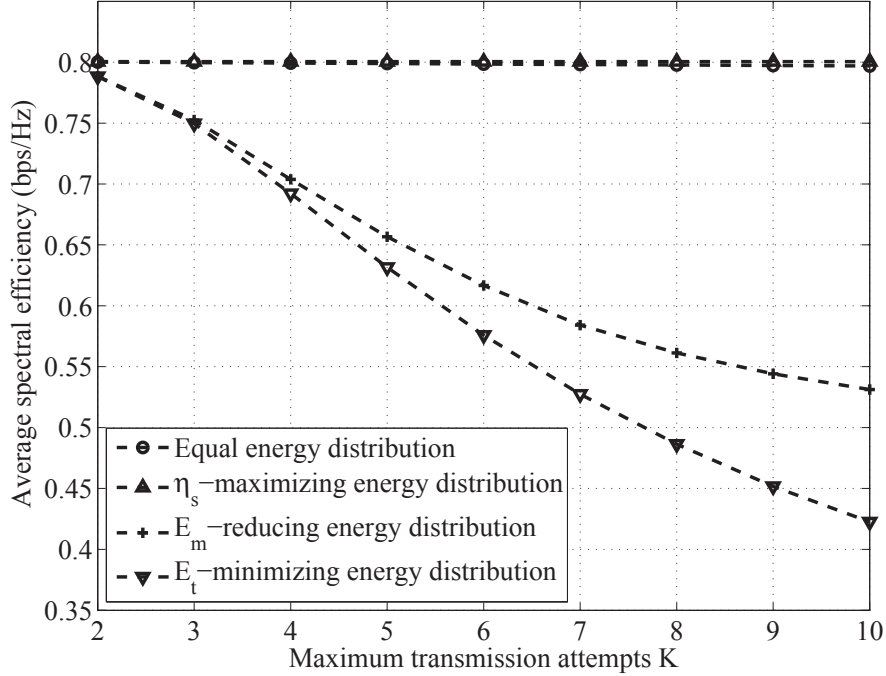
Figure 4.4: Average SE versus maximum number of transmission attempts K under various system configurations.

4.8   Conclusions

The optimum energy distributions for delay-constrained transmissions employing HARQ with Chase combining have been discussed in this paper. Three optimization metrics have been studied: the average energy per information bit E_t, the average spectral efficiency η_s, and the SE-normalized average energy per information bit $E_m = \frac{E_t}{\eta_s}$. The optimum energy distributions with respect to the EE and SE have been derived through analytical studies. An algorithm has been provided to reduce the E_m to achieve a balanced tradeoff between EE and SE. The feasible energy distributions for the E_t-minimizing and E_m-reducing systems are obtained with a new backward sequential calculation method, where the optimum energy is expressed as closed-form expressions of a large number of system parameters. Numerical results demonstrated that the proposed schemes can achieve significant performance gains over conventional HARQ systems with equal energy distribution. When the
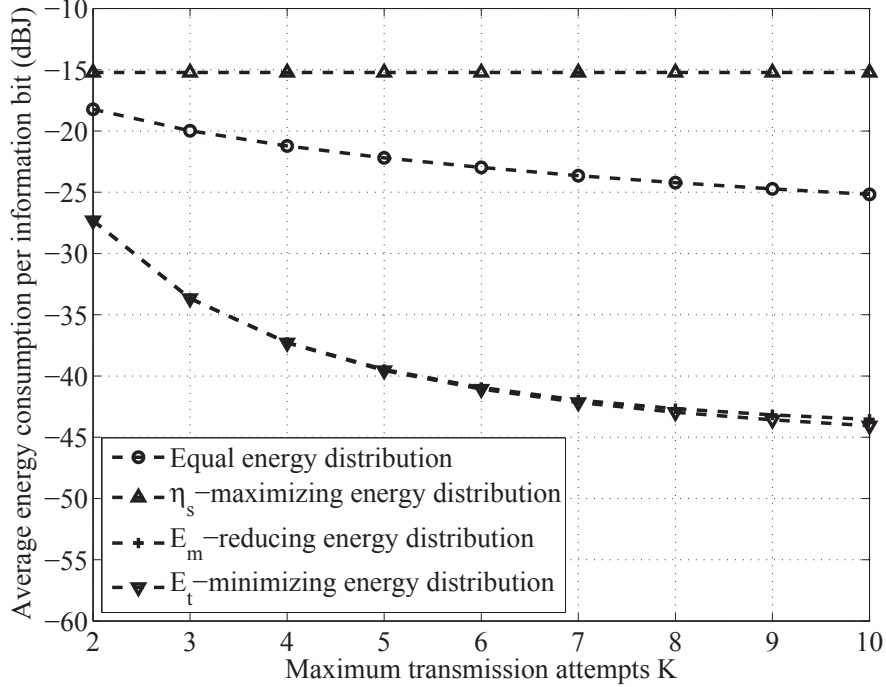
Figure 4.5: Average energy consumption per information bit versus maximum number of transmission attempts K under various system configurations.

maximum number of retransmissions is $K = 10$, the E_t-minimizing system outperformed the η_s-maximizing system by as much as 29.1 dB, but at the cost of 0.39 bps/Hz SE loss. The E_m-reducing system achieves a balanced tradeoff between EE and SE. The E_m-reducing system can increase the SE of the E_t-minimizing system by as much as 26.7% yet with only negligible cost in terms of EE.

The three different design metrics correspond to different tradeoff points on the EE-SE tradeoff curves, and they can be applied based on different system design requirements. The E_t-minimizing scheme can be used for applications with extremely stringent energy constraints, such as many wireless sensor networks designed for environment monitoring or structure health monitoring, which need to operate over a long period of time with very limited energy supplies, but only require a low or moderate data rate due to the slow changing nature of the monitored objects. The η_s-maximizing scheme should be used in applications demanding high SE, such as multimedia communications, at the cost of a higher energy consumption. The E_m-reducing scheme achieves a balanced tradeoff between

EE and SE. It can achieve a higher SE than the E_t-minimizing scheme, without incurring apparent energy cost. Thus the E_m-reducing scheme can be used for a wide range of wireless applications.

4.9 Appendix of Proofs

4.9.1 Proof of Uniqueness of (4.33)

We will show here that (4.33) is the unique solution to (4.30). The uniqueness is proved in two steps. First, it will be shown that there is only one solution to (4.30). Second, it will be shown that (4.33) is a valid solution to (4.30).

The left hand side of (4.30) is a strictly increasing function in γ_n with values ranging from 0 to ∞. Denote the right hand side of (4.30) as $f(\gamma_n)$. Since C_n, C_{n-1}, γ_ω, and γ_{n+1} are all independent of γ_n, $f(\gamma_n)$ is a strictly decreasing function in γ_n, with values ranging from $f(0) = \sqrt{\gamma_\omega} \times \left(\sqrt{\gamma_{n+1} + C_n} - \sqrt{C_{n-1}} \right)$ to $-\infty$ as γ_n goes from 0 to ∞. Since $C_n > C_{n-1}$, there is $f(0) > 0$. Therefore, there exists one and only one solution to (4.30).

The second order linear equation in (4.32) is obtained as $\gamma_n^2 = f^2(\gamma_n)$, which has two solutions, $\gamma_n^* = f(\gamma_n^*)$, and $-\gamma_n' = f(\gamma_n')$, and the valid solution should be γ_n^*. One solution to (4.32) is given in (4.33), and the other solution is $\frac{1}{2}\gamma_w + \sqrt{\frac{\gamma_w^2}{4} + \gamma_w \sqrt{\gamma_w(\gamma_{n+1} + C_n)} + \gamma_w C_{n-1}} + \sqrt{\gamma_w(\gamma_{n+1} + C_n)}$. It can be easily shown that both solutions are positive. Since $f(\gamma_n)$ is a strictly decreasing function in γ_n, we have $\gamma_n^* < \gamma_n'$ because $f(\gamma_n') < 0 < f(\gamma_n^*)$. Therefore γ_n^* should be the one given in (4.33), which is the smaller one of the two possible solutions to (4.32).

4.9.2 Proof of Lemma 1

From (4.40a) and (4.40b), we have

$$\frac{\partial \eta_s}{\partial \gamma_{K-1}} - \frac{\partial \eta_s}{\partial \gamma_K} = \Lambda \vartheta_{K-1} \varphi_{K-1} \left[\frac{1}{K-1} - \frac{1 - \varepsilon_K}{K} \right] > 0.$$

From (4.40a), we have

$$\frac{\partial \eta_s}{\partial \gamma_n} - \frac{\partial \eta_s}{\partial \gamma_{n+1}} = \Lambda \vartheta_n \varphi_n \times \left[\frac{1}{n} - \frac{1 - \sigma_{n+1}}{n+1} - \sum_{k=n+2}^{K} \frac{1 - \varepsilon_k}{k} \prod_{i=n+1}^{k-1} \varepsilon_i \right] \quad \text{for} \quad n = 1, 2, \cdots, K-2. \quad (4.49)$$

Setting $n = K - 2$ in the above equation yields

$$\frac{\partial \eta_s}{\partial \gamma_{K-2}} - \frac{\partial \eta_s}{\partial \gamma_{K-1}} = \Lambda \vartheta_{K-2} \varphi_{K-2} \times \left[\frac{1}{K-2} - \frac{1 - \sigma_{K-1}}{K-1} - \frac{1 - \varepsilon_K}{K} \sigma_{K-1} \right]. \quad (4.50)$$

The term in the bracket of (4.50) can be rewritten as

$$\left(\frac{1}{K-2} - \frac{1}{K-1} \right) (1 - \sigma_{K-1}) + \left(\frac{1}{K-2} - \frac{1 - \sigma_K}{K} \right) \sigma_{K-1},$$

which is greater than 0. Therefore $\frac{\partial \eta_s}{\partial \gamma_{K-2}} > \frac{\partial \eta_s}{\partial \gamma_{K-1}}$.

Proof by induction. Assume $\frac{\partial \eta_s}{\partial \gamma_n} - \frac{\partial \eta_s}{\partial \gamma_{n+1}} > 0$ for $n = m + 1$ with $1 \le m \le K - 2$. Then from (4.49) we have

$$\frac{1 - \sigma_{m+2}}{m+2} + \sum_{k=m+3}^{K} \frac{1 - \varepsilon_k}{k} \prod_{i=m+2}^{k-1} \varepsilon_i < \frac{1}{m+1}. \quad (4.51)$$

We next need to prove $\frac{\partial \eta_s}{\partial \gamma_n} - \frac{\partial \eta_s}{\partial \gamma_{n+1}} > 0$ for $n = m$. Simplifying (4.49) leads to

$$\frac{1}{\Lambda \vartheta_m \varphi_m} \left(\frac{\partial \eta_s}{\partial \gamma_m} - \frac{\partial \eta_s}{\partial \gamma_{m+1}} \right) = \left(\frac{1}{m} - \frac{1}{m+1} \right) (1 - \sigma_{m+1}) + \sigma_{m+1} \times \left\{ \frac{1}{m} - \sum_{k=m+2}^{K} \frac{1 - \varepsilon_k}{k} \prod_{i=m+2}^{k-1} \varepsilon_i \right\}. \quad (4.52)$$

Substituting the inequality of (4.51) into the above equation and simplifying, we have

$$\frac{\partial \eta_s}{\partial \gamma_m} - \frac{\partial \eta_s}{\partial \gamma_{m+1}} > \Lambda \vartheta_m \varphi_m \sigma_{m+1} \left(\frac{1}{m} - \frac{1}{m+1} \right) > 0. \quad (4.53)$$

In addition, since

$$\frac{\partial \eta_s}{\partial \gamma_K} = \frac{\Lambda}{K} \vartheta_K \prod_{i=1}^{K-1} \varepsilon_i > 0, \quad (4.54)$$

the proof is complete.

4.9.3 Proof of Proposition 1

We first prove the SE-maximizing energy distribution by contradiction. Assume $\gamma_n > 0$, for any $n = 2, \cdots, K$. Then we can always increase the average SE by moving an infinitesimal value $\Delta\gamma$ from the n-th transmission attempt to the first one, i.e., $\gamma'_n = \gamma_n - \Delta\gamma$ and $\gamma'_1 = \gamma_1 + \Delta\gamma$, because the net change in spectral efficiency would be

$$\left(\frac{\partial\eta_s}{\partial\gamma_1} - \frac{\partial\eta_s}{\partial\gamma_n}\right)\Delta\gamma > 0. \tag{4.55}$$

Therefore, to obtain the maximum average SE, the optimum energy distribution is $\gamma_1 = \gamma_0$ and $\gamma_n = 0$, for $n = 2, \cdots, K$. The result states that, in order to maximize the average SE, all the transmission energy should be allocated to the first transmission attempt under a total energy constraint. This is intuitive because more retransmissions mean a longer delay, which results in a smaller SE. The maximum average SE can then be calculated as $\eta_1 p_1 = \frac{L_b}{L_b+L_0}\frac{r\log_2 M}{1+\alpha}(1-\delta)$.

Similarly, the SE can be minimized by setting $\gamma_K = \gamma_0$ and $\gamma_n = 0$ for $n = 1, 2, \cdots, K-1$, because we can always achieve a higher SE by moving an infinitesimal value $\Delta\gamma$ from the K-th transmission to the n-th transmission, $\forall n$. The minimum average SE in this case is $\eta_K p_K = \frac{L_b}{L_b+L_0}\frac{r\log_2 M}{1+\alpha}\frac{1-\delta}{K}$.

4.10 Documentation of multi-authored chapter

UNIVERSITY OF
ARKANSAS.

College of Engineering
Department of Electrical Engineering

June 11, 2014

To Whom It May Concern

This letter is to certify that Mr. Gang Wang, a Ph.D. candidate under my supervision at the Department of Electrical Engineering, has contributed more than 51% of the work for the following paper,

Gang Wang, Jingxian Wu, and Yahong R. Zheng, "Optimum energy and spectral efficient transmissions for delay constrained hybrid ARQ systems," accepted in IEEE Transactions on Vehicular Technology, May 2014.

It should be noted that in my research group, the first author of a paper is usually the person who comes up with the original idea. For this paper, I came up with the original idea, and Mr. Gang Wang improved upon the idea and contributed more than 51% of the work to the paper.

Sincerely Yours,

Jingxian Wu
Director, Wireless Information Network Lab
Associate Professor of Electrical Engineering
University of Arkansas
Email: wuj@uark.edu
Tel: (479) 575-6584

3217 Bell Engineering Center • Fayetteville, AR 72701 • 479-575-3005 • Fax: 479-575-7967 • www.uark.edu
The University of Arkansas is an equal opportunity/affirmative action institution.

4.11 References

[1] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transaction on Wireless Communications*, vol. 4, pp. 2349 - 2360, Sept. 2005.

[2] G. Miao, N. Himayat, G. Y. Li, and A. Swami, "Cross-layer optimization for energy-efficient wireless communications: a survey," *Wiley J. Wireless Communication and Mobile Computing*, vol. 9, pp. 529 - 542, Apr. 2009.

[3] J. Wu, G. Wang, and Y. R. Zheng, "Energy efficiency and spectral efficiency tradeoff in type-I ARQ systems," in *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 356 - 366, Feb. 2014.

[4] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Transaction on Information Theory*, vol. 48, pp. 1319 - 1343, June 2002.

[5] Z. He, W. Chen, and X. Chen, "Energy minimization of portable video communication devices based on power-rate-distortion optimization," *IEEE Transactions on Circuits and System for Video Technology*, vol. 18, pp. 596 - 608, May 2008.

[6] H. Seo and B. G. Lee, "Optimum transmission power for single- and multi-hop links in wireless packet networks with ARQ capability," *IEEE Transactions on Communications*, vol. 55, pp. 996 - 1006, 2007.

[7] M. V. D. Schaar and S. S. N, "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms," *IEEE Wireless Communications*, vol. 12, pp. 50 - 58, 2005.

[8] K. T. Phan, H. Jiang, C. Tellambura, S. A. Vorobyov, and R. Fan, "Joint medium access control, routing and energy distribution in multi-hop wireless networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 5244 - 5249, 2008.

[9] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "The optimal transmission power per round for hybrid-ARQ Rayleigh fading links," *IEEE International Conference on Communications, ICC*, 2010.

[10] W. Su, S. Lee, D. A. Pados, and J. D. Matyjas, "Optimum power assignment for minimizing the average total transmission power in hybrid-ARQ Rayleigh fading links, *IEEE Transactions on Communications*, vol. 59, no. 7, pp.1867-1877, July 2011.

[11] P. Grover, K. Woyach, and A. Sahai, "Towards a communication-theoretic understanding of system level power consumption," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1744 - 1755, Sept. 2011.

[12] G. Wang, J. Wu, and Y. R. Zheng, "Cross-layer design of energy efficient coded ARQ systems," in *Proc. IEEE Global Telecommun. Conf. Globecom'12*, Dec. 2012.

[13] C. Xiong, G. Y. Li, S. Zhang, Y. Chen, and S. Xu, "Energy- and spectral-efficiency tradeoff in downlink OFDMA networks," *IEEE Transaction on Wireless Communications*, vol. 10, pp. 3874 - 3886, Nov. 2011.

[14] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communication Magazine*, vol. 49, pp. 30 - 37, June 2011.

[15] J. Wu, G. Wang, and Y. R. Zheng, "Energy and spectral efficient transmissions of coded ARQ systems," in *Proc. IEEE International Conference on Communications, ICC'13*, June 2013.

[16] D. Chase, "Code Combining–A Maximum-Likelihood Decoding Approach for Combining an Arbitrary Number of Noisy Packets," *IEEE Transactions on Communications*, vol. 33, pp. 385 - 393, May 1985.

[17] Y. Geng, Y. Wan, J. He and K. Pahlavan, "An Empirical Channel Model for the Effect of Human Body on Ray Tracing," *IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[18] G. Wang, J. Wu, and Y. R. Zheng, "Optimum energy efficient communications for hybrid ARQ systems," *IEEE Global Telecommun. Conf. Globecom'13*, 2013.

[19] G. Wang, J. Wu, and Y. R. Zheng, "An accurate frame error rate approximation of coded diversity systems with non-identical diversity branches," accepted in *IEEE International Conference on Communications, ICC'14*, June 2014.

[20] Y. Wang, G. Yue, S. Rangarajan, R. Sankar and S. Morgera, "Buffer-Aware Adaptive Scheduling for Downlink Multiuser Systems", in *Proc. IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[21] Y. Wang, S. Cui, R. Sankar, and S. Morgera, "Delay-throughput tradeoff with opportunistic relaying in wireless networks", in *Proc. IEEE Global Telecommun. Conf. Globecom'11*, Houston, Dec. 2011.

[22] K. Xu, D. Tipper, P. Krishnamurthy, and Y. Qian, "An efficient hybrid model and dynamic performance analysis for multihop wireless networks", in Proc. *IEEE International Conference on Computing, Networking and Communications, ICNC'13*, Jan. 2013.

[23] K. Xu, D. Tipper, Y. Qian, P. Krishnamurthy, and S. Tipmongkonsilp, "Time-varying performance analysis of multihop wireless networks with CBR traffic", accepted in *IEEE Transactions on Vehicular Technology*, 2014.

[24] Y. Geng, J. Chen and K. Pahlavan, "Motion Detection Using RF Signals for the First Responder in Emergency Operations," *Proc. IEEE Personal, Indoor and Mobile Radio Communications, PIMRC*, London, Sep. 2013.

[25] S. Verdu, "On channel capacity per unit cost," *IEEE Transaction on Information Theory*, vol. 36, pp. 1019 - 1030, Sept. 1990.

[26] I. Chatzigeorgiou, I. J. Wassell, and R. Carrasco, "On the frame error rate of transmission schemes on quasi-static fading channels," *42nd Annual Conference on Information Sciences and Systems, CISS 2008*, pp. 577 - 581, 2008.

[27] J. Cioffi, *Digital Communications.* Stanford Univ. Press, 2001.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge Univ. Press, 2004.

Chapter 5

Collision-Tolerant Media Access Control for Asynchronous Users over Frequency Selective Channels

Gang Wang, Jingxian Wu, Guoqing Zhou, and Geoffrey Ye Li

5.1 Abstract

In this paper, a frequency-domain cross-layer *collision-tolerant* (CT) *media access control* (MAC) scheme is proposed for the up-links of broadband wireless networks with asynchronous users. The collision tolerance is achieved with a *frequency-domain on-off accumulative transmission* (FD-OOAT) scheme, where the frequency-selective spectrum is divided into a large number of orthogonal sub-channels, and each symbol is transmitted over a small subset of the sub-channels to reduce the probability of collisions. Such a radio resource management scheme renders a special signal structure that enables *multi-user detection* (MUD) in the physical layer to resolve the collisions at the MAC layer. Most existing MUDs require precise symbol level synchronization among the users. The proposed scheme, however, can operate with asynchronous users. A new theoretical framework is provided to study the impacts of time-domain user delays on system performance, and the theoretical results provide guidelines on system designs. Both analytical and simulation results demonstrate that the proposed FD-OOAT structure with time-domain oversampling is robust to user delays and the timing phase offset caused by the sampling clock difference between the transmitter and the receiver. It is shown that the proposed scheme can achieve significant performance gains, in terms of both the number of users supported and the normalized throughput.

5.2 Introduction

With the growing demands for reliable and high data rate wireless communications, the limited radio spectrum resource is expected to support a large number of simultaneous broadband users with stringent quality of service (QoS) requirements. On the other hand, the design of reliable broadband multi-user systems faces a number of challenges, such as frequency-selective fading due to broadband communications, the competitions for the limited spectrum resource among users, and the lack of precise synchronization, etc. The main objective of this work is to develop a spectrum efficient communication technique that can address all these challenges by exploiting the interactions between *physical* (PHY) layer and *media access control* (MAC) layer in a communication network.

In many conventional MAC schemes, such as slotted ALOHA or *carrier sensing multiple access* (CSMA), signals collided at a receiver will be discarded and retransmitted. This results in a waste of the precious energy and spectrum resources. Various *collision-tolerant* (CT) MAC protocols have been developed to extract the salient information contained in the collided signals by resorting to cross-layer designs [1]– [10]. *Multi-packet reception* (MPR) in [1]– [4] assumes that the receiver can recover a fraction of the collided signals by signal processing in the PHY layer. Iterative *interference cancellation* (IC) methods are used to resolve multi-user collisions in a *contention-resolution diversity slotted ALOHA* (CRDSA) [5] and an *irregular repetition slotted ALOHA (IRSA)* scheme [6] and [7]. In the CRDSA and IRSA schemes, each packet is transmitted multiple times at random slots in a frame. If one of the packet is detected, then it can be used to subtract the interference caused by its replicas. The IC process is performed iteratively. These schemes work well under low offered loads. However, the throughput drops dramatically once the normalized offered load exceeds certain point, because the IC schemes in CRDSA or IRSA are unable to find at least one collision-free signal at the receiver to initiate the iterative IC process under heavy load. In addition, all above MAC techniques

rely on perfect synchronization among the users, which is difficult, if not impossible, to achieve in practical systems.

The limitations of iterative IC can be partly solved by using *multi-user detection* (MUD), which performs simultaneous detection of signals from two or more users collided at the receiver. MUD in the PHY layer can be combined with MAC techniques to improve the spectrum and energy efficiency in wireless networks [8]– [10]. MUD techniques are often designed with multi-dimensional signals in the PHY layer, such as *code-division multiple access* (CDMA) [8] or *orthogonal frequency division multiplexing* (OFDM) [9]. An *on-off accumulative transmission* (OOAT) scheme in [10] can support more simultaneous users than the dimension of the signals by repeating the same signal multiple times and using silence periods between two consecutive repetitions to reduce collisions. However, the OOAT scheme in [10] works only work in flat fading channels, yet broadband communications dictate an operation environment of frequency-selective fading.

In a multi-user system, two types of synchronizations are needed: the synchronization among the users, denoted as *multi-user synchronization* (MUS), and the synchronization of the sampling phase between the transmitter and receiver clocks, denoted as *sampling phase synchronization* (SPS). The SPS is usually based on correlation between a specially designed training sequence and the received signals [10], [14] and [15]. In multi-user systems, the *base station* (BS) first estimates the relative delays of all the users by correlation-based SPS. The estimated timing information can either assist the detection process [10], or be fed back to the users through a down-link control channel to achieve MUS [16]. All these schemes have residual timing offsets or synchronization errors, which could cause additional *multiple access interference* (MAI) and/or destroy the special signal structure critical to MUD [11]. The residual SPS errors may also introduce timing phase offset that will increase *inter-symbol interference* (ISI) and degrade *signal-to-noise ratio* (SNR) at the receiver [12,13]. If the users

in a multi-user system are not aligned perfectly in the time-domain, then there is always timing phase offset for some users.

In this paper, we propose a new cross-layer CT-MAC scheme that can support a large number of simultaneous users operating in frequency-selective fading, require neither MUS nor SPS, and is also robust to timing phase offsets. Most existing CT-MAC schemes in the literature are developed for flat fading channels [5]– [7], [10]. For example, time dispersion caused by frequency-selective fading will destroy the special signal structure that is critical to the original time-domain OOAT scheme [10]. We address this problem by developing a new *frequency-domain OOAT* (FD-OOAT), where a frequency-selective channel is divided into multiple orthogonal sub-channels in the frequency domain with the help of OFDM. Different from conventional OFDM, each symbol is transmitted over several sub-channels with a certain on-off pattern in our scheme. The frequency-domain repetition increases the degree-of-freedom (DoF) of the signals at the receiver, thus enables the collision tolerance of the system. With the FD-OOAT, the relative transmission delays among the users in the time-domain are manifested as phase shifts in the frequency domain, and our theoretical analysis shows that they have negligible impacts on system performance. Therefore, FD-OOAT does not require precise MUS or SPS, yet synchronization is critical to most existing CT-MAC systems. More importantly, the frequency-domain operations allow us to minimize the number of users colliding on each sub-channel by using simple on-off patterns that are radically different from those used by the original time-domain OOAT schemes [10].

Another important contribution of this work is the development of a new theoretical framework that quantifies the impacts of timing phase offset on system performance in multi-user multi-carrier systems. New analytical expressions of the frequency-domain channel coefficients are developed as functions of the timing phase offsets, and they provide guidelines on system designs. Both theoretical

and simulation results demonstrate that time-domain oversampling can effectively remove the effects of timing phase offset for multi-carrier systems. Therefore, the proposed scheme can operate in an asynchronous environment without incurring additional interference or SNR degradation. The collisions in FD-OOAT are resolved by using optimum and sub-optimum MUDs, which do not require precise synchronization as most existing MUD schemes. An analytical performance bound is derived to quantify the performance of the proposed scheme.

The rest of this paper is organized as follows. The FD-OOAT scheme with time-domain oversampling is presented in Section 5.3. The optimum and sub-optimum detection methods that can resolve collisions and collect the diversity gains are described in Section 5.4. In Section 5.5, theoretical studies are performed to quantify the impacts of multipath diversity gain and timing phase offset. Simulation results are given in Section 5.6, and Section 5.7 concludes the paper.

5.3 Frequency-Domain OOAT with Time-Domain Oversampling

The model and operations of the proposed FD-OOAT scheme with time-domain oversampling are presented in this section.

5.3.1 Proposed System Structure

Consider a wireless network with N users transmitting to the same receiver through a shared channel. Each MAC frame contains K symbols. To achieve collision tolerance in the MAC layer, users employ the FD-OOAT in the PHY layer as shown in Fig. 5.1.

The entire available bandwidth, B, is divided into KM sub-channels, denoted as sub-channels $0, 1, \cdots, KM - 1$ in order, with a bandwidth $B_0 = \frac{B}{KM}$ each. Each symbol uses M sub-channels uniformly spread over the entire frequency band, that is, the M sub-channels with indices, $\{mK + k\}_{m=0}^{M-1}$, are assigned for the k-th symbol in the frame, for $k = 0, \cdots, K-1$. During each transmission,
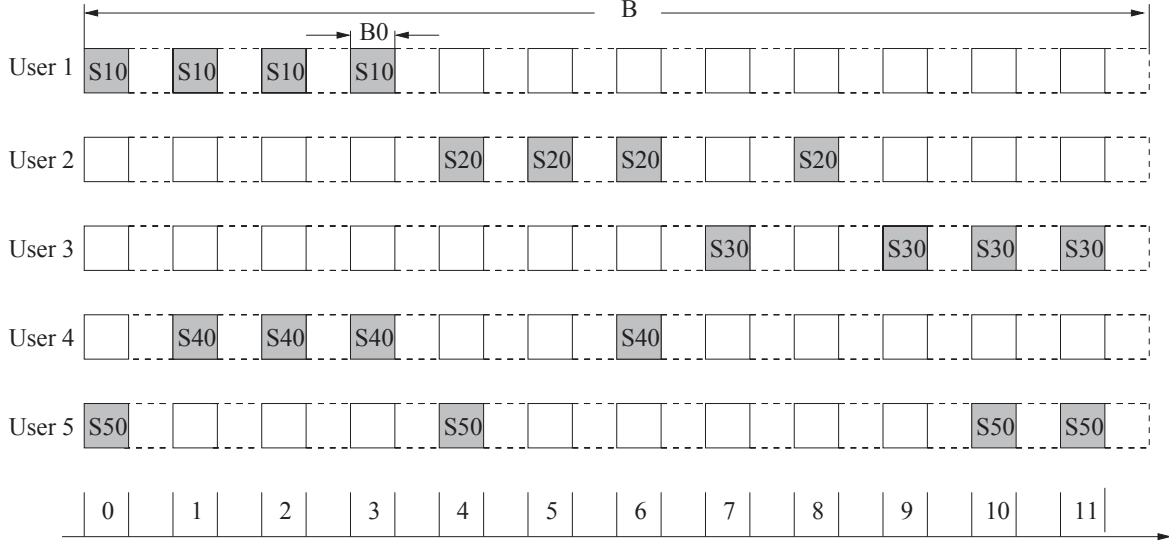
Figure 5.1: A frequency-domain OOAT system with $N = 5$ users, $R = 4$ sub-channels occupied out of $M = 12$ sub-channels for each symbol.

only R sub-channels from the M ones for each symbol are occupied. The indicator vector of the occupied sub-channels for the n-th user can be represented by a binary vector of length M, $\mathbf{p}_n = [p_n[0], \cdots, p_n[M-1]]^T \in \mathcal{B}^{M \times 1}$, where $\mathcal{B} = \{0, 1\}$, with $p_n[m] = 1$ if the k-th symbol is transmitted at the $\{mK + k\}$-th sub-channel, and $p_n[m] = 0$ otherwise. Symbols of the same user use the same transmission pattern $\mathbf{p}_n$. With such a scheme, each symbol is repeated over R sub-channels (accumulative transmission), and the utilization of the sub-channels are determined by an on-off transmission pattern $\mathbf{p}_n$. In the example shown in Fig. 5.1, there are $N = 5$ users, $M = 12$ available sub-channels per symbol, and $R = 4$ out of the 12 available sub-channels are occupied. It is assumed that all users use the same carrier frequency, thus the same set of sub-channels. As a result, signals from different users are aligned in the frequency domain as shown in Fig. 5.1.

Based on the above description, the signal transmitted on the $(mK + k)$-th sub-channel of the n-th user can be represented as

$$d_n[mK + k] = p_n[m]s_{nk}, \tag{5.1}$$

where s_{nk} is the k-th symbol from user n. Consequently, the signal vector of the n-th user can be

89

expressed as

$$\mathbf{d}_n = [d_n[0], d_n[1], \cdots, d_n[KM-1]]^T \in \mathcal{S}_+^{L \times 1}, \tag{5.2}$$

where $L = KM$, $\mathcal{S}_+ = \{\mathcal{S}, 0\}$, and $\mathcal{S}$ is the modulation constellation set with a cardinality $S = |\mathcal{S}|$.

The signal vector, $\mathbf{d}_n$, is converted to the time-domain by applying an L-point *inverse discrete Fourier transform* (IDFT) as

$$\mathbf{x}_n = \mathbf{F}_L^H \cdot \mathbf{d}_n, \tag{5.3}$$

where $\mathbf{x}_n = [x_n[0], x_n[1], \cdots, x_n[L-1]]^T$ is the time-domain signal vector, and $\mathbf{F}_L \in \mathcal{C}^{L \times L}$ is the L-point discrete Fourier transform (DFT) matrix with the $(r+1, c+1)$-th element being

$$[\mathbf{F}_L]_{r,c} = \frac{1}{\sqrt{L}} \exp\left(-j2\pi \frac{r \cdot c}{L}\right), \quad r, c = 0, 1, \cdots, L-1. \tag{5.4}$$

The space between two consecutive time-domain samples is $T_1 = \frac{1}{B}$.

Before transmission, a length-l_{cp} *cyclic prefix* (CP) is added to the time-domain signal $\mathbf{x}_n$ to avoid interference between consecutively transmitted frames. The time-domain signals pass through a transmit filter, $\varphi_1(t)$, and then transmitted over a quasi-static frequency-selective fading channel with impulse response $g_n(t)$. In a quasi-static channel, the fading is constant inside a frame, and varies independently from frame to frame. At the receiver, the received signals pass through a receive filter, $\varphi_2(t)$. Define the *composite impulse response* (CIR) of the channel as

$$h_{nc}(t) = \varphi_1(t) \odot g_n(t) \odot \varphi_2(t), \tag{5.5}$$

where $\odot$ is the convolution operator. The CIR, $h_{nc}(t)$, includes the effects of the physical channel and the transmit and receive filters. The transmit receive filters are used to limit the bandwidth of the transmitted signal. For systems with root-raised-cosine (RRC) filters as the transmit and receive filters, the bandwidth of $h_{nc}(t)$ is $\frac{1+\alpha}{T_1}$, where α is the roll-off factor of the filter, and T_1 is the space between two consecutive time domain samples at the transmitter.

The output of the receive filter is

$$y_c(t) = \sum_{n=1}^{N} \sum_{l=-\infty}^{+\infty} \sqrt{\frac{E_s}{R}} x_n[l] h_{nc}(t - lT_1 - \tau_n) + z_c(t), \qquad (5.6)$$

where E_s is the energy per symbol, τ_n is the relative delay of the n-th user, $x_n[l]$ is the l-th time-domain sample from the n-th user with a sample period T_1, $z_c(t) = p_2(t) \odot v_c(t)$ is the noise component at the output of the receive filter, with $v_c(t)$ being the *additive white Gaussian noise* (AWGN) with one-sided power spectral density N_0. The relative delay, τ_n, introduces mis-match between the receive filter and transmit filter, and the effects are captured as a time shift in the CIR $h_{nc}(t)$. It should be noted that the effects of frequency offsets are not considered in (5.6). In case of non-zero frequency offsets, they can be estimated at the receiver then compensated at the transmitter with the help of a feedback channel.

The output of the receive filter is sampled at the time instant $t = iT_2$, where $T_2 = T_1/u$ is the sampling period at the receiver, with the oversampling factor, u, being an integer. Denote the relative delays among the users as $\tau_n = l_n T_2 + \tau_{n0}$, where l_n represents the mis-alignment among the users in terms of receive samples, and $\tau_{n0} \in [0, T_2]$ is the timing phase offset between the sampling clocks at the transmitter and receiver. The discrete-time samples are

$$y_T[i] = \sum_{n=1}^{N} \sum_{l=0}^{ul_c-1} \sqrt{\frac{E_s}{R}} x_{nT}[i - l - l_n] h_{nT}[l] + z_T[i], \qquad (5.7)$$

where $y_T[i] = y_c(iT_2)$ and $z_T[i] = z_c(iT_2)$ are the T_2-spaced samples of the received signals and noise components, respectively, $h_{nT}[l] = h_{nc}(lT_2 - \tau_{n0})$ is the sampled version of the continuous-time CIR $h_{nc}(t)$, and $x_{nT}[i]$ is the oversampled version of $x_n[i]$ as $x_{nT}[i] = x_n[i/u]$, if i/u is an integer, and 0 otherwise. It is assumed that the length of the CIR, ul_c, is an integer multiple of u, with l_c being the length of the CIR without oversampling, which can be always met by appending zeros to the CIR. The timing phase offset τ_{n0} is incorporated in the discrete-time CIR $h_{nT}[l]$. We will study in Section

5.5 through frequency-domain analysis the impacts of τ_{n0} on the statistical properties of the channel coefficients and the system performance.

With the discrete-time system model given in (5.7), the length of the CP should satisfy $l_{cp} \geq l_c + l_d/u - 1$, where $l_d = \max\{l_n\}$ is the maximum relative transmission delay among the users. It should be noted that the proposed method can work for arbitrary value of l_d, and a larger l_d means a longer CP. To achieve better spectral and energy efficiency, it is assumed in the simulation that $l_d \in [0, uK)$.

Due to the time span of the transmit and receive filters, the CIR coefficients, $\{h_{nT}[l]\}_{l=0}^{uL_c-1}$, are correlated, even though the underlying channel might undergo uncorrelated scattering. The correlation coefficient, $c_n[l_1, l_2] = \mathbb{E}\left[h_{nT}[l_1]h^*_{nT}[l_2]\right]$, can be calculated as [18, eqn. (17)].

$$c_n[l_1, l_2] = \int_{-\infty}^{+\infty} R_\varphi(l_1 T_2 - \tau_{n0} - \tau)R^*_\varphi(l_2 T_2 - \tau_{n0} - \tau)\zeta(\tau)d\tau, \tag{5.8}$$

where $\zeta(\tau)$ is the power delay profile of the physical channel, and $R_\varphi(t)$ is the convolution of the transmit and receive filters.

After the removal of the CP, the received symbols can be written in a matrix form as

$$\mathbf{y}_T = \sqrt{\frac{E_s}{R}} \sum_{n=1}^{N} \mathbf{H}_{nT} \cdot \mathbf{x}_n + \mathbf{z}_T, \tag{5.9}$$

where $\mathbf{y}_T = [y_T[0], \cdots, y_T[uL-1]]^T \in \mathcal{C}^{uL\times 1}$, $\mathbf{z}_T = [z_T[0], \cdots, z_T[uL-1]]^T \in \mathcal{C}^{uL\times 1}$, $\mathbf{H}_{nT} = [\mathbf{h}_{n,1}, \mathbf{h}_{n,u+1}, \cdots, \mathbf{h}_{n,(L-1)u+1}] \in \mathcal{C}^{uL\times L}$, with $\mathbf{h}_{n,k} \in \mathcal{C}^{uL\times 1}$ being the k-th column of a circulant matrix $\mathbf{H}_n \in \mathcal{C}^{uL\times uL}$. The first column of $\mathbf{H}_n \in \mathcal{C}^{uL\times uL}$ is $\mathbf{h}_{n,1} = [\mathbf{0}_{l_n}^T, h_{nT}[0], h_{nT}[1], \cdots, h_{nT}[ul_c - 1], \mathbf{0}_{uL-l_n-ul_c}^T]^T$, and $\mathbf{0}_a$ is a length-a all-zero vector. With the equivalent discrete-time CIR representation, mis-alignments among users are represented in the form of time shifts in the columns of the circulant channel matrix $\mathbf{H}_n$, and the timing phase offsets are incorporated in the statistical properties of the discrete-time CIR $h_{nT}[l]$. They might negatively affect system performance if not

handled properly. We will study their impacts on system performance through frequency-domain analysis.

Due to the time span of the receive filter and the oversampling operation, the time-domain noise vector is also correlated. The vector, $\mathbf{z}_T$, is zero-mean complex Gaussian distributed with a covariance matrix $\mathbf{R}_{\mathbf{z}_T} = \mathbb{E}(\mathbf{z}_T \mathbf{z}_T^H) = N_0 \mathbf{R}_\varphi \in \mathcal{C}^{uL \times uL}$, where the (m,n)-th element of $\mathbf{R}_\varphi$ is $\int_{-\infty}^{+\infty} \varphi_2((m-n)T_2 + \tau)\varphi_2^*(\tau)d\tau$ [19, Lemma 2].

The uL-point DFT is applied to the vector $\mathbf{y}_T$ to convert the signal to the frequency domain as

$$\mathbf{y}_F = \sqrt{\frac{E_s}{R}} \sum_{n=1}^{N} \mathbf{H}_{nF} \cdot \mathbf{d}_n + \mathbf{z}_F, \tag{5.10}$$

where $\mathbf{y}_F = \mathbf{F}_{uL} \mathbf{y}_T$ and $\mathbf{z}_F = \mathbf{F}_{uL} \mathbf{z}_T$ are the frequency-domain signal vector and noise vector, respectively, and $\mathbf{H}_{nF} = \mathbf{F}_{uL} \mathbf{H}_{nT} \mathbf{F}_{uL}^H \in \mathcal{C}^{uL \times L}$ is the frequency-domain channel matrix, with the $(r+1, c+1)$-th element of the uL-point DFT matrix $\mathbf{F}_{uL} \in \mathcal{C}^{uL \times uL}$ being

$$[\mathbf{F}_{uL}]_{r,c} = \frac{1}{\sqrt{uL}} \exp\left(-j2\pi \frac{r \cdot c}{uL}\right), r, c = 0, 1, 2, \cdots, uL - 1. \tag{5.11}$$

Due to the correlation among the noise samples in the time domain, they are still correlated in the frequency domain. The covariance matrix of $\mathbf{z}_F$ is $\mathbf{R}_{\mathbf{z}_F} = N_0 \mathbf{F}_{uL} \mathbf{R}_\varphi \mathbf{F}_{uL}^H$. It should be noted that due to the on-off transmission, only RK out of the $L = MK$ elements in $\mathbf{d}_n$ are non-zero.

From (5.10), signals from different users are aligned in the frequency domain, even though they are asynchronous in the time domain. The time-domain mis-alignment is incorporated in the frequency-domain channel matrix $\mathbf{H}_{nF}$. The matrix $\mathbf{H}_{nF}$ can be partitioned into a stack of u sub-matrices as $\mathbf{H}_{nF} = [\mathbf{G}_{n0}^T, \cdots, \mathbf{G}_{n(u-1)}^T]^T$, where $\mathbf{G}_{nv} \in \mathcal{C}^{L \times L}$. The matrix, $\mathbf{G}_{nv}$, is a diagonal matrix, with the $(m+1)$-th diagonal element being [19, Corollary 1]

$$G_{nv}[m] = \frac{\exp\left[-j2\pi \frac{l_n \cdot (vL+m)}{uL}\right]}{\sqrt{u}} \times \sum_{l=0}^{ul_c-1} h_{nT}[l] \exp\left[-j2\pi \frac{(vL+m) \cdot l}{uL}\right]. \tag{5.12}$$

In (5.12), the relative delay, l_n, in the time-domain is manifested as a phase shift in the frequency domain as $\exp\left[-j2\pi\frac{l_n\cdot(vL+m)}{uL}\right]$.

With the model given in (5.10) and (5.12), each $d_n[m]$ is equivalently transmitted over u sub-channels with coefficients $\{G_{nv}[m]\}_{v=0}^{u-1}$. It should be noted that due to oversampling, u sub-channels at the receiver occupy the same bandwidth as one sub-channel at the transmitter. However, the relative alignment of the sub-channels from different users remain unchanged. Consider the example in Fig. 5.1, with the block diagonal structure of $\mathbf{H}_{nF}$, the sub-channels at the receiver side can be obtained by duplicating the diagram in Fig. 5.1 u times in the frequency domain, then reduce the bandwidth of each sub-channel by a factor of u. Since each symbol is repeated R times at the transmitter, each modulated symbol, s_{nk}, is equivalently transmitted over uR sub-channels in the frequency domain. Therefore, frequency diversity is achieved with the proposed FD-OOAT scheme. The uR sub-channels spread over the entire frequency band to maximize the frequency diversity. We will quantify the frequency diversity order by resorting to an analytical performance bound in Section 5.5.

5.3.2 Collision Tolerance

With the frequency-domain system representation in (5.10), the received information at the m-th sub-channel at the BS is the superposition of a set of signals, $\{d_n[m]\}_{n=1}^N$. The value of $d_n[m]$ is 0 if $p_n[i_m] = 0$. Therefore, only a subset of the users will collide at the m-th sub-channel. The collision order at the m-th sub-channel is $N_c[m] = \sum_{n=1}^N p_n[i_m]$. The collision order of the network is then defined as $N_c = \max_m N_c[m]$. We have $N_c = 2$ for the system shown in Fig. 5.1. For a system with N users, R repetitions, and M sub-channels per symbol, there are exactly NR repetitions transmitted over M sub-channels, thus it can be easily shown that the minimum collision order is $N_c = \left\lceil\frac{NR}{M}\right\rceil$, with $\lceil a \rceil$ being the smallest integer no less than a.

There are many different ways to construct the position vectors to achieve the minimum collision order. Here we present one simple construction scheme through cyclic shifting.

Definition 1: Given M and R, define the position vector of the first user as $\mathbf{p}_1 = [\mathbf{1}_R^T, \mathbf{0}_{M-R}^T]^T$, where $\mathbf{1}_R$ is an all-one vector of length R. The position vector of the n-th user can then be obtained by cyclically shifting $\mathbf{p}_1$ to the right by $(n-1)R$ positions, for $n = 2, \cdots, N$.

Lemma 2: Consider an FD-OOAT system with N users, R repetitions, and M sub-channels per symbol. If the position vectors are constructed with the cyclic shifting scheme described in Definition 1, then the collision order of the system is $N_c = \lceil \frac{NR}{M} \rceil$.

Proof: Without loss of generality, consider sub-channel with index 0. Based on the cyclic shifting construction method, user n will transmit on sub-channel 0 if and only if there exists a non-negative integer q such that

$$(n-1)R \le qM \le (n-1)R + R - 1 < nR. \tag{5.13}$$

Since q is an integer, the inequality in (5.13) can be alternatively written as

$$\lceil \frac{(n-1)R}{M} \rceil \le q < \lceil \frac{nR}{M} \rceil. \tag{5.14}$$

For a system with N users, we thus have $\max(q) \le \lceil \frac{NR}{M} \rceil - 1 < \lceil \frac{NR}{M} \rceil$. On the other hand, $\lceil \frac{NR}{M} \rceil - 1 \le \lceil \frac{(N-1)R}{M} \rceil \le \max(q)$. Therefore $\max(q) = \lceil \frac{NR}{M} \rceil - 1$. The minimum value of q is 0. Therefore there are at most $\lceil \frac{NR}{M} \rceil$ values of q satisfying the inequality. Each value of q uniquely determines an n, thus there are at most $\lceil \frac{NR}{M} \rceil$ users transmitting at sub-channel 0. The collision orders on the other sub-channels can be bounded in a similar manner by using the cyclic shifting property. ■

It should be noted that the construction described in Definition 1 is not unique. We can get a set of position vectors that achieve the minimum collision order by performing the same permutations on

all the position vectors obtained from Definition 1. Since all users permutate their position vectors following the same pattern, the relative collision relationship among the N users remains unchanged.

The oversampled FD-OOAT scheme contributes to the performance improvement of the wireless network from the following perspectives. First, the on-off transmission across the sub-channels will reduce the collision order. Second, the transmission of R identical sub-symbols with oversampling results in a uR-dimensional received signal in the frequency domain, which can be used for the detection of the N_c-dimensional signal in the space domain. Third, frequency diversity is achieved by transmitting the k-th symbol in uR sub-channels. Fourth, the relative delays among the users in the time domain are represented as phase shifts in the frequency domain, thus the user mis-alignment does not affect the collision order in the frequency domain. It should be noted that user mis-alignment has significant impacts on collision orders for the original time-domain OOAT [10].

5.4 Collision Resolution with Optimum and Sub-optimum Detections

In this section, optimum and sub-optimum detectors are developed for the oversampled FD-OOAT system to resolve the collisions among the users and to collect the inherent frequency diversity. The detectors do not require precise synchronizations among the users. The complexity of the receiver is also studied.

5.4.1 Muti-user Detection

Since the time-domain mis-alignment among the users does not affect the user alignment in the frequency domain as shown in Fig. 5.1, the k-th symbol from one user will only interfere the k-th symbols from the other users. This is different from the time-domain OOAT [10], where the k-th symbol from one user might interfere adjacent symbols from the other users due to the signal mis-alignment in the time-domain.

The k-th symbols from all the N users, $\{s_{nk}\}_{n=1}^N$, can be jointly detected by using a block of uM received signal samples $\mathbf{r}_k = [\mathbf{y}_0^T, \cdots, \mathbf{y}_{u-1}^T]^T \in \mathcal{C}^{uM \times 1}$ with $\mathbf{y}_v = [y_F[vL + k], y_F[vL + K + k], \cdots, y_F[vL + (M-1)K + k]]^T \in \mathcal{C}^{M \times 1}$. The vector $\mathbf{r}_k$ defined above is obtained by extracting uM elements from the frequency domain vector $\mathbf{y}_F$, and it can be alternatively represented as

$$\mathbf{r}_k = \mathbf{B}_k \mathbf{y}_F, \tag{5.15}$$

where $\mathbf{B}_k \in \mathcal{B}^{uM \times uL}$ is obtained by extracting uM rows from a size-uL identity matrix $\mathbf{I}_{uL}$. The indices of the extracted rows are $vL + mK + k$, for $v = 0, \cdots, u-1$ and $m = 0, \cdots, M-1$.

From (5.10), we have

$$\mathbf{r}_k = \sqrt{\frac{E_s}{R}} \mathbf{H}_k \cdot \mathbf{s}_k + \mathbf{w}_k, \tag{5.16}$$

where $\mathbf{s}_k = [s_{1k}, s_{2k}, \cdots, s_{Nk}]^T \in \mathcal{S}^{N \times 1}$ and $\mathbf{w}_k = \mathbf{B}_k \mathbf{z}_F \in \mathcal{C}^{uM \times 1}$ are the modulation symbol vector and noise vector, respectively, $\mathbf{H}_k = [\mathbf{G}_0^T, \cdots, \mathbf{G}_{u-1}^T]^T$, and $\mathbf{G}_v \in \mathcal{C}^{M \times N}$ is the frequency-domain channel matrix with the $(m+1, n)$-th element being $p_n[m]G_{nv}[mK + k]$.

Since the elements of $\mathbf{w}_k$ are extracted from $\mathbf{z}_F$, they are mutually correlated with the covariance matrix $\mathbf{R}_{\mathbf{w}_k} = N_0 \mathbf{B}_k \mathbf{F}_{uL} \mathbf{R}_p \mathbf{F}_{uL}^H \mathbf{B}_k^H$. The covariance matrix might be rank deficient. Define the pseudo-inverse of $\frac{1}{N_0} \mathbf{R}_{\mathbf{w}_k}$ as

$$\mathbf{\Phi_k} = \mathbf{V_k} \mathbf{\Omega_k^{-1}} \mathbf{V_k^H} \in \mathcal{C}^{\mathbf{uM \times uM}}, \tag{5.17}$$

with

$$\mathbf{V}_k = [\mathbf{v}_{k1}, \mathbf{v}_{k2}, \cdots, \mathbf{v}_{ku_k}] \in \mathcal{C}^{uM \times u_k} \tag{5.18a}$$

$$\mathbf{\Omega}_k = \mathrm{diag}[\omega_{k1}, \omega_{k2}, \cdots, \omega_{ku_k}] \in \mathcal{C}^{u_k \times u_k}, \tag{5.18b}$$

where u_k is the number of non-zero eigenvalues of $\frac{1}{N_0} \mathbf{R}_{\mathbf{w}_k}$, $\mathbf{\Omega}_k$ is a diagonal matrix, with the elements, $\{\omega_{ki}\}_{i=1}^{u_k}$, being the non-zero eigenvalues of $\frac{1}{N_0} \mathbf{R}_{\mathbf{w}_k}$, and $\{\mathbf{v}_{ki}\}_{i=1}^{u_k}$ are the corresponding orthonormal eigenvectors.

97

Define the noise whitening matrix $\mathbf{D}_k = \boldsymbol{\Omega}_k^{-1/2}\mathbf{V}_k^H$. Applying $\mathbf{D}_k$ on both sides of (5.16) leads to an equivalent system

$$\bar{\mathbf{r}}_k = \sqrt{\frac{E_s}{R}}\bar{\mathbf{H}}_k \cdot \mathbf{s}_k + \bar{\mathbf{w}}_k, \tag{5.19}$$

where $\bar{\mathbf{r}}_k = \mathbf{D}_k\mathbf{r}_k$, $\bar{\mathbf{H}}_k = \mathbf{D}_k\mathbf{H}_k$, and $\bar{\mathbf{w}}_k = \mathbf{D}_k\mathbf{w}_k$. The covariance matrix of the noise vector, $\bar{\mathbf{w}}_k$, in the equivalent system can be calculated as $\mathbf{R}_{\bar{\mathbf{w}}_k} = \mathbf{D}_k\mathbf{R}_{\mathbf{w}_k}\mathbf{D}_k^H = N_0\mathbf{I}_{u_p}$. Therefore, the original system in (5.16) with uM outputs and a colored noise is converted into an equivalent system with u_k outputs and a white noise.

The optimum maximum likelihood (ML) detection of (5.19) is

$$\hat{\mathbf{s}}_k = \underset{\mathbf{s}_k \in \mathcal{S}^N}{\operatorname{argmin}} \left(\bar{\mathbf{r}}_k - \sqrt{\frac{E_s}{R}}\bar{\mathbf{H}}_k\mathbf{s}_k\right)^H \left(\bar{\mathbf{r}}_k - \sqrt{\frac{E_s}{R}}\bar{\mathbf{H}}_k\mathbf{s}_k\right). \tag{5.20}$$

The ML detection requires the exhaustive search of a set of S^N possible signal vectors. The complexity of the optimum detector grows exponentially with the modulation level S and the number of users N.

A low-complexity detection algorithm is presented here to balance the performance-complexity tradeoff. The sub-optimum algorithm is developed by employing an iterative *soft input soft output* (SISO) *block decision feedback equalizer* (BDFE) [20], which performs *soft successive interference cancellation* (SSIC) among the N symbols in $\mathbf{s}_k$.

The soft input to the iterative BDFE equalizer is the *a priori* probability of the symbols, $P(s_{nk} = S_i)$, for $n = 1, \cdots, N$ and $i = 1, \cdots, S$, where $S_i \in \mathcal{S}$. The *a priori* information is obtained from the previous detection round with an iterative detection method. The soft output of the equalizer is the *a posteriori* probability of the symbols, $P(s_{nk} = S_i|\bar{\mathbf{r}}_k)$, for $n = 1, \cdots, N$ and $i = 1, \cdots, S$. With the soft output at the equalizer, define the *a posteriori* mean, $\hat{s}_{nk}$, and the extrinsic information, $\beta_{nk}[i]$,

of the symbol $s_n(k)$ as

$$\hat{s}_{nk} \quad = \quad \sum_{i=1}^{S} P(s_{nk} = S_i | \bar{\mathbf{r}}_k) S_i \tag{5.21a}$$

$$\beta_{nk}[i] \quad = \quad \log P(s_{nk} = S_i | \bar{\mathbf{r}}_k) - \log P(s_{nk} = S_i). \tag{5.21b}$$

The *a posteriori* mean, $\hat{s}_{nk}$, is used as soft decisions for the SSIC during the SISO-BDFE process. Details of the SISO-BDFE detection can be found in [20].

In the proposed sub-optimum detection, the SISO-BDFE with SSIC will be performed iteratively. At the first iteration, the *a priori* probability is initialized to $P(s_{nk} = S_i) = \frac{1}{S}$. The extrinsic information at the output of the v-th iteration will be used as the soft input of the $(v+1)$-th iteration as $P(s_{nk} = S_i) = c_{nk} \exp[\beta_{nk}[i]]$, where c_{nk} is a normalization constant to make $\sum_{i=1}^{S} P(s_{nk} = S_i) = 1$. At the final iteration, hard decision will be made based on the *a posteriori* probability generated by the SISO-BDFE as

$$\hat{s}_{nk} = \operatorname*{argmax}_{S_i \in \mathcal{S}} P(s_{nk} = S_i | \bar{\mathbf{r}}_k). \tag{5.22}$$

Simulation results show that the performance of the iterative detection algorithm usually converges after 4 iterations. The sub-optimum iterative detection algorithm can achieve a performance that is very close to its optimum counterpart, but with a much lower complexity.

5.4.2 Complexity Analysis

The complexity of the proposed oversampled FD-OOAT scheme is mainly contributed by the operations at the receiver, because the transmitter only involves simple linear operations. Thus here we only consider the complexity at the receiver side. The complexity is measured by the number of complex multiplications (CM). For BDFE detection with $u = 1$, the complexity is on the order of $\mathcal{O}(MN^2)$ [23]. When $u > 1$, the complexity of the whitening operation is on the order of $\mathcal{O}(u^3 M^3)$, and the complexity of the BDFE is on the order of $\mathcal{O}(uMN^2)$. Thus the overall complexity of the

oversampled system with BDFE is on the order of $\mathcal{O}(uMN^2 + u^3M^3)$. For the optimum detection with ML detector, the computational complexity is on the order of $\mathcal{O}(S^N + u^3M^3)$, where the first term is due to the exhaustive search and the second term is due to the noise whitening operation. The complexity of the ML detection grows exponentially with the constellation size S and the number of users N, yet the complexity of the BDFE detectors scales only with N^2. In addition, the complexity of the oversampled system is slightly higher than that of the system with $u = 1$. We will show through both theoretical analysis and simulations that $u = 2$ is sufficient to render the optimum performance for a system with at most 100% excessive bandwidth. Therefore the complexity difference between systems with $u = 1$ or $u = 2$ is very small. However, it will be shown through simulations that the oversampled systems can achieve significant performance gains over their non-oversampled counterparts.

5.5 Performance Analysis

Theoretical analysis is performed in this section to quantify the impacts of timing phase offsets on the performance of the proposed FD-OOAT scheme.

5.5.1 An Analytic Performance Lower Bound

An analytic performance lower bound on the *bit-error rate* (BER) of the proposed frequency-domain CT-MAC scheme with *binary phase shift keying* (BPSK) is developed by employing the genie-aided detector [21], where a genie provides side information of the symbols from all other users such that interference-free detection can be performed. The genie-aided bound is the same as the exact BER of a single user system, because it assumes that interferences from all other users can be removed. It will be shown through simulations that the bound is very tight even when the number of users is large due to the collision-tolerance properties of the proposed scheme.

With the interference-free assumption, the received signal corresponding to the k-th symbol of the n-th user is

$$\mathbf{r}_{nk} = \sqrt{\frac{E_s}{R}} \mathbf{g}_{nk} \cdot s_{nk} + \mathbf{w}_{nk}, \tag{5.23}$$

where $\mathbf{r}_{nk} = [\mathbf{y}_{n0}^T, \cdots, \mathbf{y}_{n(u-1)}^T]^T \in \mathcal{C}^{uR \times 1}$ with $\mathbf{y}_{nv} = [y_F[vL + n_1K + k], \cdots, y_F[vL + n_RK + k]]^T \in \mathcal{C}^{R \times 1}$, n_r is the r-th non-zero position in $\mathbf{p}_n$, $\mathbf{w}_{nk} = [\mathbf{z}_{n0}^T, \cdots, \mathbf{z}_{n(u-1)}^T]^T \in \mathcal{C}^{uR \times 1}$ with $\mathbf{z}_{nv} = [z_F[vL + n_1K + k], \cdots, z_F[vL + n_RK + k]]^T \in \mathcal{C}^{R \times 1}$, and $\mathbf{g}_{nk} = [\tilde{\mathbf{G}}_{n0}^T, \cdots, \tilde{\mathbf{G}}_{n(u-1)}^T]^T \in \mathcal{C}^{uR \times 1}$ with $\tilde{\mathbf{G}}_{nv} = [G_{nv}[n_1K + k], \cdots, G_{nv}[n_RK + k]]^T \in \mathcal{C}^{R \times 1}$ being the channel coefficient vector.

From the system model in (5.23), R repetitions of each symbol is equivalently transmitted over uR sub-channels, which is equivalent to a *single-input multiple-output* (SIMO) system. The SIMO system has correlated channel taps and is corrupted by colored noise.

The channel coefficient vector, $\mathbf{g}_{nk}$, can be represented as

$$\mathbf{g}_{nk} = \sqrt{L} \mathbf{B}_{nk} \cdot \mathbf{F}_{uL} \cdot \mathbf{h}_{n,1}, \tag{5.24}$$

where $\mathbf{h}_{n,1}$ is the first column of the circulant time-domain channel matrix $\mathbf{H}_n$, and $\mathbf{B}_{nk} \in \mathcal{B}^{uR \times uL}$ is a binary matrix, with the $(vR + r, vL + n_rK + k + 1)$-th element being 1, for $r = 1, \cdots, R$, $v = 0, \cdots, u - 1$, and all other elements of $\mathbf{B}_{nk}$ being zero.

The auto-correlation matrix, $\mathbf{R}_{nk} = \mathbb{E}[\mathbf{g}_{nk}\mathbf{g}_{nk}^H]$, can then be calculated as

$$\mathbf{R}_{nk} = L\mathbf{B}_{nk}\mathbf{F}_{uL}\mathbf{R}_{\tilde{h}n}\mathbf{F}_{uL}^H\mathbf{B}_{nk}^T, \tag{5.25}$$

where $\mathbf{R}_{\tilde{h}n} = \mathbb{E}(\mathbf{h}_{n,1}\mathbf{h}_{n,1}^H)$ is the auto-correlation matrix of the time-domain CIR vector. $\mathbf{R}_{\tilde{h}n}$ can be written as a block matrix as

$$\mathbf{R}_{\tilde{h}n} = \begin{bmatrix} \mathbf{0}_{l_n \times l_n} & \mathbf{0}_{l_n \times ul_c} & \mathbf{0}_{l_n \times l_r} \\ \mathbf{0}_{ul_c \times l_n} & \mathbf{R}_{hn} & \mathbf{0}_{ul_c \times l_r} \\ \mathbf{0}_{l_r \times l_n} & \mathbf{0}_{l_r \times ul_c} & \mathbf{0}_{l_r \times l_r} \end{bmatrix}, \tag{5.26}$$

where $l_r = L - l_n - ul_c$, and the (l_1, l_2)-th element of $\mathbf{R}_{hn} \in \mathcal{C}^{ul_c \times ul_c}$ is $c_n[l_1, l_2]$ defined in (5.8).

The covariance matrix $\mathbf{R}_{\mathbf{w}_{nk}}$ of the colored noise $\mathbf{w}_{nk}$ can be represented as

$$\mathbf{R}_{\mathbf{w}_{nk}} = N_0 \mathbf{B}_{nk} \mathbf{F}_{uL} \mathbf{R}_{\varphi} \mathbf{F}_{uL}^H \mathbf{B}_{nk}^T. \tag{5.27}$$

The covariance matrix might be rank deficient. Define the pseudo-inverse of the noise covariance matrix $\mathbf{R}_{\mathbf{w}_{nk}}$

$$\mathbf{R}_{\mathbf{w}_{nk}}^+ = \mathbf{U}_{nk} \mathbf{\Lambda}_{nk}^{-1} \mathbf{U}_{nk}^H \in \mathcal{C}^{uR \times uR}, \tag{5.28}$$

with

$$\mathbf{U}_{nk} = [\mathbf{u}_{nk,1}, \mathbf{u}_{nk,2}, \cdots, \mathbf{u}_{nk,v_k}] \in \mathcal{C}^{uR \times v_k} \tag{5.29a}$$

$$\mathbf{\Lambda}_{nk} = \mathrm{diag}[\lambda_{nk,1}, \lambda_{nk,2}, \cdots, \lambda_{nk,v_k}] \in \mathcal{C}^{v_k \times v_k}, \tag{5.29b}$$

where v_k is the number of non-zero eigenvalues of $\mathbf{R}_{\mathbf{w}_{nk}}$, $\mathbf{\Lambda}_{nk}$ is a diagonal matrix with $\{\lambda_{nk,i}\}_{i=1}^{v_k}$ being the non-zero eigenvalues of $\mathbf{R}_{\mathbf{w}_{nk}}$, and $\{\mathbf{u}_{nk,i}\}_{i=1}^{v_k}$ are the corresponding orthonormal eigenvectors.

Define the noise whitening matrix $\mathbf{D}_{nk} = \mathbf{\Lambda}_{nk}^{-1/2} \mathbf{V}_{nk}^H$. Applying $\mathbf{D}_{nk}$ to both sides of (5.28) yields an equivalent system

$$\bar{\mathbf{r}}_{nk} = \sqrt{\frac{E_s}{R}} \bar{\mathbf{g}}_{nk} \cdot s_{nk} + \bar{\mathbf{w}}_{nk}, \tag{5.30}$$

where $\bar{\mathbf{r}}_{nk} = \mathbf{D}_{nk} \mathbf{r}_{nk}$, $\bar{\mathbf{g}}_{nk} = \mathbf{D}_{nk} \mathbf{g}_{nk}$, and $\bar{\mathbf{w}}_{nk} = \mathbf{D}_{nk} \mathbf{w}_{nk}$ with the covariance matrix of $\bar{\mathbf{w}}_{nk}$ being $\mathbf{R}_{\bar{\mathbf{w}}_{nk}} = \mathbf{D}_{nk} \mathbf{R}_{\mathbf{w}_{nk}} \mathbf{D}_{nk}^H = N_0 \mathbf{I}_{v_k}$.

The SNR of (5.30) can be written as

$$\gamma = \mathbf{g}_{nk}^H \mathbf{R}_{\mathbf{w}_{nk}}^+ \mathbf{g}_{nk} \frac{\gamma_0}{R}, \tag{5.31}$$

where $\gamma_0 = \frac{E_s}{N_0}$ is the SNR without fading. For systems with BPSK and Rayleigh fading, the error probability for s_{nk} is [12]

$$P_{nk}(E) = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} \prod_{r=1}^{\tilde{L}_{nk}} \left[1 + \frac{\delta_{nkr} \gamma_0}{R \sin^2 \theta}\right]^{-1} d\theta, \tag{5.32}$$

where $\tilde{L}_{nk}$ is the rank of the product matrix, $\mathbf{D}_{nk}\mathbf{R}_{nk}\mathbf{D}_{nk}^H$, and δ_{nkr}, for $r = 1, \cdots, \tilde{L}_{nk}$, are the corresponding non-zero eigenvalues. A closed-form expression of (5.32) can be obtained through partial fraction expansion. In case all the eigenvalues are unique, which is the case for most practical systems, the closed-form expression of (5.32) can be written as

$$P_{nk}(E) = \frac{1}{2} \sum_{r=1}^{\tilde{L}_{nk}} \prod_{l=1,l\neq r}^{\tilde{L}_{nk}} \frac{\delta_{nkr}}{\delta_{nkr} - \delta_{nkl}} \left[1 - \sqrt{\frac{\delta_{nkr}\gamma_0}{R + \delta_{nkr}\gamma_0}} \right] \tag{5.33}$$

The average BER can then be calculated as

$$P(E) = \frac{1}{NK} \sum_{n=1}^{N} \sum_{k=1}^{K} P_{nk}(E). \tag{5.34}$$

In the above analysis, the order of multipath diversity is $\tilde{L}_{nk}$, which is the rank of $\mathbf{D}_{nk}\mathbf{R}_{nk}\mathbf{D}_{nk}^H$. The matrix $\mathbf{R}_{nk}$ corresponds to correlation of the frequency domain channel coefficients, and $\mathbf{D}_{nk}^H\mathbf{D}_{nk} = \mathbf{R}_{\mathbf{w}_{nk}}^+$ is the pseudo-inverse of the noise covariance matrix, $\mathbf{R}_{\mathbf{w}_{nk}}$.

The off-diagonal elements of the matrix $\mathbf{R}_{\mathbf{w}_{nk}}$ are contributed by the correlation of the colored noise. The uR elements in the noise vector, $\mathbf{w}_{nk}$, are extracted from the size-uL frequency-domain noise vector $\mathbf{z}_F$ based on the transmission pattern $\mathbf{p}_n$, and there is at least K sub-channels between any two samples in $\mathbf{w}_{nk}$. As a result, the mutual correlation between the samples in $\mathbf{w}_{nk}$ is usually very small. To measure the mutual correlation of the samples in $\mathbf{w}_{nk}$, define a metric

$$\rho = \frac{1}{NK} \sum_{n=1}^{N} \sum_{k=1}^{K} \frac{\|\mathbf{R}'_{\mathbf{w}_{nk}}\|_2}{\|\mathbf{R}_{\mathbf{w}_{nk}}\|_2}, \tag{5.35}$$

where $\mathbf{R}'_{\mathbf{w}_{nk}}$ is a diagonal matrix obtained by setting all off-diagonal elements of $\mathbf{R}_{\mathbf{w}_{nk}}$ to 0, and $\|\mathbf{A}\|_2$ is the Frobenius norm of the matrix $\mathbf{A}$. The metric $0 \leq \rho \leq 1$ measures the percentage of energy on the diagonal of $\mathbf{R}_{\mathbf{w}_{nk}}$, and $\rho = 1$ means that $\mathbf{R}_{\mathbf{w}_{nk}}$ is a diagonal matrix. Table 1 shows the values of $1 - \rho$ with $u = 2$, $M = 12$, $R = 2$, and various values of K. It is clear that ρ is very close to 1, and the difference between ρ and 1 decreases as K increases. The results in Table 1 demonstrate that

103

Table 5.1: The metric $1 - \rho$ under various values of K ($M = 12$, $R = 2$ and $u = 2$).

K	1	10	20	50	100
$1-\rho$	4.9×10^{-3}	2.3×10^{-4}	8.9×10^{-5}	4.0×10^{-5}	1.8×10^{-5}

the off-diagonal elements of $\mathbf{R}_{\mathbf{w}_{nk}}$ are negligible compared to its diagonal elements. Therefore, the mutual correlation among the noise samples is very small or negligible.

If we ignore the off-diagonal elements of $\mathbf{R}_{\mathbf{w}_{nk}}$ and approximate the noise vector $\mathbf{w}_{nk}$ as white noise with correlation matrix $\mathbf{R}'_{\mathbf{w}_{nk}}$, then we can simplify the error performance analysis. With the white noise assumption, the SNR in (5.31) can be approximated by

$$\gamma'_{nk} = \frac{\gamma_0}{R} \sum_{v=0}^{u-1} \sum_{r=1}^{R} |G_{nv}[n_r K + k]|^2 \phi_{nk}[vR + r], \tag{5.36}$$

where $\phi_{nk}[r] = q_{nk}^{-1}[r]$ if $q_{nk}[r] \neq 0$ with $q_{nk}[r]$ being the r-th diagonal element of $\mathbf{R}_{\mathbf{w}_{nk}}$, and $\phi_{nk}[r] = 0$ otherwise. The error probability in (5.32) and (5.34) can then be approximated by using the eigenvalues of the product matrix $\mathbf{D}'_{nk} \mathbf{R}_{nk} \mathbf{D}'^H_{nk}$, with $\mathbf{D}'_{nk} = \text{diag}\left\{ \phi_{nk}^{1/2}[0], \cdots, \phi_{nk}^{1/2}[uR - 1] \right\}$ being a diagonal matrix. The BER results calculated with the white approximation in (5.36) is very close to the exact genie-aided bound in (5.32), from our simulation since ρ is very close to 1.

5.5.2 Impacts of Relative Delays

In this subsection, a theoretical framework is provided to study the impacts of the relative delays among the users on the performance of the proposed FD-OOAT scheme. From the analysis in the previous subsection, the performance of the system is dominated by the statistical properties of the SNR γ'_{nk} defined in (5.36), which in turn depends on the squared amplitude of the channel coefficients, $|G_{nv}[m]|^2$. It should be noted that the power and the auto-correlation of the noise components are independent of the relative delays τ_n as evident in (5.27).

The relative delay can be expressed as $\tau_n = l_n T_2 + \tau_{n0}$, where l_n represents the mis-alignment

among the asynchronous users, and $\tau_{n0} \in [0, T_2]$ is the timing phase offset of the sampler. It is clear from (5.12) that l_n has no impact on the squared amplitude $|G_{nv}[m]|^2$. Next we will study the impact of τ_{n0} on $|G_{nv}[m]|^2$.

Define the *discrete-time Fourier transform* (DTFT) of the T_2-spaced discrete-time CIR, $h_{nT}[l]$, as

$$H_{nT}(f) = \sum_{l=0}^{ul_c-1} h_{nT}[l] e^{-j2\pi lf}, 0 \le f \le 1 \tag{5.37}$$

Since $h_{nT}[l] = h_{nc}(lT_2 - \tau_{n0})$, based on the sampling theorem, the DTFT can be expressed as

$$H_{nT}(f) = \frac{1}{T_2} \sum_{i=-\infty}^{\infty} \mathcal{H}_{nc}\left(\frac{f-i}{T_2}\right) \exp\left(-j2\pi\tau_{n0}\frac{f-i}{T_2}\right), \tag{5.38}$$

where $\mathcal{H}_{nc}\left(\frac{f}{T_2}\right)$ is the Fourier transform of the CIR $h_{nc}(t)$.

From (5.12), (5.37), and (5.38), we can write the frequency-domain channel coefficient, $G_{nv}[m]$, as

$$G_{nv}[m] = \frac{\exp\left(-j2\pi\frac{l_n \cdot (vL+m)}{uL}\right)}{T_2\sqrt{u}} \times \sum_{i=-\infty}^{\infty} \mathcal{H}_{nc}\left(\frac{vL+m}{uLT_2} - \frac{i}{T_2}\right) \exp\left(-j2\pi\tau_{n0}\frac{vL+m-uLi}{uLT_2}\right). \tag{5.39}$$

The CIR, $h_{nc}(t)$, includes the effects of the physical channel and the transmit and receive filters. From (5.5), we have

$$\mathcal{H}_{nc}\left(\frac{f}{T_2}\right) = \mathcal{P}_1\left(\frac{f}{T_2}\right) \mathcal{G}_n\left(\frac{f}{T_2}\right) \mathcal{P}_2\left(\frac{f}{T_2}\right), \tag{5.40}$$

where $\mathcal{P}_i\left(\frac{f}{T_2}\right)$ and $\mathcal{G}_n\left(\frac{f}{T_2}\right)$ are the Fourier transforms of $\varphi_i(t)$ and $g_n(t)$, respectively. If the roll-off factor of the transmit and receive filters is α, then the frequency domain support of $\mathcal{P}_i\left(\frac{f}{T_2}\right)$ is $\left|\frac{f}{T_2}\right| \le \frac{1+\alpha}{2T_1}$, or $|f| \le \frac{1+\alpha}{2u}$. All practical systems have at most 100% excessive bandwidth, *i.e.*, $\alpha \le 1$. Therefore, $\mathcal{H}_{nc}\left(\frac{f}{T_2}\right) = 0$ for $|f| > \frac{1}{u}$.

5.5.2.1 $u = 1$

For a system without oversampling, we have $T_1 = T_2$, and the frequency domain support of $\mathcal{P}_i\left(\frac{f}{T_1}\right)$ and $\mathcal{H}_{nc}\left(\frac{f}{T_1}\right)$ are $\left|\frac{f}{T_1}\right| < \frac{1+\alpha}{2T_1}$. In this case, due to the excessive bandwidth of the transmitted signal

105

when $\alpha > 0$, the sampling operation at the receiver causes spectrum aliasing as shown in (5.38) and (5.39). It is apparent from (5.39) that the frequency domain channel coefficient is a function of τ_{n0}. Therefore, the performance of the system with $u = 1$ will be affected by the timing phase offset τ_{n0}.

5.5.2.2 $u \geq 2$

The frequency domain support of $\mathcal{P}_i \left(\frac{f}{T_2} \right)$ and $\mathcal{H}_{nc} \left(\frac{f}{T_2} \right)$ are $|\frac{f}{T_2}| < \frac{1+\alpha}{2uT_2} \leq \frac{1}{2T_2}$ for $\alpha \leq 1$. Therefore, the sampling rate $\frac{1}{T_2}$ is at least twice as much as the signal bandwidth, thus there is no spectrum aliasing after the sampling operation. The frequency-domain channel coefficient in (5.39) can be simplified to

$$G_{nv}[m] = \frac{\exp\left(-j2\pi\frac{l_n \cdot (vL+m)}{uL}\right)}{T_2\sqrt{u}} \times \mathcal{H}_{nc}\left(\frac{vL+m}{uLT_2}\right) \exp\left(-j2\pi\tau_{n0}\frac{vL+m}{uLT_2}\right). \tag{5.41}$$

The squared amplitude of the channel coefficient can then be expressed as

$$|G_{nv}[m]|^2 = \frac{1}{T_2\sqrt{u}} \left| \mathcal{H}_{nc}\left(\frac{vL+m}{uLT_2}\right) \right|^2. \tag{5.42}$$

It is interesting to note that $|G_{nv}[m]|^2$ is independent of the user mis-alignments l_n or the timing phase offset τ_{n0}. Since the system performance is dominated by the squared amplitude of the channel coefficient as shown in the SNR defined in (5.36), the user mis-alignments or timing phase offset has a very small, if any, impact on the performance of the system when $u \geq 2$. Specifically, for systems with at most 100% excessive bandwidth, an oversampling factor of 2 is sufficient to avoid spectrum aliasing at the receiver, thus removes the impacts of τ_{n0}. The above analysis is corroborated by simulation results with both optimum and sub-optimum detectors.

5.6 Simulation Results

Simulation results are presented in this section to demonstrate the performance of the oversampled FD-OOAT scheme with the optimum or low-complexity detection. The effects of time-domain

oversampling and the timing phase offset on the system performance are also studied in this section.

In the simulation examples, the sample period at the transmitter is set to $T_1 = 3.69 \ \mu s$, and RRC filters with a roll-off factor $\alpha = 1.0$ is used for both the transmit and receive filters. The relative delays among the users, τ_n, is uniformly distributed between $[0, KT_1]$ with $K = 50$ unless stated otherwise. The frequency-selective fading channel follows the *Typical Urban* (TU) *power delay profile* (PDP) [22].

Fig. 5.2 shows the BER results of the proposed CT-MAC system under various system configurations. There are $M = 12$ sub-channels per symbol and each symbol is transmitted with $R = 2$ repetitions. The sub-optimum BDFE detection is performed with 4 iterations. The analytical results are obtained with both (5.33) and the white approximation as in (5.36), and the two results overlap. Only the one obtained with (5.33) is shown in the figure. We have the following observations about the results. First, when $N = 1$, the analytical and simulation results match perfectly for both $u = 1$ and 2. Second, with the BDFE receiver, increasing N has less impacts on the oversampled system with $u = 2$ than the system with $u = 1$. At BER $= 2 \times 10^{-3}$, increasing N from 1 to 10 results in a 1.5 dB and a 0.8 dB performance loss for systems with $u = 1$ and $u = 2$, respectively. This indicates that the proposed FD-OOAT system can operate properly even when there are a large number of users and collisions. In addition, when $u = 2$ and $N = 10$, the sub-optimum BDFE receiver achieves almost the same performance as the optimum ML receiver, but with a much lower complexity. Third, the oversampled system consistently outperforms the system without oversampling. The performance improvement is contributed by the additional multipath diversity and the insensitivity to the timing phase offset due to the oversampling operation. At BER $= 2 \times 10^{-3}$ and $N = 10$, the oversampled system outperforms its non-oversampled counterpart by 5.6 dB when BDFE is used.

The effects of the receiver timing phase offset on the system performance are studied through
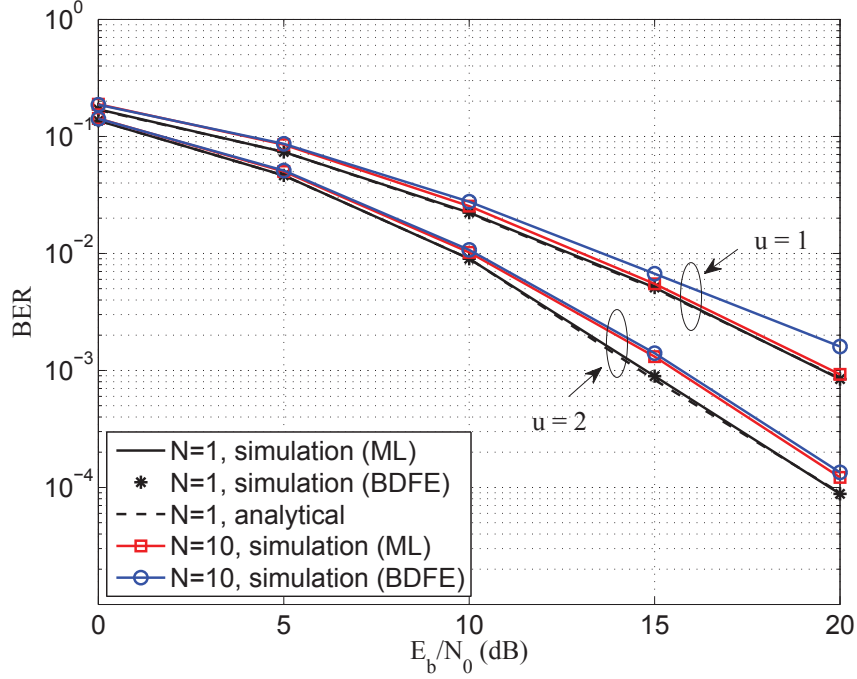
Figure 5.2: BER performance comparison of systems with $M = 12$ sub-channels per symbol, $R = 2$ repetitions, and different number of users.

simulations in Fig. 5.3 for single-user systems and Fig. 5.4 for multi-user systems, respectively. In Fig. 5.3, there are $M = 12$ sub-channels per symbol, and each symbol is transmitted with $R = 2$ repetitions. To have a better understanding on the effects of timing phase offset, it is assumed that τ_{n0} is fixed at 0 or $0.5T_2$ in Fig. 5.3. The performance of the system with $u = 1$ varies as τ_{n0} changes, yet the performance of the oversampled system is independent of τ_{n0}.

A similar observation is obtained in Fig. 5.4 for systems with multiple users, where the BER is shown as a function of τ_{n0}. The mis-alignment among the asynchronous users, l_n, is uniformly distributed between $[0, uK]$. The E_b/N_0 is 10 dB. The BER of the oversampled system stays constant regardless of the values of τ_{n0}, for both the optimum and sub-optimum algorithms with different number of users. On the other hand, the BER of the system with $u = 1$ is a function of τ_{n0}. The simulation results corroborate the theoretical analysis that twice oversampling is sufficient to
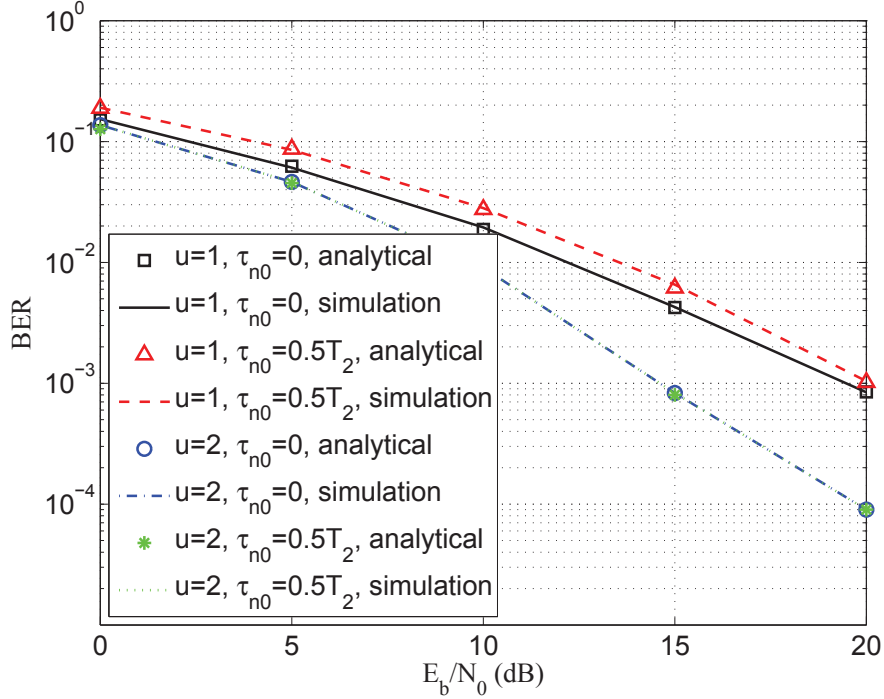
Figure 5.3: The effects of the receiver timing phase offset on the BER performance of the system (There are $N = 1$ user, $M = 12$ sub-channels per symbol, and $R = 2$ repetitions).

remove the effects of τ_{n0} for a system with at most 100% excessive bandwidth. Therefore, the proposed oversampled FD-OOAT scheme can operate effectively at the presence of both multi-user interference, user mis-alignment, and timing phase offset.

Fig. 5.5 demonstrates the impacts of the number of iterations on the *frame error rate* (FER) with the sub-optimum BDFE detector through simulation results. There are $N = 10$ active users, $M = 12$ sub-channels per symbol, $R = 2$ repetitions. As seen from the figure, the largest performance gain is achieved at the second iteration and the performance converges at the fourth iteration for systems with $u = 1$ or $u = 2$. At the fourth iteration and FER $= 4 \times 10^{-2}$, the FER performance of the oversampled system outperforms the one without oversampling by 5.6 dB, which is consistent with the BER improvement observed in Fig. 5.2.

Fig. 5.6 shows the normalized throughput as a function of the normalized offered load for various
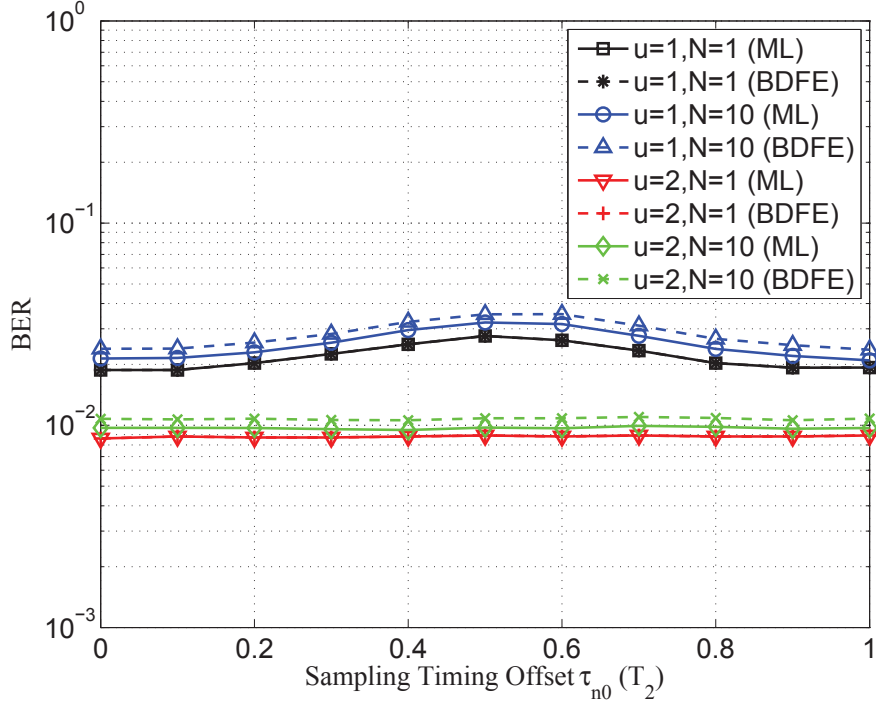
Figure 5.4: BER v.s. timing phase offset ($E_b/N_0 = 10$ dB. There are $M = 12$ sub-channels per symbol, and $R = 2$ repetitions).

MAC schemes. For the FD-OOAT system, there are $M = 10$ sub-channels per symbol, and $R = 2$ repetitions. All other systems have $M = 10$ slots per frame. The normalized offered load of all the systems is calculated as $G = \frac{N}{M}$. The normalized throughput is defined as the amount of data successfully delivered to the receiver per unit time per unit bandwidth. The normalized throughput for the FD-OOAT scheme is calculated as $\frac{N}{M}(1 - \text{FER})$. Details of the calculation of the normalized offered load and normalized throughput can be found in [10]. For the slotted ALOHA, CRDSA, and IRSA systems, the simulations are performed under the assumption of noise-free communication, $i.e.$, the only source of errors for these systems is the unresolvable signal collisions among the users. Results obtained under the noise-free assumption represent the best possible performance under any channel configurations. On the other hand, the results of the proposed FD-OOAT systems are obtained in a frequency-selective fading channel with $E_b/N_0 = 15$ dB. As shown in the figure, the slotted ALOHA,
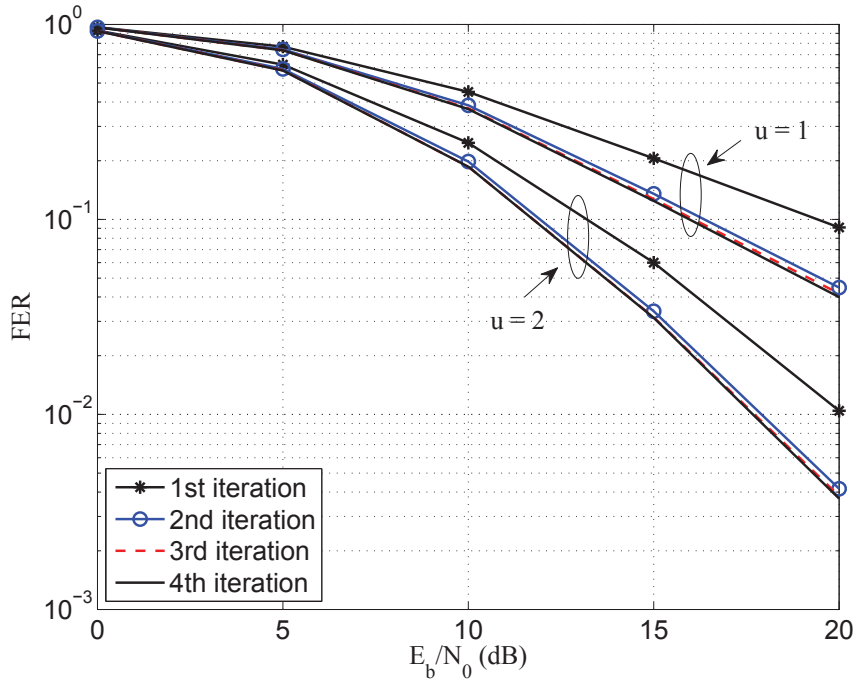
Figure 5.5: FER performance of systems with the BDFE receiver (There are $N = 10$ users, $M = 12$ sub-channels, and $R = 2$ repetitions).

CRDSA and IRSA achieve their respective peak throughputs when $G \leq 1$, and the throughputs drop dramatically when $G > 1$. This is similar to a traditional orthogonal frequency division multiple access (OFDMA) system, where the maximum number of users supported is the same as the number of sub-channels. The proposed FD-OOAT scheme achieves the maximum throughput 1.03 bps/Hz at $G = 1.6$ when $u = 1$. For the oversampled system with $u = 2$, the maximum throughput 2.06 bps/Hz is achieved at $G = 2.6$. Therefore, the FD-OOAT system can be overloaded by supporting more users than the number of sub-channels, yet all the other MAC or OFDMA schemes must operate with $G \leq 1$. Employing FD-OOAT increases both the number of users supported and peak throughput. In addition, time-domain oversampling allows the FD-OOAT system to support 60% more users than the system with $u = 1$, and improves the throughput by 100%.
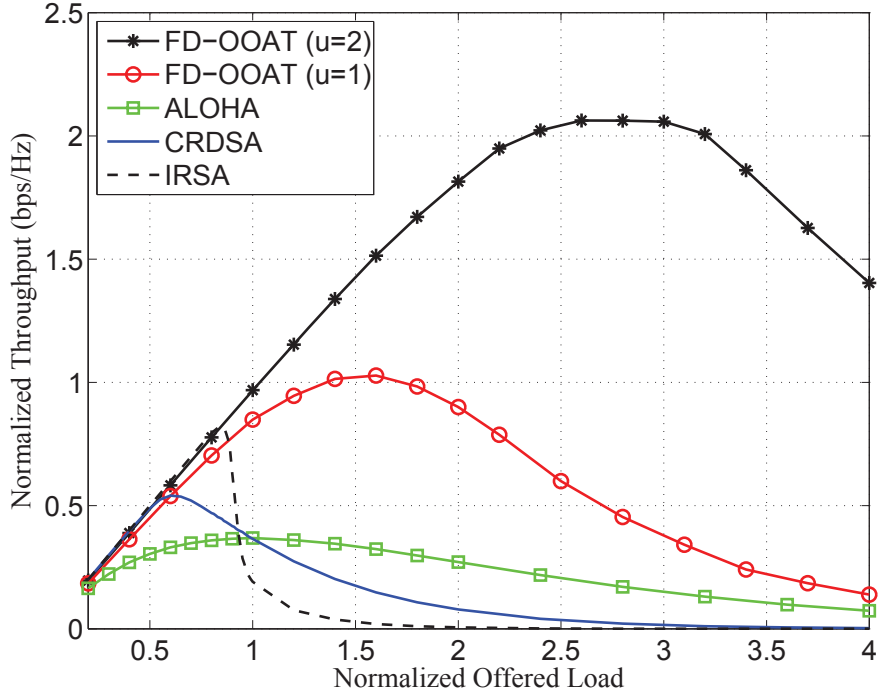
Figure 5.6: Normalized throughput v.s. normalized offered load.

5.7   Conclusions

A cross-layer CT-MAC scheme with frequency-domain OOAT and time-domain oversampling has been proposed for broadband wireless networks operating in frequency-selective fading channels. With the help of time-domain oversampling, the proposed scheme can operate without precise synchronization, and it is insensitive to timing phase offsets between the sampling clocks at the transmitter and receiver. The collision tolerance in the MAC layer was achieved by performing MUD over the specially designed FD-OOAT signal in the PHY layer. Simulation results demonstrated that 1) the performance of the oversampled FD-OOAT system was insensitive to user mis-alignment or sampler timing phase offset; 2) significant multipath diversity gain was achieved with the oversampled FD-OOAT scheme; 3) the proposed scheme could achieve a high spectral efficiency by supporting a large number of simultaneous broadband users. An oversampled FD-OOAT with M sub-channels per symbol could support up to $N = 2.6M$ simultaneous users and has a normalized throughput

peak at 2.06 bps/Hz with BPSK modulation.

5.8 Documentation of multi-authored chapter

UNIVERSITY OF ARKANSAS.

College of Engineering
Department of Electrical Engineering

June 11, 2014

To Whom It May Concern

This letter is to certify that Mr. Gang Wang, a Ph.D. candidate under my supervision at the Department of Electrical Engineering, has contributed more than 51% of the work for the following paper,

Gang Wang, Jingxian Wu, Guoqing Zhou, and Geoffrey Ye Li, "Collision-tolerant media access control for asynchronous users over frequency selective channels," IEEE Transactions on Wireless Communications, vol. 12, no. 10, pp. 5162-5171, Sept. 2013.

It should be noted that in my research group, the first author of a paper is usually the person who comes up with the original idea. For this paper, I came up with the original idea, and Mr. Gang Wang improved upon the idea and contributed more than 51% of the work to the paper.

Sincerely Yours,

Jingxian Wu
Director, Wireless Information Network Lab
Associate Professor of Electrical Engineering
University of Arkansas
Email: wuj@uark.edu
Tel: (479) 575-6584

3217 Bell Engineering Center • Fayetteville, AR 72701 • 479-575-3005 • Fax: 479-575-7967 • www.uark.edu
The University of Arkansas is an equal opportunity/affirmative action institution.
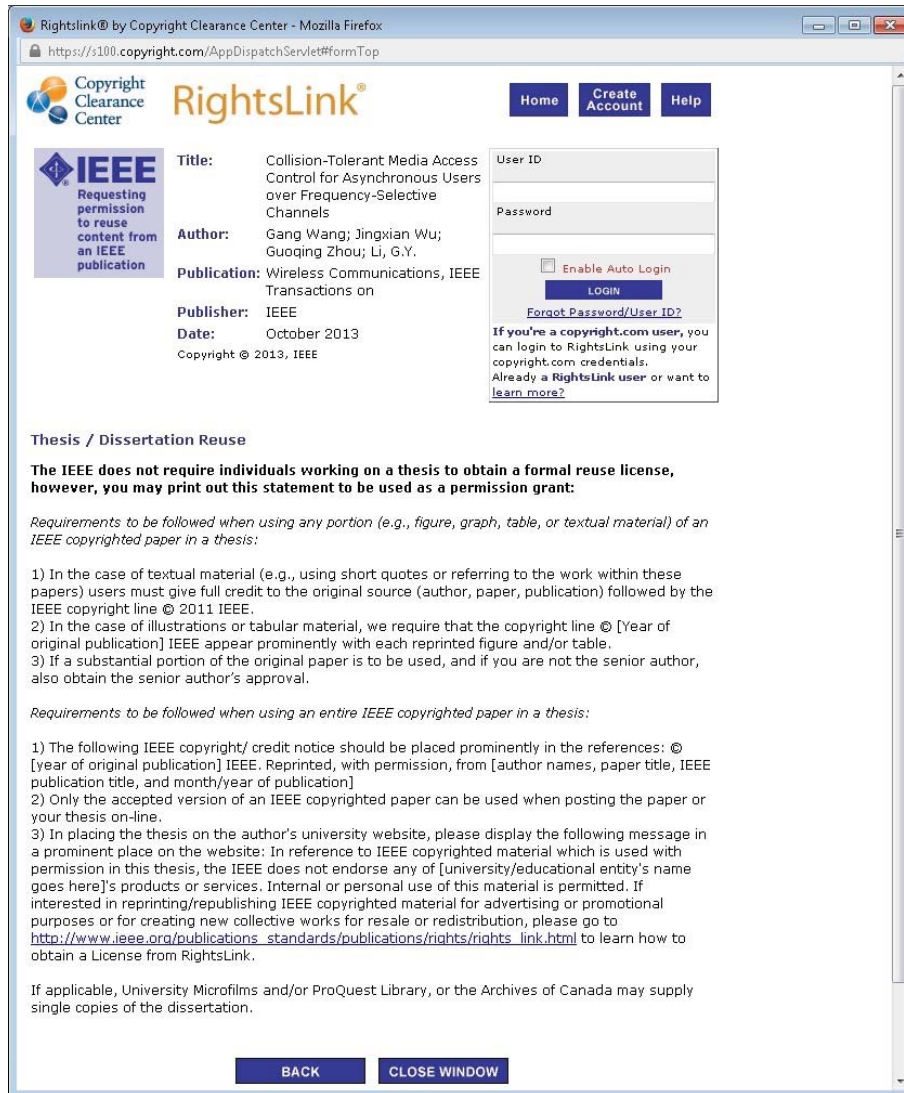
5.9   References

[1] G. Mergen and L. Tong, "Receiver controlled medium access in multihop ad hoc networks with multipacket reception," in *Proc. IEEE Military Commun. Conf. MILCOM 2001*, vol. 2, pp. 1014 - 1018, 2001.

[2] L. Tong, V. Naware, and P. Venkitasubramaniam, "Signal processing in random access," *IEEE Sig. processing Mag.*, vol. 21, pp. 29-39, Sept. 2004.

[3] Q. Zhao and L. Tong, "A dynamic queue protocol for multiaccess wireless networks with multipacket reception," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 2221 - 2231, Nov. 2004.

[4] R. Samano-Robles, M. Ghogho, and D. C. Mclernon, "An infinite user model for random access protocols assisted by multipacket reception and retransmission diversity," in *Proc. IEEE Sig. Processing Advances Wireless Commun. SPAWC 2008*, pp. 111 - 115, 2008.

[5] E. Casini, R. De Gaudenzi, and O. del Rio Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): an enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1408 - 1419, Apr. 2007.

[6] G. Liva, "A slotted ALOHA scheme based on bipartite graph optimization," in *Proc. Intern. ITG Conf. Source and Channel Coding SCC'10*, Jan. 2010.

[7] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, pp. 477 - 487, Feb. 2011.

[8] B. Lu, X. Wang, and J. Zhang, "Throughput of CDMA data networks with multiuser detection, ARQ, and packet combining," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1576 - 1589, Sept. 2004.

[9] Y. J. Zhang and K. B. Letaief, "An efficient resource-allocation scheme for spatial multiuser access in MIMO/OFDM systems," *IEEE Trans. Commun.*, vol. 53, pp. 107-116, Jan. 2005.

[10] J. Wu and G. Y. Li, "Collision-tolerant media access control with on-off accumulative transmission," *IEEE Trans. Wireless Commun.*, vol. 12, pp. 50-59, Jan. 2013.

[11] D. Marabissi, R. Fantacci and S. Papini, "Robust multiuser interference cancellation for OFDM systems with frequency offset," *IEEE Trans. Wireless Commun.*, vol. 5, pp. 3068 - 3076, Nov. 2006.

[12] J. Wu and C. Xiao, "Performance analysis of wireless systems with doubly selective Rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 56, pp. 721 - 730, Mar. 2007.

[13] J. Wu, Y. R. Zheng, K. B. Letaief and C. Xiao, "On the error performance of wireless systems with frequency-selective fading and receiver timing phase offset," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 720 - 729, Feb. 2007.

[14] S. Tian, K. Panta, H. A. Suraweera, B. J. C. Schmidt, S. McLaughlin and J. Armstrong, "A novel timing synchronization method for ACO-OFDM based optical wireless communications," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 4958 - 4967, Dec. 2008.

[15] B. Park, H. Cheon, C. Kang, and D. Hong, "A novel timing estimation method for OFDM systems," *IEEE Commun. Lett.*, vol. 7, pp. 239-241, May 2003.

[16] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," *IEEE Trans. Commun.*, vol. 52, pp. 296-306, Feb. 2004.

[17] J. Lee and S. Kim, "Time and frequency synchronization for OFDMA uplink system using the SAGE algorithm," *IEEE Trans. on Wireless Commun.*, vol. 6, pp. 1176 - 1181, Apr. 2007.

[18] C. Xiao, J. Wu, S.-Y. Leong, Y. R. Zheng and K. B. Letaief, "A discrete-time model for triply selective MIMO Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1678 - 1688, Sept. 2004.

[19] J. Wu, and Y. R. Zheng, "Oversampled orthogonal frequency division multiplexing in doubly selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 59, pp. 815 - 822, Mar. 2011.

[20] J. Wu and Y. R. Zheng, "Low complexity soft input soft output block decision feedback equalization," *IEEE J. Selected Areas Commun.*, vol. 26, pp. 281 - 289, 2008.

[21] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communications*, 3rd Ed., Springer, 2003.

[22] ITU-R Recommendation M.1225, "Guidelines for evaluation of radio transmission techniques for IMT-2000," 1997.

[23] J. Tao, J. Wu, and Y. R. Zheng, "Reliability-based turbo detection," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 2352 - 2361, 2011.

5.10 Copyright Clearance

Chapter 6

Conclusions

This chapter summarizes the main contributions of this dissertation and lists some possible directions for the future research.

6.1 Contributions

The contents presented in this dissertation focus on the energy efficiency and spectral efficiency for wireless systems, and the main contributions are summarized as follows.

First, we developed an accurate FER approximation formula for quasi-static Rayleigh fading channel, a general assumed channel model for wireless networks. The FER approximation was obtained with a threshold-based method, and the threshold value modelled as a linear function of the frame length in the log domain. The analytical FER approximation was expressed as an explicit function of a large number of system parameters related to modulation, channel code, frame length, and signal-to-noise ratio. The formula has been verified with several practical channel coding schemes in junction with different modulation schemes. The practical channel coding schemes include a popular convolutional code with coding rate $r = \frac{1}{2}$, generator polynomial $[171, 133]_8$, and a constraint length 7, commonly utilized in space communications; a convolutional code with coding rate $r = \frac{1}{3}$, generator polynomial $[557, 663, 711]_8$, and a constraint length 9, adopted in W-CDMA standard; the turbo code with a rate $r = \frac{1}{3}$ code with generator polynomials $[1, 5/7, 5/7]_8$ and a block interleaver; and the LDPC code with a variable coding rate irregular code with maximum 20 iteration decoding number set and the frame length less than 3819 bits. Besides the FER approximation for single transmission, we also developed the FER approximation for general receiver diversity systems, such

as for the SIMO system with MRC and HARQ system with Chase combining. It shows our proposed FER approximations can accurately predict the actual FER performances for many practical wireless communications.

Second, the energy efficiency design has been studied for both ARQ and HARQ systmes, where the average energy consumption per information bit is minimized. A practical energy consumption model is firstly built considering many system parameters, including the hardware power consumption, modulation, channel coding, and FER in the physical layer, and frame length and protocol overhead in the media access control layer. For ARQ system, to minimize the energy consumption, the joint design for the transmission energy and frame length has been optimized. For HARQ system with Chase combining, to minimize the energy consumption while ensuring the QoS requirement, the transmission energy sequence in each transmission round has been optimized. The closed-form solutions have been derived for both ARQ and HARQ systems while minimizing the energy consumption per information bit.

Third, the fundamental trade-off between energy efficiency and spectral efficiency has been analyzed for both ARQ system and HARQ system with Chase combining. A simple but effective new metric has been proposed to facilitate the analysis. The new metric is the normalized energy consumption per information bit by the spectral efficiency. It shows that for both systems, the minimum energy consumption is achieved at the cost of spectral efficiency, and vice versa. However, the spectral efficiency can always be largely increased while sacrificing littler energy consumption, comparing with the minimum energy consumption case, by minimizing or reducing our proposed new metric. Therefore, it exists the balanced trade-off between these two metrics. For ARQ system, the closed-form solution for the joint design of transmission energy and frame length has been provided to achieve this balanced trade-off. For HARQ system with Chase combining, to achieve the balance,

the numerical search algorithm has been provided for the determination of the transmission energy in each round.

Fourth, besides the above fundamental analysis for point-to-point system, we designed a frequency-domain on-off accumulative transmission (OOAT) scheme for the network energy efficiency and spectral efficiency. Due to the nature of wireless transmission channel, the interference is always existed for wireless communications. Therefore, to achieve both high energy efficiency and spectral efficiency for wireless networks, we take advantage of each received packet, even it is collided by interferences. We resort to the powerful digital signal processing capability and frequency resource management among multiple nodes to achieve this goal. It shows that each node can achieve the collision-tolerate transmission with our proposed frequency-domain OOAT scheme, even multiple nodes transmit simultaneously.

6.2 Future Works

There are several possible directions for the future works, some of them we are exploring.

First, for the point-to-point system, we only fully considered the ARQ and HARQ case for the energy efficiency, spectral efficiency, and balanced trade-off between EE and SE, where the transmitter retransmits its own information without the help of other nodes. Therefore, it arises a question that what is the optimum transmission scheme for either energy efficiency, spectral efficiency, or the balanced trade-off between these two metrics, for cooperative communications. We want to firstly investigate a basic situation, where there is only one relay node. For the relay node, it purely help to forward the information bits from the transmitter to receiver, but not transmit its own information bits. There are two types of basic cooperative communication, which are the amplifier-and-forward (AF) scheme and decode-and-forward (DF) scheme. We are exploring the

optimum transmission schemes for these two types of cooperative communication for different metrics of performance evaluation. We will then be able to get some conclusions that in what situations non-cooperative communication, AF cooperative scheme, or DF cooperative scheme is preferred for different performance requirements in different scenarios.

Second, we want to extend our work to network performance evaluation, such as the network energy efficiency, network spectral efficiency, and the balanced trade-off between the two metrics. For non-cooperative communication, our results for ARQ and HARQ can be easily extended to wireless networks. However, what are the optimum transmission schemes for cooperative wireless networks with AF scheme and DF scheme? Here we will consider each relay node not only forward the information bits from other nodes, but also transmit its own information bits to its destination. Therefore, it arises three questions that 1) Which relay node(s) should a general node choose among multiple potential relay nodes? 2) For each node, how many bits should be transmitted for its own information, and how many bits should be forwarded for other nodes? 3) For each node with energy constraint, how much energy should be assigned to transmit its own information and how much should be utilized for forwarding information bits from other nodes?

Finally, we believe there exist advanced communication techniques can improve both the energy efficiency and spectral efficiency. Therefore, we are also exploring some other advanced communication schemes, such as the conjunction of network coding, interference alignment management, rateless code, and effective DSP detection algorithms.

Chapter 7

Vitae

Gang Wang was born in 1986, Hubei, China. He received the B.S. degree from the Northwestern Polytechnical University, Xi'an, China, in 2005, and the M.S. degree from Beihang University (with honor), Beijing, China, in 2008, both in electrical engineering. He then served as a wireless system engineer in New Postcom Equipment Co., Ltd, in Beijing, from 2008 to 2010. He is now a Ph.D. student at the Department of Electrical Engineering, University of Arkansas. During July 2013 to December 2013, he was an intern at Intelligent Fusion Technology, Inc., Maryland, mainly contributed in two projects related with large mobile network performance evaluation and jammer mitigation techniques development. His current research interests include energy-efficient system, convex optimization and nonlinear programming, cross-layer design, wireless networks, cooperative communications, and game theory, etc. He is a member of Golden Key International Honour Society.