Energy Aware Network Coding in Wireless Networks

by

Xiaomeng Shi

B.Eng. Electrical Engineering,
University of Victoria, Canada (2005)
S.M. Electrical Engineering and Computer Science,
Massachusetts Institute of Technology (2008)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

Author ....
Department of Electrical Engineering and Computer Science
August 30, 2012

Certified by ....
Muriel Médard
Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Leslie A. Kolodziejski
Chairman, Department Committee on Graduate Students

Energy Aware Network Coding in Wireless Networks

by

Xiaomeng Shi

Submitted to the Department of Electrical Engineering and Computer Science
on August 30, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in
Electrical Engineering and Computer Science

Abstract

Energy is one of the most important considerations in designing reliable low-power wireless communication networks. We focus on the problem of energy aware network coding. In particular, we investigate practical energy efficient network code design for wireless body area networks (WBAN). We first consider converge-cast in a star-shaped topology, in which a central base station (BS), or hub, manages and communicates directly with a set of nodes. We then consider a wireless-relay channel, in which a relay node assists in the transmission of data from a source to a destination. This wireless relay channel can be seen as a simplified extended star network, where nodes have relay capabilities. The objective is to investigate the use of network coding in these scenarios, with the goal of achieving reliability under low-energy and lower-power constraints.

More specifically, in a star network, we propose a simple network layer protocol, study the mean energy to complete uploads of given packets from the nodes to the BS using a Markov chain model, and show through numerical examples that when reception energy is taken into account, the incorporation of network coding offers reductions in energy use. The amount of achievable gains depends on the number of nodes in the network, the degree of asymmetry in channel conditions experienced by different nodes, and the relative difference between transmitting and receiving power at the nodes. We also demonstrate the compatibility of the proposed scheme with the IEEE 802.15.6 WBAN standard by describing ways of incorporating network coding into systems compliant to the standard.

For a wireless relay channel, we explore the strategic use of network coding according to both throughput and energy metrics. In the relay channel, a single source communicates to a single sink through the aid of a half-duplex relay. The fluid flow model is used to describe the case where both the source and the relay are coding, and Markov chain models are proposed to describe packet evolution if only the source or only the relay is coding. Although we do not attempt to explicitly categorize the optimal network coding strategies in the relay channel under different system parameters, we provide a framework for deciding whether and where to code, taking into account of throughput maximization and energy depletion constraints.

Thesis Supervisor: Muriel Médard
Title: Professor of Electrical Engineering

Acknowledgments

My deepest gratitude goes to my advisor, Professor Muriel Médard, for her guidance, care and support, and encouragements throughout the course of my graduate studies. She has taught me, starting from the very basics, how to approach a problem, how to think critically, and how to work through technical details without losing focus of the central issue to be resolved. Her immense technical knowledge, expertise, and insights have always been a constant source of inspiration. I feel greatly privileged to have had the chance of working with her. On a personal level, Muriel has been a pillar of support for me through many, many tough times. Her interactions with people have been a great lesson in professionalism and compassion, while her appreciation of professional activities has had a deep impact on my view towards duties and responsibilities. I am deeply indebted to her for her invaluable advice and care, both professionally and personally.

I would like to thank my thesis committee members, Professor Anantha Chandrakasan and Professor Dina Katabi, for their guidance and insightful feedbacks during the completion of this thesis. I was very fortunate to have had the opportunity to work under their advice on several occasions. Many thanks to Professor Katabi and her students for providing the resources for experimental works related to this thesis.

I have met many good friends while at MIT, who have made my years here pleasant and memorable. Without listing everyone by name, I want to express special thanks to my officemates and groupmates, past and present, for the innumerable wonderful discussions, technical and otherwise. As colleagues, you all have helped me in my learning process in this group. As friends, you all have given me the gift of companionship.

Last but not least, I would like to thank my parents and my husband Li, for always putting up with my tantrums. Thank you for giving me strength. I could not have finished this thesis without your unconditional love and care.

This thesis is based on work supported by

- The Interconnect Focus Center (IFC), one of the six research centers funded under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation (SRC) program, under subcontract # RA306-S1.
- The NSERC Postgraduate Scholarship (PGS) issued by the Natural Sciences and Engineering Research Council of Canada.
- The Claude E. Shannon Research Assistantship from the Research Laboratory of Electronics at MIT.
- The Air Force Office of Scientific Research (AFOSR) under award number 016974-002 S.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Past developments in wireless communication systems have focused mostly on increasing throughput, maximizing system capacity, reducing end-to-end transmission delays, and enhancing data security. However, recent advancements in down-scaling the size of electronic devices and integrating systems onto single chips have enabled a design shift to network architectures composed of low power, short range, wireless devices. An example where energy efficiency becomes especially important is wireless sensor networks (WSN), in which small scale sensor nodes are typically powered by batteries or energy harvesting devices, and are required to operate for extended time periods, without much heat dissipation.

In this thesis, we consider energy efficient network protocol designs in a special WSN – Wireless Body Area Networks (WBANs), which consist of energy-constrained sensors attached to or implemented in the human body, and an energy-abundant control base station (BS) placed within close proximity to the sensors. Sensors can be arranged into a star topology, where each node communicates directly with the base station; they can also be arranged into relay networks, with 2-hop links between sensors and the base station. We consider the use of linear network codes as a mean to reduce the overall expected energy used for successfully delivering a packet from a sensor to the data-collecting base station. Although still a relatively young field, network coding has shown great potential in addressing issues such as network throughput, robustness, and security [29,63]. Our goal is to develop protocols to incorporate network coding into the system architecture of WBANs, and to study the potential gains in energy consumption and throughput that network coding may offer. In short, we seek to determine the best coded scheme to ensure, with low energy, sensed signals are reliably communicated to the base station.

In the remaining sections of this chapter, we shall discuss briefly considerations pertaining to energy efficient network protocol designs, give some background on WBANs, and present a few selected works to give an overview of network coding. An outline of each chapter and a summary of the main contributions of this thesis are given at the end.

1.1 Energy Efficient Network Protocol Designs and WBANs

A typical wireless radio operates in four modes: transmit, receive, idle, and sleep, where maximum power is consumed during transmission, and least is consumed during sleep. For energy efficiency, circuits can be partially turned off during idle mode to maintain only essential functions such as keeping an internal clock, or receiving beacon signals for network synchronization. However, the process of waking up from hibernation (idle or sleep) states, turning off to enter the hibernation states, and switching between transmit and receive modes can also consume a non-trivial amount of power.

In addition to low-power hardware design, energy-efficient wireless systems require low-power strategies throughout the entire network protocol stack [36]. In the physical and data link layers, error control schemes such as Forward Error Correction (FEC) coding and Automatic Repeat reQuests (ARQ) can be used to conserve power. During packetized data transmissions, collisions should be avoided as much as possible in the medium access (MAC) layer to reduce wasteful retransmissions. Wireless terminals should also be given data transmission starting and stopping times to facilitate switching between different operating modes: a transceiver should fall back into idle or sleep mode whenever it determines that it will not be transmitting or receiving for a period of time. Also, wireless terminals should be allocated contiguous slots for transmission or reception to reduce energy consumption for turn-arounds. Moreover, computation of the transmission schedule should be relegated to an central energy-abundant base station. For one, individual sensor nodes may not hear reservation requests from all other nodes; for two, distributed computation usually consume more power collectively. Depending on network lifetime, connectivity, and coverage constraints, prioritized transmissions from energy-scarce nodes may also be arranged. On the network layer, routing schemes can be established under energy constraints, such that nodes are uniformly depleted of battery power, maximizing the connectivity of the network. Depending on the networking protocol under consideration, transport and application layer techniques may also be employed to further reduce the energy consumed per successfully delivered packet.

Body Area Networks (BANs) present numerous application opportunities in healthcare, sports, and other areas where personal information is to be stored and shared with another individual or a central database [69]. One example is wearable medical monitors which can relay patients' vital information to physicians or paramedics in real time; another is performance monitors which can assist athletic training. A BAN can be constructed on either wearable electronic textiles [64], or through wireless links using new or existing wireless protocols. Sensors are embedded in or attached to the human body. They also function as transceivers to send measurements to a personal server, which we call a base station (BS). The base station can be a PDA, a cell phone, or a dedicated device. This central receiver communicates with remote servers or databases. The focus of this thesis will be on wireless body area networks (WBANs), where each sensor communicates to the base station through radio links. The central communication problem is to ensure reliable transmission of measured data to the base station in a timely and robust fashion.

Various design challenges exist in the construction of WBANs. Unlike in the wirelined case, where communication links do not vary significantly when the body under monitoring is in motion, in a WBAN, body movements can bring frequent and unpredictable link failures, as well as changes in channel fading, as movements cause antennas to change in orientation. Unlike in general WSNs, a WBAN is in close proximity to the human body, where absorption of emitted power can also alter the channel response. Reference [91] provides a summary of channel modeling studies conducted and submitted to the IEEE 802.15 WPAN Task Group 6 (TG6) [2], which was formed in 2007 with the goal of establishing a standard suitable for WBANs. Depending on the type of sensors (implant, body surface, or external), the location of sensors on the body, and the frequency band under consideration (narrow or ultrawide), propagation and path loss models can be established with different parameters. In this thesis, we do not take into account details of the physical channel characteristics. Instead, we consider a packet erasure channel abstraction for the link layer model.

Another issue in the design of WBANs is that, compared with cellular networks where packet losses can sometimes be tolerated for voice transmissions, medical data to be transmitted in a WBAN have much more stringent requirements on reliability. Most importantly, while terminals in wireless networks such as cellular or WiFi usually have abundant memory and enough power, sensors in WBANs are limited in both size and energy. As sensors are attached to the human body, WBANs are under severe energy and power constraints. Long system lifetimes and low emitting power are desired. To solve the challenges on reliability and energy conservation, efficient error control, channel access, and resource allocation mechanisms are needed. Medium Access Control (MAC) protocol design is therefore one of the most important issues in WBAN development.

One approach to establishing a WBAN is to modify existing WSNs to suite the need of WBAN systems. References [11, 15, 83] compare WBAN with traditional WSNs and give comprehensive overviews of recent research efforts in the design of WBAN systems, particularly in terms of sensor device design, physical layer schemes and data link layer protocols. In WSNs, energy is often wasted in medium access collisions, idle listening, and protocol overheads when the desired data rate is low. WBANs, on the other hand, typically contain only a limited number of nodes, all positioned close to the base station. A single master-slave architecture with time division medium access (TDMA) therefore usually suffices to remove much of the energy wastage seen in a WSN. For example, Reference [67] implements such an architecture, with adjustable wakeup fallback times to mitigate possible slot overlaps. References [59], on the other hand, considers a TDMA scheme in a two level hierarchical topology, where intermediate master nodes also communicate with a central monitoring station. Other low-power MAC protocols such as T-MAC [86], S-MAC [92], Wise-MAC [23], DQBAN-MAC [68], and BodyMAC [26] have also been proposed to introduce various degrees of synchronization into the transmission schedule. Moreover, a WBAN-specific MAC protocol can be obtained by adjusting parameters of the IEEE 802.15.4 standard [4] to achieve energy efficiency [49, 50]. However, direct modification of existing standards also introduces redundant communication modes, which may lead to unnecessary complexities during practical implementations of WBAN systems.

Very recently, the IEEE 802.15.6 standard [6] has been published by the IEEE 802.15 Task Group 6 (TG6) [2] to overcome the limitations of other Personal Area Network (PAN) standards such as IEEE 802.15.4 (Zigbee) [4] and IEEE 802.15.1 (Bluetooth) [1]. Since its initial draft proposal [21, 22, 46], many studies have been conducted to analyze the performance of medium access mechanisms of this standard. For example, [84] presents theoretical throughput and delay limits and bandwidth efficiency results for ideal channels without transmission errors, under different payload sizes, while [77] analyzes the energy lifetime of periodic scheduled access modes. In this thesis, we are also interested in assessing the potential of incorporating the use of network codes into the WBAN standard. As part of Chapter 2, we will give a quick overview of IEEE 802.15.6, and show how network coding can be included with minimal changes to the MAC layer, with discussions on the potential benefits of using a coded system.

1.2 An Overview of Network Coding

Unlike traditional routing, network coding encourages nodes within a network to mix data before forwarding [28, 29, 63, 93]. Data flows are viewed as information which can be combined algebraically. As long as enough degrees of freedom, in the form of linearly independent coded packets or symbols, are received at a destination, the original uncoded data can be recovered.

The concept of network coding is introduced by Ahlswede *et al.* [7] in their seminal paper, which shows that digital network coding achieves the network capacity of wirelined multicast connections. Traditional routing techniques can achieve the capacity bounds of unicast connections; network coding thus extends the Min-Cut Max-Flow theorem from unicast to multicast. Li. *et al.* [52] then show the sufficiency of linear codes for multicast. Koetter and Médard [44] provide an algebraic framework for such coded networks, and show that the problem of finding a capacity-achieving network code is equivalent to satisfying algebraic conditions abstracted from the underlying network topology. Ho *et al.* [33]

introduce random linear network codes (RLNC), showing that it achieves multicast capacity probabilistically in a distributed manner, exponentially in the finite field size for coding operations. Other low complexity linear network codes are concurrently proposed for multicast connections [34, 35]. Aside from scalar algebraic codes, block network codes [60] and convolutional network codes [24, 27, 51] have also been proposed to mitigate constraints on large field size [31], to accommodate cycles within a network, and for achieving network capacity bounds in non-multicast settings.

In addition to its network capacity achieving advantages over traditional routing, network coding can provide reliability and robustness to errors and losses. From a code design perspective, Koetter and Kschischang [43] consider the problem of error-control in random linear network coding by exploiting the vector-space preserving property of linear network codes, while Silva *et al.* [75] investigate subspace codes under the rank-metric to show its application to error control in random network coding. From an operational perspective, Lun *et al.* discuss the use of random linear network codes in packet erasure networks, both wirelined and wireless, by separating the problem into two parts: network coding, and subgraph selection [56, 61]. The network coding problem decides what coding operations are performed at each individual node, given the rates at which packets are injected onto each link; the subgraph selection problem decides appropriate packet injection rates. Efficiency in the coding operations can be measured through different metrics. For example, benefits in energy reduction can be demonstrated by considering transmission energy as the objective cost during subgraph selection. Traskov *et al.* [79] further extend this framework with a scheduling constraint for wireless networks. Furthermore, Dana *et al.* [20] derive the capacity for a class of wireless erasure networks with no interference at reception and show that linear coding suffices to achieve the capacity region.

The two-hop tandem network shown in Figure 1-1 is a straight forward example to show the advantage of using linear network codes over end-to-end forward error protection schemes. Assume that data can be injected into outgoing links at the rate of one per time slot by the source s and the intermediate node r. Let packet erasure rates be ε_{sr} and ε_{rd} on the two links. With end-to-end erasure coding, the achievable rate from s to d is $(1 - \varepsilon_{sr})(1 - \varepsilon_{rd})$. On the other hand, network coded data flow can be re-encoded at r through

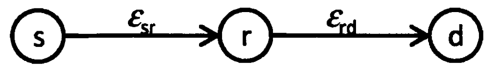


Figure 1-1: A tandem network consisting of two point-to-point links.

the generation of additional mixtures from those received successfully from s. In this case, a higher rate of $\min(1-\varepsilon_{sr}, 1-\varepsilon_{rd})$ can be achieved. Observe that coded data can be transmitted in a rateless fashion, similar to fountain codes. The advantage of network coding here lies in its composability – no decoding is necessary at the intermediate node r.

Several practical approaches of digital network coding have been proposed to bridge theoretical studies to real applications. To take into account randomly generated coding coefficients when RLNC is applied, a generation-based scheme with concurrently transmitted coefficients is considered in [16] to allow network coding to co-exist with current network protocols. Simple XOR codes can be used in the implementation of network codes in two-way relay networks and four-way cross topologies [40] by inserting an additional network coding layer between the IP and MAC layers of the 802.11 protocol. Here opportunistic coding schemes are experimentally shown to provide substantial throughput gains. Reference [96] further explores the effect of fairness has on coding opportunities. In addition to throughput gains, a network coding approach can be shown to eliminate much of the centralized coordination necessary for opportunistic routing [12]. Apart from such generation-based schemes, a sliding-window approach very similar to that used in the Transmission Control Protocol (TCP) for flow and congestion control can be introduced for network coding, such that end-to-end packet losses in wireless networks using the TCP protocol can be masked [42,62,76].

The important potential of network coding in providing improved efficiency in lossy networks in a distributed manner makes it an attractive option for energy constrained WBANs. The rateless nature of network codes can lower energy consumption by reducing the amount of medium access control needed, and reducing energy used by sensor nodes to wake up periodically to receive timing allocations from the control base station [74]. In this thesis, we want to understand when and where to use network coding in a WBAN, and how to exploit it to achieve reliability with reduced energy use. We start by extending the single-hop time division duplex system discussed in [54] and [55] to a star topology, and analyze the optimal transmission scheme to minimize energy use.

In addition to performing network coding in the digital domain, in which data are processed above the physical layer, an alternative scheme that explores the random nature of wireless links is analog network coding (ANC) [37, 38, 95]. With analog network codes, data are mixed on the signal level, with collisions caused by interferences from multiple nodes viewed as naturally occurring codes. By joint relaying and analog network coding, a throughput gain of two can be achieve in a two-way relay channel at high SNR levels [39]. A very useful application of analog network coding is in combating the hidden terminal scenario in an 802.11 wireless network. Instead of using a contention scheme where collisions from different terminals lead to exponential backoffs, consecutive collisions can be viewed as linear combinations of the same set of packets. An access point can therefore decode the received mixtures in a zig-zag fashion [30]. One potentially critical issue with the application of analog network codes in a body area network is that it operates with acceptable packet error rates only at high SNR levels [58]. It is not immediately clear what fundamental tradeoffs exist between digital and analog network codes, and how to determine an operating point, given a network topology and wireless channel model. In this thesis, we focus on the use of digital network codes only.

1.3 Main Contributions and Thesis Outline

The main contributions of this thesis are as follows.

- Application of network coding to a star-shaped wireless body area network: we propose a Markov chain to model the packet transmission process by a given number of sensor nodes to a central data-collecting base station. This model illustrates the interdependence of energy use among sensor nodes, which wake up at pre-determined times to receive new medium access schedules. The Markov chain formulation and the analytical framework we provide can be used to measure the performance of the proposed coded transmission scheme in aspects other than energy, as along as an appropriate metric is defined. We also discuss a simple heuristic that can be used in practical implementations of WBANs. Part of the work presented in this chapter was published in [72].

- Incorporation of network coding into the IEEE 802.15.6 standard: we examine the recently published IEEE 802.15.6 WBAN standard, and consider the transmission of coded data using this standard. We provide discussions on both a generation-based approach, and a sliding window approach, under a modified block acknowledgment mode, or a proposed Group-Block acknowledgment (GB-Ack) mode. With advancements in low-power network coding accelerator design, it is more likely that such incorporation of network coding into systems compliant to the WBAN standard will occur in the future.

- Analysis to determine when and where to code in a wireless relay network, thus assessing the potential of using relays in a WBAN. We propose a Markov chain to model the evolution of innovative packets at each node within a half-duplex relay network, and numerically study the expected energy for delivering a packet successfully to the receiver, when coding is performed at either or both the source and the relay. Our analysis is parametrized to provide a framework for evaluating the cost and benefits of using network coding, even though we do not attempt to explicitly categorize the optimal network coding strategies under different system parameters. Part of the work presented in this chapter was published in [73].

This thesis is organized as follows. Chapter 2 considers the problem of finding an optimal transmission schedule in a coded star network, and present a Markov Chain model for the analysis. A simple heuristic is provided as an estimate for the optimal scheme. We then discuss in detail the incorporation of network coding into the IEEE 802.15.6 standard. Chapter 3 considers the problem of finding an optimal transmission schedule in a coded wireless relay network, with the goal of determining where and when to code, taking into account of throughput and energy performances. We consider a dedicated half-duplex relay that assists in data packet transmission from a source node to a destination node. Finally, Chapter 4 presents a summary of the thesis and provides brief discussions on future research directions.

There are several other previous publications, [31,41,78], not directly incorporated into this thesis. These may be of interest to the readers as supplementary readings.

Chapter 2

Packet Erasure Star Network

Unlike general wireless sensor networks, a wireless body area network (WBAN) contains only a limited number of nodes, all positioned close to the base station. A single-hop master-slave architecture with time division medium access (TDMA) thus seems to be a natural choice for the system setup. A star-shaped WBAN differs from a one-hop cellular network, as the amount of data uploaded from the sensors to the base station much outweighs the amount of control signals downloaded from the base station. In this chapter, we consider a simple scheduled MAC protocol and study the mean energy to complete the overall converge-cast transmission from the sensors to the base station.

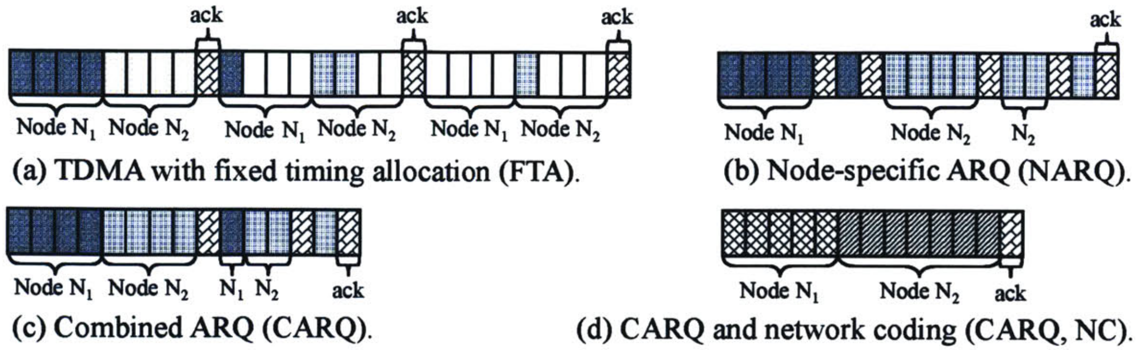
Our proposed scheme and the corresponding analysis differ from previous studies on energy use in WBANs in three aspects. Firstly, we take into account of energy spent by sensors on the reception of acknowledgment packets from the base station. Generally, acknowledgment packets are much shorter than data packets. Nonetheless, in WBANs, data signals can be infrequent and very short, depending the biometric being measured. Thus data and acknowledgments can be comparable in terms of packet length. In addition, depending on the underlying circuit design, the energy cost of listening and transitioning between transmit and receive modes can be significant. Secondly, we allow the use of linear network codes in the transmission process, simplifying the acknowledgment signal, while allowing redundancies to be added on the packet level, similar to forward error correction codes on the symbol level. For one, network coding can be viewed as a reliability mechanism here on the link level. For two, a coded scheme allows the nubmer of acknowledgments to be reduced when listening is costly. Lastly, our analysis takes into account dependencies among energy use by individual sensor nodes, through periodic wakeups to receive acknowledgments and scheduling information from the base station.

The rest of the chapter is organized as follows. Section 2.1 gives a simple example to show the potential energy gains of applying network coding to transmissions in a WBAN. Section 2.2 describes the network and energy model, and the network coded algorithm. Section 2.3 provides a Markov chain model to analyze the optimal number of packets to transmit by each sensor node, and gives a heuristic as an estimate of the optimal scheme. Numerical results are presented in Section 2.3, comparing the network coded scheme with uncoded scheme in terms of completion energy. Section 2.4 discusses the IEEE 802.15.6 WBAN standard and the incorporation of network coding to block data transmissions. Section 2.5 looks at some of the issues and challenges we encountered in a preliminary software-defined radio implementation of the proposed coded transmission scheme. Section 2.6 concludes this chapter.

2.1 Example: Network Coding Benefits in a Star Network

Before considering a general star network model, we first show through a simple example why network coding can be beneficial. As stated the beginning of this chapter, one category of energy use often overlooked in wireless networks is the reception energy spent on listening to control signals from the base station. In an energy-constrained system such as a WBAN, such reception energies can have more significant effects on node depletion time since data rate is much lower, but control signals need to be transmitted frequently for medium access purposes. In addition, energy cost for turnarounds between transmission and reception modes can be significant, thus it may be beneficial to reduce the total number of times acknowledgments are communicated from the base station to individual sensors.

Let us consider a two-sensor star network with nodes N_1 and N_2, each trying to directly upload 4 packets to a base station through the same frequency band. In the link layer, assume the packet erasure probabilities are time invariant, at 0.2 and 0.4 respectively. Figure 2-1 shows instances of four different possible communication schemes, all based on



(a) TDMA with fixed timing allocation (FTA).

(b) Node-specific ARQ (NARQ).

(c) Combined ARQ (CARQ).

(d) CARQ and network coding (CARQ, NC).

Figure 2-1: Example comparing overall completion energy for 2 nodes, each with 4 packets to upload; the erasure probabilities are $p_1 = 0.2$, $p_2 = 0.4$. Shaded blocks represent data packets in transmission and ack packets in reception; white blocks represent time during which nodes are idle.

TDMA with Automatic Repeat reQuests (ARQs). Shaded blocks represent data packets in transmission and acknowledgment (ack) packets in reception; white blocks represent time during which nodes are idle. Some packets are lost during transmission probabilistically. Define a transmission round to be the transmission of data packets by one or more sensor nodes, followed by a broadcasted ack packet. Both nodes wake up at the end of a transmission round to listen to the ack, which contains retransmission requests and scheduled allocation periods for the next round.

Note that, here we have assumed the same number of 4 packets to be uploaded by each sensor. In real WBAN applications, it is possible that sensors have different data measurement and transmission rates. Also note that, we assume the given 4 packets are to be transmitted as a batch, with no new data arriving afterwards. Such an assumption is reasonable since medical data are often measured at low rates.

(a) *Fixed Timing Allocation (FTA)*: each node is allocated 4 slots per round. N_1 transmits its 4 packets first. With $p_1 = 0.2$, on average, one packet is expected to be lost. Similarly for N_2, two packets are expected to be lost as $p_2 = 0.4$. At the end of the first round, the base station acknowledges and identifies the missing packets for retransmission. Both nodes listen to this broadcasted ack signal to determine if they need to retransmit any lost packets or if they should start sending the next batch of packets. In this particular example, one packet is retransmitted by N_1, and two packets are retransmitted by N_2. Note that, since packets are uncoded, the acknowledgment is required to specify the packet indices of those to be retransmitted.

(b) *Node-specific ARQ (NARQ)*: each node transmits until all of its packets are received successfully by the base station. The two nodes transmit one after the other. The ack packet broadcasted by the base station contains retransmission requests for the actively transmitting node and scheduling information for both nodes. Note that, during the ack signal for N_1, node N_2 also needs to listen, otherwise N_2 does not know when its own transmission can start. The same applies to N_1: even after completing its own transmission, it needs to wakeup to receive the ack packet, otherwise it does not know when to start sending its next batch of packets. Similar to (a), the ack signal is required to specify the indices of packets to be retransmitted.

(c) *Combined ARQ (CARQ)*: both nodes are allocated specific transmission periods each round, with a combined ack packet broadcasted at the end. Ack for individual nodes are concatenated and transmitted together. All nodes wake up to listen during ack periods. Compared with (a), this approach uses the same amount of energy in transmission and reception, but provides a higher throughput, since no time slot is wasted in idle waiting.

(d) *Combined ARQ and network coding (CARQ-NC)*: in this scheme, linear network coding is employed, with a generation size of 4. The code can be random, with coefficients attached to each packet; it can also be deterministic, with an index attached to each coded packet, such that the base station can look up the coding coefficients in a pre-determined table. In addition to simply mixing the data packets, each node can send more coded packets than the number of degrees of freedom required for decoding. This design compensates for anticipated packet losses. The number of coded packets to transmit depends on the required number of degrees of freedom and the packet erasure rates. Since $p_2 > p_1$, more coded packets are sent by N_2 if the same number of degrees of freedom is requested by the base station from N_1 and N_2. Keep in mind that transmitting more packets to compensate for erasures is part of the coding scheme. This setup is similar to that proposed in [55], which studies the unicast, time division duplexing channel.

Table 2.1: Comparison of completion energy per accepted data packet; there are 2 nodes in the star each with 4 packets to upload. E_{TX} = total transmission energy; E_A = energy spent on listening to acknowledgement packets; E_{tot} = total completion energy per accepted data packet, η = throughput.

	E_{TX}	E_A	E_{tot}	η
(a) FTA	$12E$	$6E$	$9E/4$	$8/27$
(b) NARQ	$12E$	$10E$	$11E/4$	$8/17$
(c) CARQ	$12E$	$6E$	$9E/4$	$8/15$
(d) CARQ-NC	$12E$	$2E$	$7E/4$	$8/13$

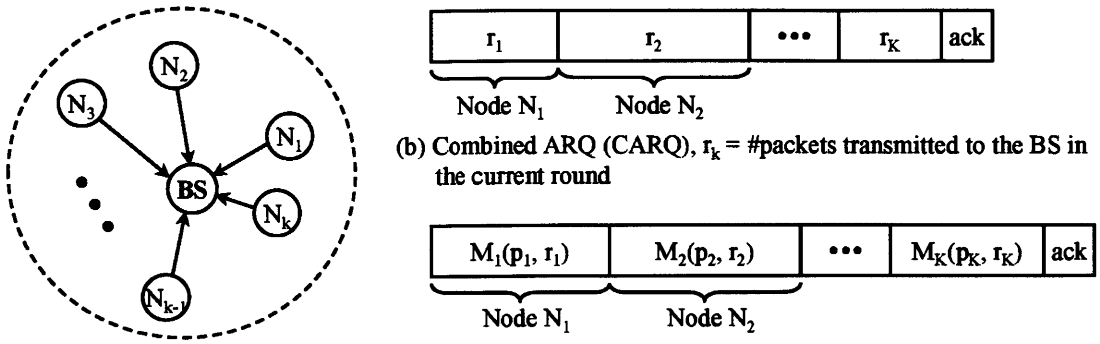
To evaluate energy use, assume every data packet transmission and every ack packet reception consumes an equal amount of E units of energy. Table 2.1 compares the total energy required for the schemes shown in Figure 2-1. Also shown is the throughput of each scheme, defined as the total number of accepted data packets divided by the total transmission time in units of packet slots. Excluding ack periods and time during which nodes are sleeping, all schemes require $12E$ in data transmission. On the other hand, the energy used for ack reception varies significantly across different schemes. CARQ-NC is the most energy efficient, since it combines acknowledgments for both nodes together, while transmitting coded packets with redundancies allows the total number of acknowledgments to be reduced. FTA requires less or the same amount of total energy than NARQ and CARQ, but is throughput inefficient. As the number of nodes in the network increases, this inefficiency will become increasingly severe. CARQ outperforms NARQ, and CARQ-NC introduces further gains.

Note that, it is not necessarily true that CARQ-NC always transmits the same total number of data packets as CARQ. In fact, CARQ-NC is expected to send *more* packets than the required number of degrees of freedom. Nonetheless, the added transmission energy is offset by reduced reception energy to give a smaller overall completion energy. This specific example is extremely simple, but very similar results can be expected as more sensor nodes are added and more packets are involved. In the remaining parts of this chapter, we will consider only the CARQ and the CARQ-NC scheme. Our goal is to determine analytically the optimal network coding and transmission scheme such that the expected completion energy for the overall transmission is minimized.

2.2 System Model and Problem Formulation

We model a WBAN with a star topology as shown in Figure 2-2(a): each of K sensor nodes communicates with the base station directly to upload M data packets. Nodes and the base station are assumed to operate in half-duplex mode, either transmitting or receiving, but not at the same time. A WBAN occupies a single frequency band, with the base station centrally coordinating a TDMA scheme. Exact synchronization among the nodes and the base station is assumed, and nodes return to sleep when not transceiving. Computation of the transmission schedule is relegated to the base station, with start and stop times allocated through acknowledgment packets from the base station. We assume ack packets are transmitted reliably, and propagation delays from base station to nodes are negligible. The channel between an individual sensor N_k, $1 \leq k \leq K$, and the base station are assumed to be memoryless, with packet erasure probability p_k, which is invariant during the time interval when the M packets are uploaded. Note that, individual sensor nodes can have different numbers of packets to upload. We have assumed the same number of M here for simplicity, but the proposed system model and the analysis in this chapter also apply to the case when data rates vary among different nodes.

The above system model may seem over-simplified, but is sufficiently accurate for the current study. As already discussed, the small physical size of a WBAN enables the use of a star topology with TDMA scheduling controlled centrally by a base station. Compared to sensor nodes, the base station is relatively unconstrained in power. Ack packets can therefore be piggybacked on a periodic synchronization signal transmitted at high power, or protected through error correction codes to ensure reliability. In an actual implementation, additional headers or beacon packets will be needed for synchronization, but such details can be safely omitted here in analyzing the data transmission energy efficiency and system throughput. Recall from Section 1.1 that an additional difficulty in WBAN design is channel modeling for physical layer designs. In this chapter, we only consider an erasure channel abstraction for the network layer model. The time-invariance assumption is a reasonable first step, since data in WBANs come in very small bursts periodically and the channel can be assumed to fade slowly over each such small periods.



(a) Star topology

(b) Combined ARQ (CARQ), r_k = #packets transmitted to the BS in the current round

(c) CAQR with network coding (NC). $M_k(p_k, r_k)$ = #coded packets transmitted to the BS in the current round, $1 \leq k \leq K$.

Figure 2-2: Uplink transmissions using combined ARQ and network coding. In a star network, individual sensors upload data directly to the base station (BS).

In the CARQ scheme, nodes take turns to transmit data packets before waiting for a combined ack packet, which contains repeat requests and scheduling information. Figure 2-2(b) illustrates one round of transmission, where r_k represents the number of packets transmitted by node N_k to the base station in the current round. Note that, each packet is acknowledged individually, thus the ack signal is required to specify the packet indices of those to be retransmitted in the next round.

In the network coding based scheme, each node linearly combines its M packets before taking turns to transmit the ensuing mixtures. Coding coefficients can be generated on the fly and attached to the data payload, or tabulated a priori, with table indices attached to each coded packet. We assume the field size is large enough such that accepted coded packets are independent from each other with very high probability. In practical implementations, the field size for coding operations is often small, and a tradeoff exists between energy used for computation, and energy used for retransmissions caused by linearly dependent packets received at the base station [8]. Since coded packets represent degrees of freedom rather than distinct data packets, the acknowledgment signal broadcasted by the base station at the end of each transmission round does not need to specify packet indices. Instead, it contains numerical values indicating the numbers of degrees of freedom still required from each sensor node for the base station to successfully decode all received packets. For example, if a total of 2 coded packets have been received successfully from node N_k, the base station requests $M - 2$ degrees of freedom from N_k for the next transmission round. In turn, each node can transmit more than the requested number of degrees of freedom to compensate for packet losses. In this example, N_k can send more than $M - 2$ coded packets in the subsequently round.

Figure 2-2(c) illustrates one round of coded transmission. M_k represents the number of coded packets for (re)transmission. M_k is a function of p_k, the erasure probability, and r_k, the remaining number of degrees of freedom needed at the base station to decode successfully. Note that, $M_k \geq r_k$. Since all nodes within the network need to wakeup from sleep modes to listen to the ack, reducing the total number of ack packets can reduce the total energy consumption, if the energy cost for listening is non-negligible. M_k should also be kept small to minimize redundant transmissions. We want to show that an optimal number, M_k, $1 \leq k \leq K$, of coded packets exists to minimize the mean completion energy.

To evaluate the amount of energy spent to successfully deliver a packet from a sensor to the base station, we assume sensor nodes operate in three modes: transmit, receive, and sleep. Denote the processing and transmission energy per data packet by $E_{p,CARQ}$ and $E_{p,NC}$ for the uncoded and coded cases respectively; also denote the reception and processing energy per ack packet by $E_{a,CARQ}$ and $E_{a,NC}$. When in sleep mode, most circuit components are assumed to be turned off such that energy consumption is negligible. Let $E_{a,CARQ} = \alpha E_{p,CARQ}$, $E_{a,NC} = \alpha E_{p,NC}$, where the parameter α can take on different positive values depending on the underlying circuit design and protocol designs. For example, in narrow-band systems where transmission power is approximately the same as receiving power, α is the ratio between lengths of ack and data packets. For short range ultra-wide band systems where transmission energy per bit is much smaller than the reception energy per bit, α can take on large values in the range of tens to the hundreds [19, 65, 71]. Moreover, assume $E_{p,NC} = (1 + \beta)E_{p,CARQ}$, where the non-negative factor β represents the additional energy needed to perform network coding [8].

2.3 Markov Chain Model

To study the energy use of uploading data from sensor nodes to the base station, we model the communication process using a Markov chain. Let state $I = (i_1, \ldots, i_K)$ represent the degrees of freedom requested by the base station from nodes $(N_1, \ldots, N_K)$ for the next round of transmission, where $0 \leq i_k \leq M$, $1 \leq k \leq K$. The overall communication process initializes in state $\mathbf{M} = (M, \ldots, M)$ and terminates in state $\mathbf{0} = (0, \ldots, 0)$. Let $c_{i_k,k} = M_k(p_k, i_k)$ denote the number of coded packets node N_k transmits when it sees a packet erasure probability of p_k and the base station requires i_k additional degrees of freedom for decoding; $c_{0,k}$ is equal to zero since no data is requested by the base station. Assume packet losses occur independently across nodes. The transition probability from state $I = (i_1, \ldots, i_K)$ to state $J = (j_1, \ldots, j_K)$ is the the product of individual terms each corresponding to a sensor node:

$$P_{IJ} = P_{(i_1,\ldots,i_K)(j_1,\ldots,j_K)} = \prod_{k=1}^{K} P_{(i_k,j_k)}\,, \tag{2.1}$$

where $P_{(i_k,j_k)} = P_{i_k j_k}$ and

$$P_{i_k j_k} = \begin{cases} \begin{pmatrix} c_{i_k,k} \\ i_k - j_k \end{pmatrix} (1-p_k)^{i_k - j_k} p_k^{c_{i_k,k} - i_k + j_k} & 0 < j_k \leq i_k \\ \sum_{l=i_k}^{c_{i_k,k}} \begin{pmatrix} c_{i_k,k} \\ l \end{pmatrix} (1-p_k)^l p_k^{c_{i_k,k} - l} & j_k = 0 \end{cases}. \tag{2.2}$$

The value of $c_{i_k,k}$ is computed by the base station, and communicated to the sensor nodes through acknowledgment packets. Observe that an ordering exists among the states of the Markov chain: since the numbers of degrees of freedom requested from each sensor node is non-increasing, let $I \succ J$ if $P_{IJ} > 0$, $I \neq J$, i.e., state transitions are only possible as self-loops or from high to low states. This Markov chain has $(M+1)^K - 1$ transient states and one recurrent state, $\mathbf{0}$, which signals completion of the transmission. Figure 2-3 illustrates the case where there are $K = 2$ sensor nodes within the WBAN. Observe that $j_k = 0$ signals that the end state is along one of the edges of the Markov chain. Moreover, observe from Eq. (2.2) that when $0 < j_k \leq i_k$, $P_{i_k j_k}$ corresponds to a binomial distribution with success probability $1 - p_k$; the total number of trials equal to $c_{i_k,k}$. When $j_k = 0$, $P_{i_k j_k}$ corresponds to the tail of the binomial distribution, with $c_{i_k,k} \geq i_k$.

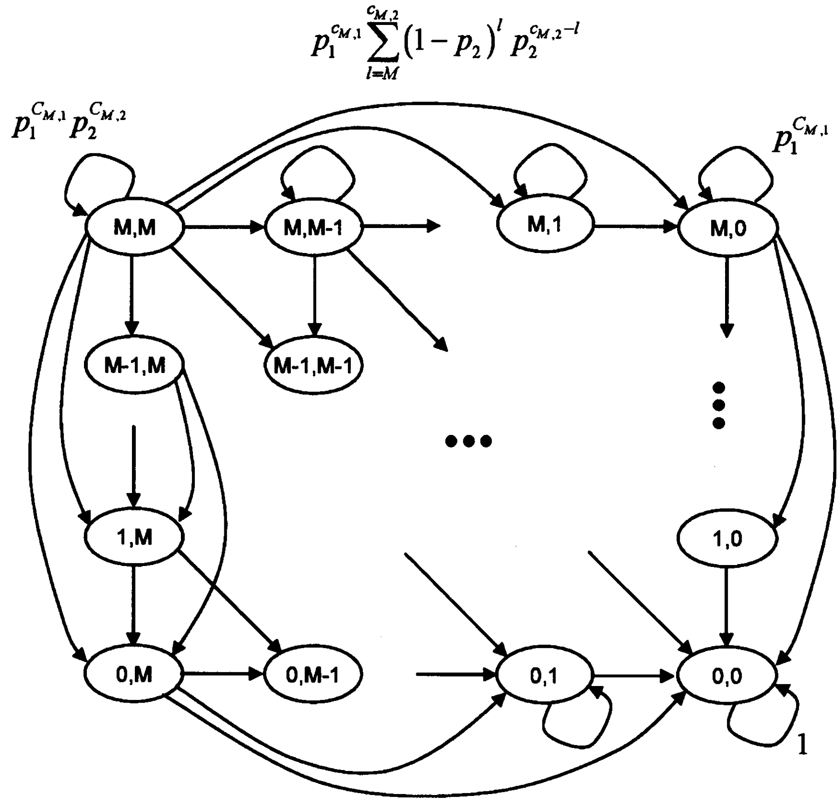


Figure 2-3: Markov chain representation of the network coded scheme. The number of nodes is $K = 2$. Each node has a total of M packets to upload to the base station. In state (i, j), the base station still requires i degrees of freedom from node N_1 and j degrees of freedom from node N_2 to successfully decode the original M data packets from each node.

2.3.1 Expected Energy for Completing Transmission

Let E_I denote the expected system completion energy when nodes $(N_1, \ldots, N_K)$ have $I = (i_1, \ldots, i_K)$ degrees of freedom to upload to the BS respectively, i.e., E_I is the expected amount of energy spent on uploading coded packets to the base station, such that i_k degrees of freedom are received from N_k by the base station at the end of the overall transmission process. The energy cost of uploading a single data packet is E_p, while the energy cost of receiving an ack packet is E_a. Thus, the following recursion holds

$$E_I = \frac{1}{1 - \prod_{k=1}^{K} p_k^{c_{i_k,k}}} \left\{ E_p \sum_{k=1}^{K} c_{i_k,k} + E_a K + \sum_{J \prec I} P_{IJ} E_J \right\}. \tag{2.3}$$

Optimal Number of Packets to Transmit

To minimize the expected completion energy $E_{\mathbf{M}}$, let

$$C = \{c_{i_k,k} | 1 \le i \le M, 1 \le k \le K\}, \quad c_{0,k} = 0, 1 \le k \le K, \tag{2.4}$$

we then have

$$C^* = \underset{C}{\operatorname{argmin}}\, E_{\mathbf{M}}, \quad E^*_{\mathbf{M}} = \min_{C} E_{\mathbf{M}}, \tag{2.5}$$

and the following recursion.

$$E^*_{\mathbf{M}} = \min_{c_{M,1} \dots, c_{M,K}} \frac{1}{1 - \prod_{k=1}^{K} p_k^{c_{M,k}}} \left\{ E_p \sum_{k=1}^{K} c_{M,k} + E_a K + \sum_{J \prec \mathbf{M}} P_{\mathbf{M}J} E^*_J \right\}. \tag{2.6}$$

Unlike the erasure probability P_{IJ}, the expected completion energy can not be separated into node-specific energy terms. One possible approach to this combinatorial optimization problem is to ignore the integer constraints and solve for $c_{i_k,k}$ iteratively by taking the partial derivatives of the objective function and finding the values that set these to zero. For the simplest case with $K = 2$ nodes and $M = 1$ packet to upload from each node, the optimal values for $(c^*_{1,1}, c^*_{0,2})$ and $(c^*_{0,1}, c^*_{1,2})$ can be computed explicitly with the Lambert W function, which is the solution to the nonlinear equation $z = W(z)e^{W(z)}$ for a given z [10, 14, 17]:

$$(c^*_{1,1}, c^*_{0,2}) = \left(\frac{1 + W(-\exp(\frac{2E_a}{E_p} \ln p_1 - 1))}{\ln p_1} - \frac{2E_a}{E_p}, 0 \right), \tag{2.7}$$

$$(c^*_{0,1}, c^*_{1,2}) = \left(0, \frac{1 + W(-\exp(\frac{2E_a}{E_p} \ln p_2 - 1))}{\ln p_2} - \frac{2E_a}{E_p} \right). \tag{2.8}$$

Note that the Lambert W function is multi-valued. We pick the negative values on the lower $W_{-1}(\cdot)$ branch in this evaluation. Reference [10] provides an analytic approximation for $W_{-1}(\cdot)$ with very small relative errors. Nonetheless, solving $(c^*_{1,1}, c^*_{1,2})$ analytically is not easy: the two unknowns can not be written in terms of each other in closed form. An alternative is to perform exhaustive numerical searches for the optimal values C^* on an integer grid. For given values of $\{p_k | 1 \leq k \leq K\}$, E_p, and E_a, we can recursively search on an M dimensional space of non-negative integers to find C^*. We will show numerical examples in the next section for such an optimal scheme.

Discussions

In a practical implementation of the proposed network coding based transmission scheme, the computation task can be imposed on the base station, not individual sensor nodes. Neither do the results need to be computed in real time. Instead, pre-computed values can be stored in a look-up table according to different packet erasure probabilities.

Nonetheless, observe from Figure 2-3 that there are $(M+1)^K$ states in the Markov chain. As M and K increase, the computational complexity of an exhaustive search approach becomes prohibitive even for an energy-abundant base station. In addition, the state transition probabilities from Eq. (2.2) are defined in terms of packet erasure rates p_k, $1 \leq k \leq K$, as well as $c_{i_k,k}$, the number of coded packets node N_k sends when the base station requires i_k additional degrees of freedom for decoding. In other words, the model proposed in this section represents a family of Markov chains parametrized by p_k and $c_{i_k,k}$. To cover the entire range of packet erasure probabilities, tabulation of the optimal $c_{i_k,k}$ values requires fine discretization of the parameter space $[0,1]^K$. Such discretizations, in turn, require considerable amount of memory for storage. It is therefore intuitive to ask whether the given model can be well approximated or simplified, or whether a heuristic for the evaluation of $c_{i_k,k}$ can be found as a good approximation or an initial search step, such that near optimal energy use can be achieved with reduced computation complexity and memory requirements.

A possible simplification of the proposed model is to approximate the value of $P_{i_k j_k}$ in the large packet number range, using Stirling's approximation for $\binom{N}{r}$. However, even though the computation of $P_{i_k j_k}$ is no longer combinatorial under this approximation, it does not reduce the complexity of the Markov chain itself. Considering the large packet number range is also not justified. We will show later with numerical examples that the benefit of taking into account reception energy is most significant when the channel conditions for different nodes are asymmetric. State transitions in the higher states (ie. when large numbers of degrees of freedom are requested from each node) by itself is *not* the bottleneck of the problem. Instead, it is the tradeoff between transmitting more packets while in the higher states and waiting for others to finish while in the lower states, that requires more delicate control of the transmission schedule.

A very simple heuristic is to use the approximation

$$\tilde{c}_{i_k,k} = \begin{cases} \left\lfloor \frac{i_k}{1-p_k} \right\rfloor & \alpha > 1 \\ c & \alpha \leq 1 \end{cases} . \tag{2.9}$$

where c is the smallest integer that satisfy the following inequality

$$\binom{c_{i_k,k}}{i_k - 1} (1-p_k)^{i_k-1} p_k^{c_{i_k,k}-i_k+1} \leq \sum_{l=i}^{c_{i,k}} \binom{c_{i,k}}{l} (1-p_k)^l p_k^{c_{i,k}-l} . \tag{2.10}$$

Recall that $\alpha = E_a/E_p$ represents the ratio between the amount of energy spent on listening and transmitting. When $\alpha > 1$, listening takes more energy, it is more desirable to send as much as possible, and to terminate the overall communication process in as few turns as possible. Thus, $\tilde{c}_{i_k,k}$ is taken as the optimal throughput of a stop-and-wait ARQ system when i_k packets are to be sent through a channel with packet erasure rate p_k. In other words, if $\alpha > 1$, each node N_k sends packets in advance on a best effort basis. Enough coded packets are sent such that i_k degrees of freedom are expected to be received at the base station. Of course, the actual number of packet lost in each round can be higher or lower than the expected value.

When $\alpha \leq 1$, transmission takes more energy. It is therefore more desirable to send a limited number of packets, and to request retransmissions as they are needed. Observe the optimization problem given in Eq. (2.6) and the structure of the Markov Chain shown in Fig. 2-3. Intuitively, the optimal path from the starting state (M, M) to the terminating state $(0, 0)$ should be along the diagonal of the Markov chain, rather than along the four edges. This is because along the edge of the Markov chain, at least one sensor node has completed its transmission, and is waking up only to receive necessary scheduling information, leading to energy wastage. On the other hand, transmitting a large number of coded packets at once leads to a higher probability of terminating the entire transmission in one step, or ending in an edge state. Such over-provisioning for packet losses should be avoided since then energy may be wasted during transmission. Taking into account of the tradeoff between energy for data packet transmission and energy for listening to ack packets, we estimate the value of $c_{i_k,k}$ by bounding the tail term of the binomial distribution in Eq. (2.2).

One note about the proposed simple heuristic is that it does not take into account the interdependence among different nodes. Unlike the worst link channel and the combined erasure effect heuristics given in [53] for a broadcast setting, here the transmission schedule of each sensor is determined independently. In other words, with this heuristic, the Markov chain is decomposed into multiple smaller ones, each representing the transmission state of a single sensor. Such a decomposition is not in line with what we have previously discussed, that since each node wakes up to listen to acknowledgment and scheduling information, even when a node does not have data to transmit, its behavior is dependent on other nodes in the network. Nonetheless, we will show through an example in the numerical results section that this heuristic can lead to a average expected completion energy that does not deviate much from the optimal scheme. Hence, our proposed Markov chain based analysis provides an optimal solution and serves as a bound for simpler transmission schemes, while our heuristic can be easily implemented in a practical system to provide a reasonable solution to the converge-cast problem in a star network.

2.3.2 Numerical Results

In this section, we provide numerical examples for the CARQ and CARQ-NC schemes to study the amount of energy reduction offered by network coding as system parameters vary.

Table 2.2 lists explicitly the solution to the optimization problem stated in Section 2.3 when there are $K = 2$ nodes within the network, each having $M = 4$ data packets to upload. The packet erasure probabilities are $p_1 = 0.2$ and $p_2 = 0.4$ respectively. Assume $\alpha = 1$ and $\beta = 0$. The first column states the remaining numbers of degrees of freedom

Table 2.2: Optimal numbers of coded packets to transmit by each node when the number of degrees of freedom requested by the base station is i_1 from node N_1, and i_2 from node N_2. There are 2 nodes in the star, each with 4 packets to upload. The packet erasure rates are $p_1 = 0.2$, $p_2 = 0.4$. Assume transmitting and receiving a packet takes the same amount of energy, $\alpha = 1$; also assume energy use for performing network coding is negligible, $\beta = 0$.

$i_1 \backslash i_2$	4	3	2	1	0
4	$(5,6)$	$(5,5)$	$(5,3)$	$(5,2)$	$(5,0)$
3	$(3,6)$	$(4,5)$	$(4,3)$	$(4,2)$	$(4,0)$
2	$(2,6)$	$(2,5)$	$(2,3)$	$(2,1)$	$(3,0)$
1	$(1,6)$	$(1,5)$	$(1,3)$	$(1,2)$	$(1,0)$
0	$(0,7)$	$(0,5)$	$(0,3)$	$(0,2)$	$(0,0)$

required by the base station from node N_1, while the first row states the remaining numbers of degrees of freedom required by the base station from node N_2. Transmission initiates at $(i_1, i_2) = (4, 4)$, and terminates at $(0, 0)$. Since N_2 sees a more challenging channel, it sends more coded packets than N_1, when the same number of degrees of freedom is requested. Observe that the number of coded packets sent by N_2 actually increases from 6 to 7 when $i_2 = 4$, and i_1 is decremented from 4 to 0. This is because N_1 is expected to complete its data transmission in a small number of rounds, thus N_2 would want to send more data packets such that it also completes its data transmission in a small number of rounds, to reduce the total number of times both wake up to listen to ack signals. The optimal solution minimizes the sum of all energy terms, taking into account of future rounds of retransmissions, and tries to reduce possible energy wastes. The optimal expected total completion energy is found to be $16.46E$, larger than $14E$ shown in Table 2.1, which illustrates only one possible channel realization.

Packet Erasure Probability p_2

Figure 2-4 extends the example in Table 2.2 to summarize the percentage reduction in expected completion energy per accepted data packet achieved by the network coding scheme when packet erasure probabilities are varied. A packet is said to be accepted by the base station if it is received successfully, and the percentage reduction is computed with respect to the CARQ scheme. Again, we assume $\alpha = 1$, i.e., $E_{p,NC} = E_{a,NC}$, $E_{p,CARQ} = E_{a,CARQ}$, and $\beta = 0$, i.e., $E_{p,NC} = E_{p,CARQ}$. The horizontal axis represents variations in the packet

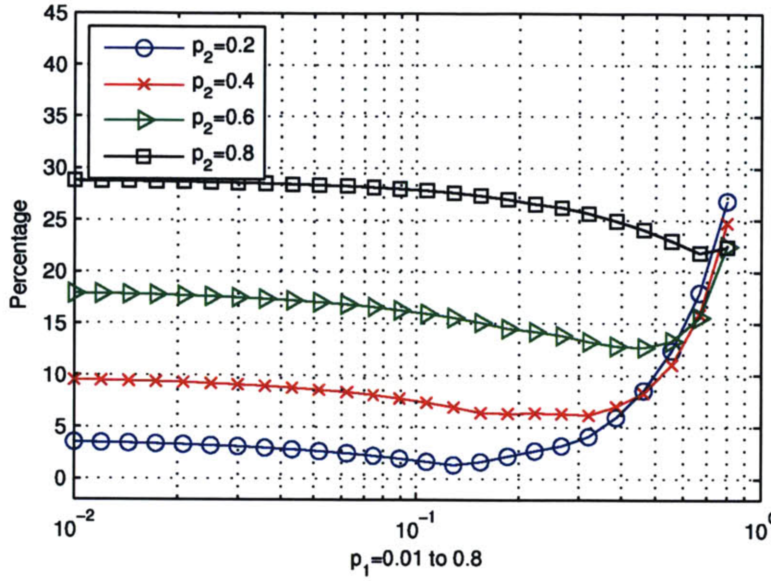


Figure 2-4: Percentage reduction in expected completion energy per accepted data packet as p_2 is varied; $M = 4$, $K = 2$, $\alpha = 1$, $\beta = 0$.

erasure probability of node N_1. Different curves correspond to changes in the erasure probability of N_2. As p_2 increases from 0.2 to 0.8, the amount of reduction in expected completion energy per accepted data packet increases from 3.5% to about 29%. Although not shown explicitly in this graph, the energy gain is derived from reduced number of transmission rounds. As p_2 degrades, the actual amount of energy spent for each accepted data packet increases, because more retransmissions are expected. Since depletion occurs more quickly when the channel condition worsens, the increased amount of saving is beneficial in extending the lifetime of sensor nodes. Another observation from Figure 2-4 is that the curves take a dip at different values of p_1. The locations of these minima correspond approximately to the values of p_2 in each case. This is because the network coding scheme achieves higher energy reduction when nodes experience more asymmetric channel conditions. When packet losses occur asymmetrically, nodes with more reliable channels complete data transmission quickly; yet they are forced to wake up for the ack signal repeatedly until other nodes with less reliable channels complete their transmissions. When nodes see similar channel conditions, with high probability, all nodes have non-zero number of packets to send each round, hence not as much energy is wasted in listening to the ack signals.

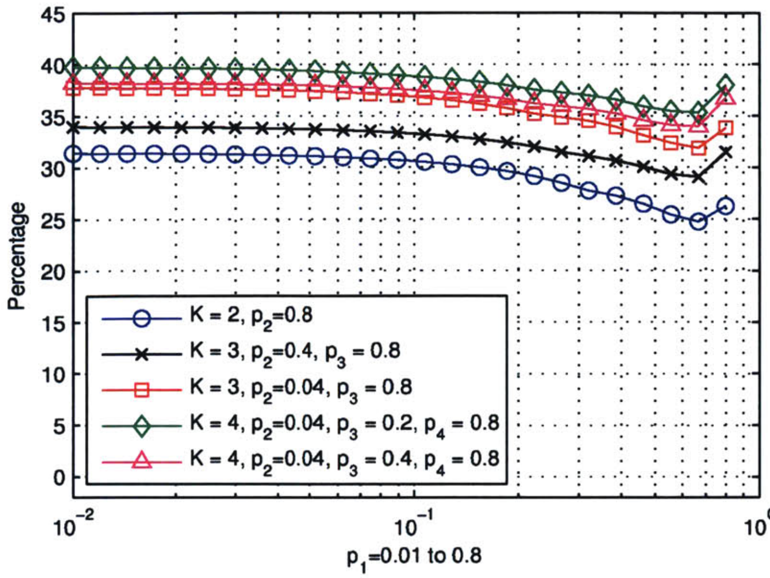


Figure 2-5: Percentage reduction in expected completion energy as the number of sensors K is increased, $M = 2, \alpha = 1, \beta = 0$.

Number of Sensors K

Figure 2-5 considers the more general case when the number of sensors is increased. The number of packets to be uploaded from each sensor is set to $M = 2$ for simplicity. As expected, more gains can be achieved when more nodes are included, especially under asymmetric channel conditions. What we show in this figure is a very limited set of channel conditions, with very few number of nodes in the network. In more general settings, similar evaluations can be completed using the analytical framework we have provided.

Generation Size M

Figure 2-6 compares the reduction in mean completion energy when the generation size, or the number M of packets to be uploaded from each sensor to the base station, is incremented from 2 to 10. Again, packet erasure rate p_2 is set to 0.8, the ratio α between reception and transmission energy per packet is set to 1, and coding energy overhead β is set to 0. The amount of energy gain over the CARQ scheme is higher when M is smaller, with a change of approximately 10% as M increases from 2 to 10. This behavior is owing to the wakeup energy being amortized over more data packets. Furthermore, Figures 2-7 shows the actual values of the expected completion energy for three sets of M. The solid

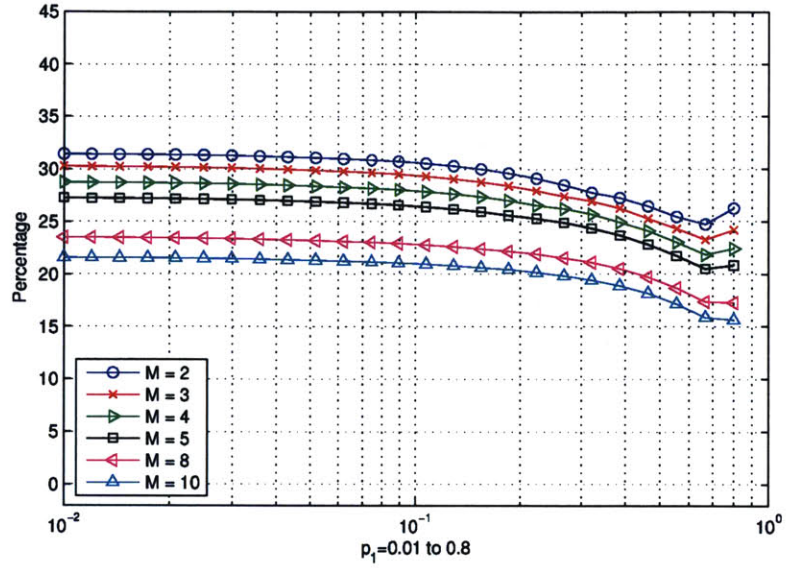


Figure 2-6: Percentage reduction in expected completion energy when M is varied, $K = 2$, $\alpha = 1, \beta = 0, p_2 = 0.8$.

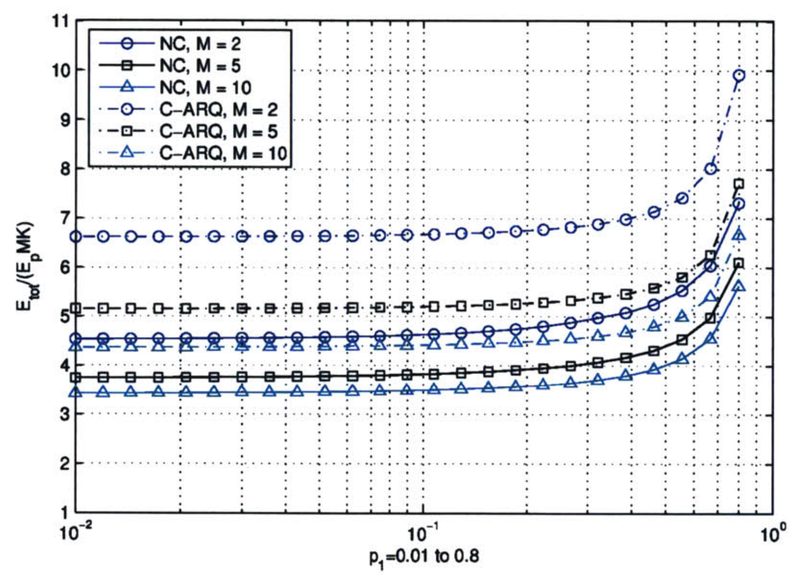


Figure 2-7: Expected completion energy when M is varied, $K = 2$, $\alpha = 1$, $\beta = 0$, $p_2 = 0.8$.

lines correspond to the coded scheme, while dotted lines correspond to the CARQ scheme. Observe that, when coding is performed, the value of M does not have as significant an effect on the overall energy use as in the uncoded CARQ scheme: the curves corresponding to the CARQ case are more spread apart, while those for the coded scheme are relatively clustered together.

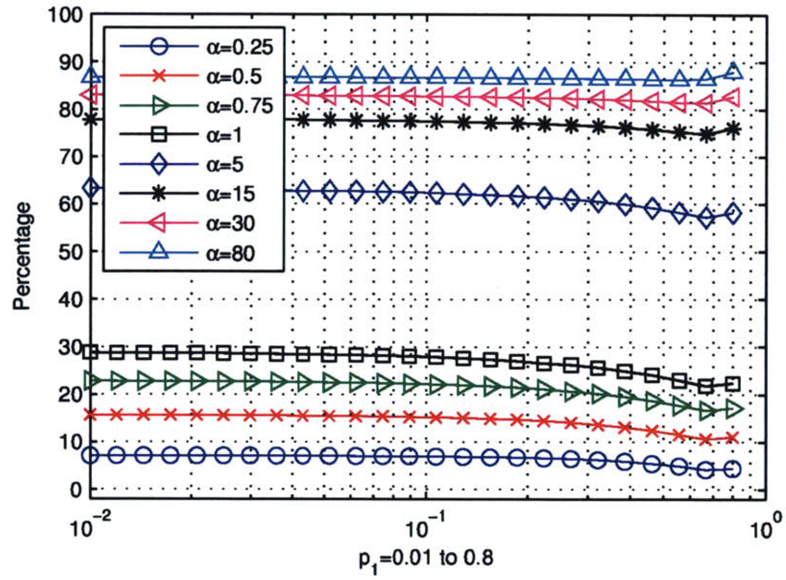


Figure 2-8: Percentage reduction in expected completion energy when α is varied, $M = 4$, $K = 2$, $\beta = 0$, $p_2 = 0.8$.

Transmission vs. Reception

So far we have examined reduction in completion energy achievable through network coding when the reception energy E_a per ack packet is the same as the transmission energy E_p per data packet, i.e., $\alpha = 1$. The actual value of α is dependent on the circuit architectures of the transmitter and receiver, and the data and ack packet payload design. For example, α can be on the order of 1 for a narrow band system, but can be one or two orders of magnitude larger for an ultra-wide band system [19, 65, 71]. Figure 2-8 shows the reduction in completion energy in using network coding with CARQ, over CARQ alone, when α is varied, where $E_a = \alpha E_p$, $E_p = E_{p,NC} = E_{p,CARQ}$ and $E_a = E_{a,NC} = E_{a,CARQ}$. Again, the computation is conducted over a 2-node network for simplicity. When $\alpha = 1$, the observed reduction is approximately 29%. When α decreases from 1, the observed gain is not as significant, for data transmission energy much outweighs ack reception energy. However, as α increases to values higher than 15, we can achieve up to 87% in energy reduction, i.e., 5 times less energy. This is equivalent to extending the lifetime of a sensor node by a factor of 5. Closer examination of the actual optimal transmission schedule shows that when α is large, the network coded scheme compensates by sending a large amount of redundancies a priori, so that the total number of transmission rounds is as small as possible. This is

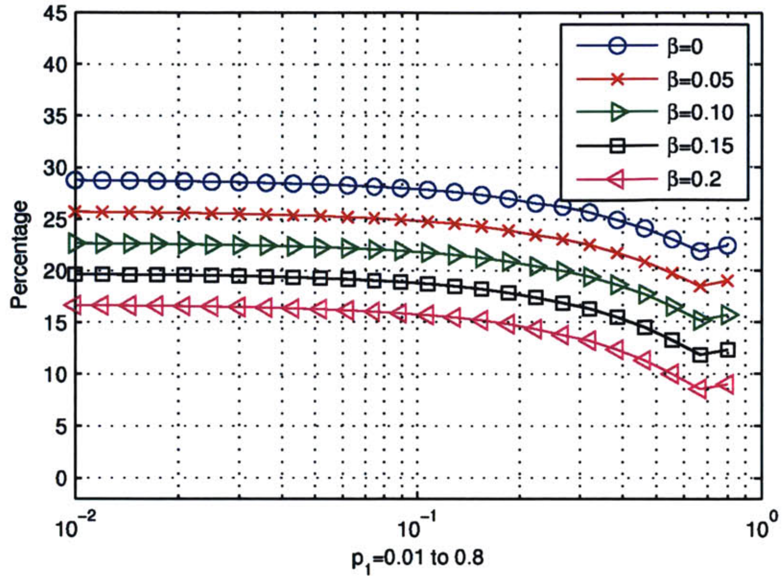


Figure 2-9: Percentage reduction in expected completion energy when β is varied, $M = 4$, $K = 2$, $\alpha = 1$, $p_2 = 0.8$.

similar to rateless transmissions, where only a single acknowledgment is needed at the end to signal successful recipient of all data packets.

Energy Overheads for Network Coding

Another assumption we have made explicitly in previous examples is that the average transmission energy E_p is the same for both boded and uncoded schemes. Indeed, recent studies in the design and implementation of network coding devices show that the energy overhead for network coding can be very small when compared with transmission energy [9]. Nonetheless, depending on the actual implementation of a WBAN, the coded scheme may require non-negligible energy overheads for coding. Figure 2-9 compares the completion energy of the two schemes when $E_{p,NC} = (1+\beta)E_{p,CARQ} = (1+\beta)\alpha E_{a,CARQ}$, $\alpha = 1$, and $E_{a,NC} = E_{a,CARQ}$. Here the energy advantage is lessened because of the added cost of coding. Nonetheless, even with a 20% overhead in coding, we can still achieve an energy reduction of about 17%. Energy overhead associated with coding may be mitigated by using systematic codes. For example, recall from Table 2.1 that when the erasure probability is low, the optimal coded scheme rarely transmit more than the requested numbers of degrees of freedom. Using a systematic code in this case means coding is carried out for only

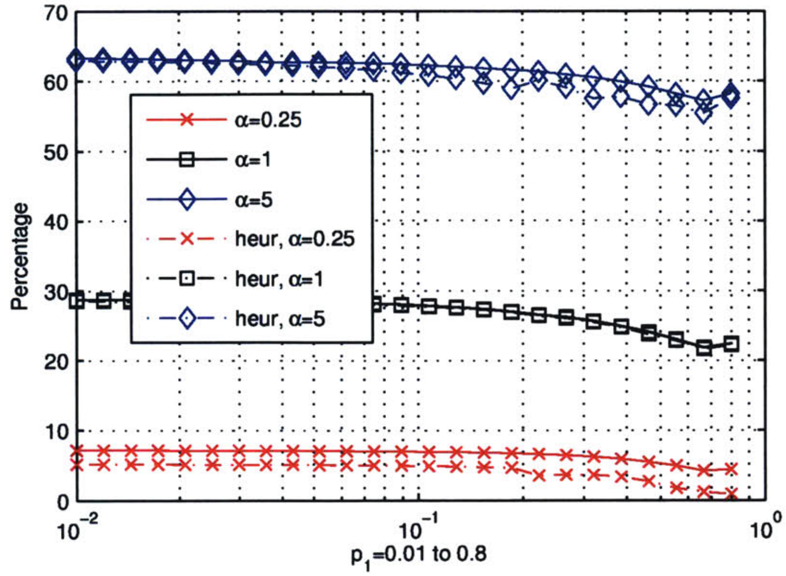


Figure 2-10: Reduction in expected completion energy computed with the given heuristic, when α is varied, $M = 4$, $K = 2$, $\beta = 0$, $p_2 = 0.8$.

a small fraction of the time, thus coding overhead may be negligible.

Heuristic Estimates

Figure 2-10 shows the reduction in completion energy in using network coding over CARQ when α is varied, and when the heuristic given in Eq (2.9) is applied. Again, we consider the two node case, each with $M = 4$ packets to upload. We also ignore the energy overheads for network coding and set β to 0. The packet erasure rate is assumed to be $p_2 = 0.8$ for node N_2. The curves plotted with solid lines refer to the cases when exhaustive search is performed; the curves plotted with dotted lines refer to the cases when the heuristic is applied. In terms of expected completion energy, the proposed heuristic provides a very good estimate of the optimal energy use, especially when $\alpha \geq 1$. Recall that $\alpha = E_a/E_p$ represents the ratio between energies spent on listening to transmission per packet. When $\alpha > 1$, it is more energy efficient to transmit more packets upfront. The heuristic in this case ignores the interdependence among nodes, yet provides a very good approximation to the performance achieved by the optimal transmission scheme found through exhaustive search. A close examination in the small packet number case shows that the exact value of $\tilde{c}_{i_k,k}$ is a reasonable approximation to the optimal scheme.

2.4 Compatibility with the the IEEE 802.15.6 Standard

So far we have assumed very simplistic channel and medium access models to study the behavior of a star-shaped wireless network when linear network coding is incorporated into the system. In realizing the practical potential of this scheme, in this section, we take a closer examination at the IEEE 802.15.6 Wireless Body Area Network standard [6], and propose minor modifications to the standard so that linear random network coding can be included. Since the standard is relatively new, instead of referring the readers to its text for details, we first provide a summary of the physical (PHY) layer and medium access control (MAC) sublayer specifications relevant to the analysis of throughput and energy use. We omit discussions on establishing the initial connection between nodes and the base station, secured or unsecured. Tables and figures given in the description of the standard are reproduced, with limited modifications, from [6].

The IEEE 802.15 Task Group 6 (TG6) was formed in November 2007 to develop "a communication standard optimized for low power devices and operation on, in or around the human body (but not limited to the humans) to serve a variety of applications including medical, consumer electronics/personal entertainment and others" [2]. There have been numerous physical layer, medium access control, and security service proposals. In February 2012, the new 802.15.6 standard was approved by the IEEE-SA Standards Board. This new standard uses the existing industrial, scientific, medical (ISM) bands as well as other frequency bands compliant to regulations by different countries. It supports communication among portable devices under extremely low power, and takes into account changes in channel characteristics due to user motion, as well as specific absorption rates (SAR) of radiation by the body. Each reserved band is further divided into individual channels to allow simultaneously operating WBAN networks. For scalability of the network, multiple medium access mechanisms are offered, including scheduled access, improvised polling, and contention-based random access. Furthermore, to improve power efficiency, macroscopic and microscopic power management schemes are allowed so that devices can enter idle and sleep states, during which the hardware is turned off partially or entirely to preserve energy.

Table 2.3: Physical layer modulation parameters for the 2400MHz to 2483.5MHz frequency ISM band. (SRRC: Square Root Raised Cosine)

Packet Component	Modulation (M)	Symbol rate $1/T_s$ (ksps)	Code rate (k/n)	Spreading factor (S)	Pulse shape	Info. data rate (kbps)
PLCP header	$\pi/2$-DBPSK ($M = 2$)	600	19/31	4	SRRC	91.9
PSDU	$\pi/2$-DBPSK ($M = 2$)	600	51/63	4	SRRC	121.4
PSDU	$\pi/2$-DBPSK ($M = 2$)	600	51/63	2	SRRC	242.9
PSDU	$\pi/2$-DBPSK ($M = 2$)	600	51/63	1	SRRC	485.7
PSDU	$\pi/4$-DQPSK ($M = 4$)	600	51/63	1	SRRC	971.4

In each wireless body area network (WBAN), there are multiple *nodes*, but one and only one *hub*. The hub coordinates medium access and power management among different nodes. We have previously called this coordinating node the base station (BS). The maximum number of nodes allowed in a WBAN is 64. These nodes can be arranged into a star-shaped topology, as we have discussed in this chapter, or into a two-hop extended star topology, where information are exchanged between the hub and a node via another relay-capable node. In the next chapter, we will consider the use of network coding in a relay network.

2.4.1 Overview of the ISM Narrowband Physical Layer Specifications

The IEEE 802.15.6 standard defines three physical layer specifications: Narrowband (NB), Ultra-Wide Band (UWB), and Human Body Communications (HBC). In the narrowband, seven different frequency ranges are considered. A compliant device shall be able to support transmission and reception in at least one of these frequency ranges. Modulation parameters are specified for each individual band. For example, in the ISM radio band in the 2400MHz to 2483.5MHz range, the modulation parameters are as listed in Table 2.3. This frequency range is further divided into 1MHz wide channels, for 79 channels in total. An active WBAN is assigned a single channel, on which all nodes within the BAN and the hub communicate by sharing the medium in the time domain.

The standard physical-layer protocol unit (PPDU) structure is as given by Figure 2-11. Each PPDU is composed of a physical-layer convergence protocol (PLCP) preamble, a PLCP header, and a physical-layer service data unit (PSDU). The PLCP preamble is used

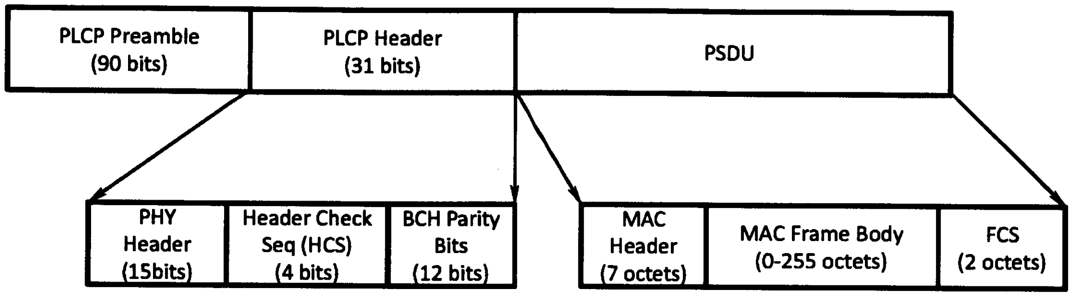


Figure 2-11: Standard PPDU structure.

for packet detection, timing synchronization, and carrier-offset recovery. The PLCP header is encoded using a rate 19/31 systematic BCH code, and contains parameters used in the decoding of the PSDU. The PSDU is encoded with a rate 51/63 systematic BCH code. The total time duration of a PPDU packet is given by the following equation (Eq. (77) from [6]).

$$t_{packet} = T_s \times \left[N_{preamble} + N_{header} \times S_{header} + \frac{N_{total}}{\log_2(M)} \times S_{PSDU} \right] \tag{2.11}$$

Here T_s, S_{header}, M and S_{PSDU} are as given in Table 2.3. $N_{preamble}$ and N_{header} are 90 and 31 bits long respectively as given in Figure 2-11, while N_{total} is defined in terms of N_{PSDU}, the number of bits in the PSDU, N_{CW}, the number of BCH codewords added after systematic encoding, and N_{pad}, the number of bits added to align symbol boundaries for modulation.

Table 2.4 lists several physical layer timing parameters for the narrowband case. The receive-to-transmit (RX-to-TX) turnaround time for a node or a hub shall be between pSIFS and pSIFS + pExtraIFS. For example, when an immediate acknowledgment (I-Ack) is requested for a frame, the I-Ack frame shall be transmitted pSIFS after the end of the received frame. We will discuss different acknowledgment modes in the next subsection. The transmit-to-receive (TX-to-RX) turnaround time for a node or a hub shall be greater than pSIFS. Data can also be transmitted in blocks, if the burst mode field is set in the PHY header. Under the burst mode, uninterrupted successive transmissions are carried out by a device, with interframe spacing between pMIFS and pMIFS + pExtraIFS. Since pMIFS < pSIFS, the burst mode supports a higher throughput. It also allows the transmission of consecutive frames without acknowledgments.

Table 2.4: Narrowband physical layer timing parameters.

PHY parameter	Value
pSIFS	75 μs
pMIFS	20 μs
pExtraIFS	10 μs
pAllocationSlotMin	500 μs
pAllocationSlotResolution	500 μs

2.4.2 Overview of the Medium Access Control Sublayer Specifications

Superframe Structure, Access Modes and Access Phases

A hub may or may not provide time referenced allocations to nodes within its WBAN. In the case where it does support time referenced allocations, the MAC sublayer uses a slotted time reference model as shown on the top of Figure 2-12. The time axis is divided into beacon periods (superframes) of equal lengths, and each beacon period is divided into up to 256 allocation intervals of equal length. A frame transmission may span more than one allocation intervals, with the start and end times not necessarily aligned to the interval boundaries. A beacon frame may be broadcasted by the hub at the beginning of or at a shifted location within each beacon period. Each allocation slot is of length pAllocationSLotMin + $L \times$ pAllocationSlotResolution, where L is the Allocation Slot Length field of a beacon frame, and pAllocationSLotMin, pAllocationSlotResolution, are physical layer dependent parameters as given in Table 2.4.

There are three possible access modes within a WBAN.

1. *Beacon mode with beacon periods (superframes)*: a beacon frame is broadcasted at the beginning of each beacon period (superframe) to specify the length of allocation intervals within the superframe, the start and stop times of the access phases present, and the number of inactive superframes to follow the current active superframe. Four types of access phases can be present in a beacon frame, in the order shown in the beacon frame structure on the bottom of Figure 2-12. The length of each access phase is given in numbers of allocation slots. Exclusive Access Phases (EAP1, EAP2), Random Access Phases (RAP1, RAP2), and Contention Access Phase (CAP) are contention-based random access phases. EAPs are reserved for high priority

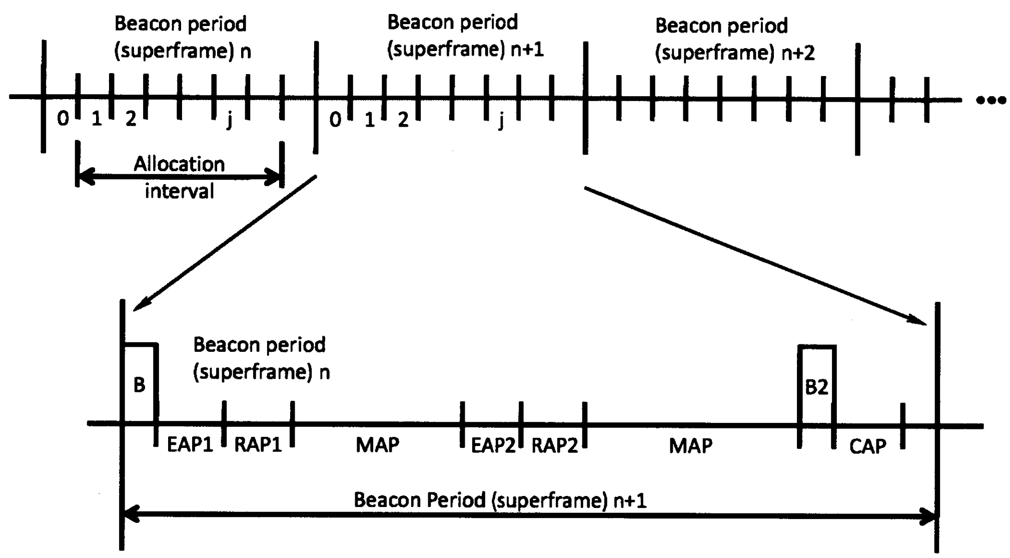


Figure 2-12: Beacon mode with beacon periods (superframe).

transfers. The method for obtaining contended allocations shall be either contention avoidance (CSMA/CA), or slotted Aloha access. In managed access phases (MAPs), allocation intervals are either scheduled, or improvised in the form of polled or posted allocations. A poll is a data request from the hub; a post is an instruction from the hub to a node. In an inactive superframe, no beacon frame is transmitted, and all access phases are absent.

2. *non-beacon mode with superframes*: the superframe contains only a managed access phase (MAP). No beacon frame is transmitted by the hub within the superframe.

3. *non-beacon mode without superframes*: in this mode, the hub may allow unscheduled polled and posted allocations, as well as CSMA/CA based random access to obtain contended allocations.

In beacon and non-beacon modes with superframes, scheduled access may be employed to obtain uplink and downlink allocations, and scheduled-polling may be employed to obtain bilink allocations. Scheduled allocations can be 1-periodic, or m-periodic. In 1-periodic allocation, a node is active in every beacon period, with one or more allocation intervals spanning the same allocation slots in every beacon period. In m-period allocation, a node is active every m beacon periods, where $m > 1$. The same slot or slots are allocated to a node in each active beacon period. A node shall not have both 1-periodic and m-periodic allocations in the same WBAN. To obtain one or more new scheduled allocations, a node sends a Connection Request frame to the hub, which in return sends a Connection Assignment frame to the node to grant access when possible. Existing scheduled allocations can be modified or terminated by the transmission of a new Connection Request frame. A scheduled allocation may also be aborted if the expected receiver does not receive any frames successfully within a pre-defined time period.

MAC Frame Format, Frame Types and Subtypes

The format of a MAC Service Data Unit (MSDU) is given in Figure 2-13. A MAC frame is consisted of a header, a variable-length frame body, and a 16-bit cyclic redundancy check sequence. The MAC frame body has a maximum possible length of 255 octets. A zero length indicates that no frame body is included. The MAC header in turn contains the sender and receiver IDs, as well as the frame control sequence.

Each WBAN is identified with an unique one-octet abbreviated BAN identifier (BAN ID). The hub also chooses a one-octet hub identifier (HID) to be included in the MAC headers of all frames sent to or from the hub. This HID should not overlap with the identifier of any node connected to it, or any neighboring hubs. Similarly, each node in the BAN is identified by a one-octet node ID (NID), the value of which depends on whether the node has established a confirmed connection with the hub of the BAN, and whether the node is sending to (unicast) or receiving from (multicast or broadcast) from the hub. Additionally, a six-octet extended unique identifier (EIU-48) can provide full addressing of the hub and nodes within the WBAN.

The contents of a MAC frame differ among different frame types and subtypes. In the MAC frame header, a Frame Type field can take on four different values, representing management, data, control and reserved frames. A management type frame can be further divided into beacon, security-related, connection-related, and reserved frame subtypes. In beacon mode with beacon periods (superframes), a beacon frame is multicasted by the hub at the beginning of every beacon period (superframe). Its frame payload contains the length of the current beacon period, and timing allocation information, such as the time duration

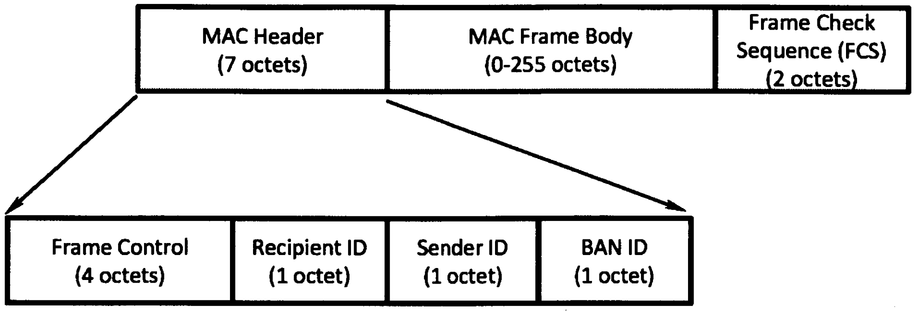


Figure 2-13: MAC frame format.

of each allocation slot and the start and stop times of each access interval. Data type frames are divided into subtypes depending on user priority levels and data subtypes. A control type frame is used by a node or a hub to acknowledge the recipient of a management type or a data type frame. A node or hub receiving a control type frame does not generate any further acknowledgments. In each transmitted management type or data type frame, an Ack Policy field in the MAC header indicates the type of acknowledgement requested by the current frame.

Modes of Frame Acknowledgments

The Ack Policy field of the MAC header indicates the type of acknowledgments requested by the current frame. It can take on different values, representing the four policies below:

1. *No acknowledgment (N-Ack)*: an N-Ack indicates that the current frame does not require an acknowledgment from the recipient. A frame with an N-Ack request can be transmitted by either the hub or a node. Control type frames, which carry acknowledgment information for management and data type frames, always have their Ack Policy fields set to N-Ack.

2. *Group acknowledgment (G-Ack)*: group acknowledgments are applicable to data type frames sent to a *hub*. A G-Ack is requested if the frame is of *data* type, its Ack Policy field is set to N-Ack, and its frame subtype field is set to the constant mG-AckDataSubtype. The hub acknowledges frames with G-Ack requests from multiple nodes together through the use of a G-Ack multicast, which is given in

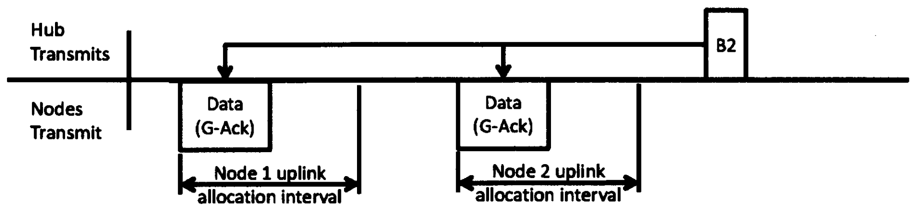


Figure 2-14: Group Acknowledgement (G-Ack) example.

the form of a frame subtype called B2 frames. A B2 frame contains a set of NIDs, indicating the nodes from which the hub has received a frame with G-Ack requests since the last transmitted B2 frame. A node may retry the frame with the G-Ack request if it fails to receive the expected B2 frame, or if its NID does not appear in a received B2 frame. Figure 2-14 gives an example of group acknowledgments.

3. *Immediate acknowledgment (I-Ack)*: a frame with an I-Ack request can be transmitted by either the hub or a node. The frame is acknowledged by an I-Ack frame immediately upon its reception. The recipient sends back an I-Ack frame pSIFS ($75\mu s$ in the narrowband case) after the end of the frame reception. Management type frames are always transmitted with an I-Ack request. Data type frames can be transmitted with a request for any one of the four acknowledgement modes.

4. *Block acknowledgment later (L-Ack) and Block acknowledgment (B-Ack)*: a frame with an L-Ack or a B-Ack request can be transmitted by either the hub or a node. The transmitted frame must contain a whole MSDU. Frames with L-Ack requests are not followed by any immediate acknowledgements. Instead, blocks of frames with L-Acks are acknowledged by a B-Ack frame pSISF after the reception of the last frame with a B-Ack request. Figure 2-15 illustrates the block transmission scheme established through the use of L-Acks and B-Acks. During block transmission, the source shall send frames in the order of non-decreasing sequence numbers, with repetitions if retransmissions take place. For example, Figure 2-15 provides two instances of retries. All frames in a block transmission should be of the same type and subtype. A block transmission may span more than one allocation intervals, as illustrated by Figure 2-15.

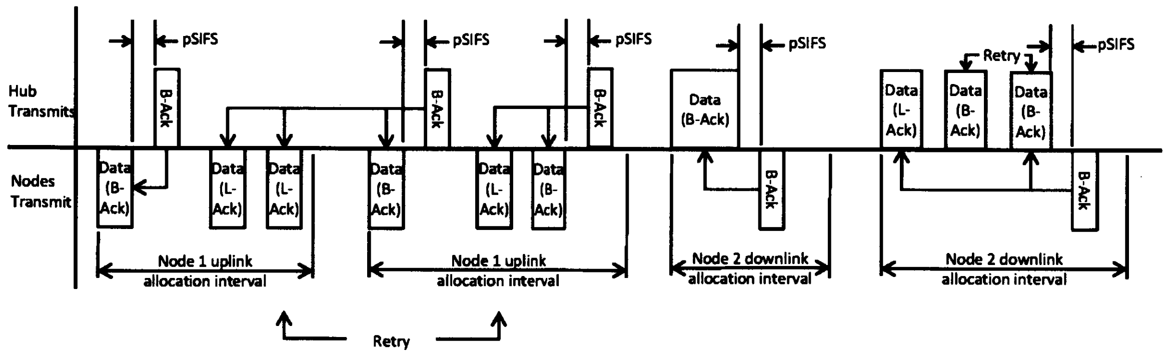


Figure 2-15: Block Acknowledgment Later (L-Ack) and Block Acknowledgment (B-Ack) example.

Three fields of a B-Ack frame function together to provide acknowledgments to a block of data: Oldest Frame Expected, Next Block Size, and Frame Status Bitmap. The Oldest Frame Expected field is not present if there are only one frame in the current block. When present, it is set to the sequence number of the oldest frame that is expected but has not been received. If all frames in the block has been received, the Oldest Frame Field is set to 1 plus the sequence number of the frame preceding the B-Ack frame. In other words, the Oldest Frame Expected field indicates the next expected frame, assuming all frames are transmitted in order. This is similar to the ack signals sent in the Transmission Control Protocol (TCP) for internet applications. The Next Block Size field is not present if only one new frame is expected and allowed. When present, it is set to the maximum number of frames that can be included in the next block transmission. The Frame Status Bitmap field is one byte long, and indicates the reception status of each successive frame in the block transmission, starting from the one immediately next to the oldest expected frame in sequence number. Successive frames are frames with successive sequence numbers. Hence, acknowledgments are packet specific; any un-acknowledged packet is retransmitted during the next block transmission. Moreover, since frames in a block transmission are required to contain whole MSDUs, the recipient may implement a timeout mechanism so that MSDUs can be released to the upper layers without the reception of all frames in the block.

Other than I-Ack, B-Ack, and B2 frames, I-Ack+Poll and B-Ack+Poll control type frames can also be transmitted by the hub to grant immediate polled allocations or to announce a future poll or post request, in addition to frame acknowledgments. Moreover, a node or a hub shall treat an expected frame, such as an I-Ack or B-Ack frame, as not arriving, after waiting for the PHY preamble of the frame for mTimeOut amount of time ($30\mu s$).

Power Management

To preserve energy, a node is allowed to be inactive over some time intervals. In particular, two power management schemes are permitted

1. *Hibernation – macroscopic power management*: by varying the Wakeup Period field in its Connection Request frame, a node may request an integer number of beacon frames during which it is to be inactive. In m-periodic scheduled access, a node is to remain in the hibernation state for $m - 1$ beacon periods in between active beacon periods.

2. *Sleep – microscopic power management*: in a beacon period during which a node is to be active, the node can return to the inactive state when it does not need to transmit to or to receive from the hub. In beacon modes with superframes, a node may be inactive during the beacon transmission time, if it does not need to receive the beacon frame. If necessary, it can become active again after its own allocation slots have passed to receive a B2 frame or to transmit in the CAP period.

2.4.3 Incorporation of Network Codes

A quick comparison between the acknowledgment modes offered by the standard and our description of the network coding based scheme given earlier in this chapter shows that network coding can be easily incorporated into the standard. In this subsection, we provide a brief description of how this can be carried out. The quantitative advantages of allowing network coded transmissions require detailed analysis of actual implementations of the standard. We do not attempt to characterize the performance of such systems here.

Generation-Based Approach

Similar to discussions provided in earlier parts of this chapter, network coding can be carried out at individual nodes using a generation-based approach [16].

First, the block acknowledgment mode with L-Acks/B-Acks can be used with coded frames transmitted by a node or the hub. At the transmitter side, the MAC frame payload of a given number of packets, to be transmitted together in a block, can be coded together before being transmitted. Each block is therefore a generation. If random linear codes [33] are used, the coding coefficients can be attached as part of the payload. If deterministic codes are used, the coding coefficients can be pre-defined. In the MAC frame control header of the transmitted frame, a reserved bit can be used to indicate whether the current packet is coded or not. With coding, redundant packets can be added to compensate for potential losses. As we have discussed in Section 2.3, a reasonable level of redundancy can be computed using the heuristic given in Eq. (2.9). In practice, the packet loss rate should be estimated and the redundancy level should be adjusted accordingly. A systematic network code can be used to minimize energy use for coding. All coded frames except the last one carry an L-Ack request to the receiver, while the last coded frame carries a B-Ack request to the receiver.

At the receiver side, assuming well chosen coding coefficients, each arriving packet is a unique degree of freedom, being innovative to the receiver with high probability. Whenever a B-Ack request is received together with a coded frame, the receiver can use the one-octet-long Frame Status Bitmap field to acknowledge to the transmitter the number of degrees of freedom it has received, or the number of degrees of freedom it is still expecting (computed using the Next Block Size field). In addition, the Oldest Frame Expected field can be modified to indicate the generation number. Recall that with uncoded transmissions, only 9 frames can be acknowledged in each block (indicated by the Oldest Frame Expected field, plus 8 bits in the Frame Status Bitmap). With coded data, more than 9 frames can be transmitted in a single block, or generation.

Second, the block and group acknowledgment modes can be combined into a Group-Block acknowledgment (GB-Ack) mode, similar to the combined ARQ scheme discussed

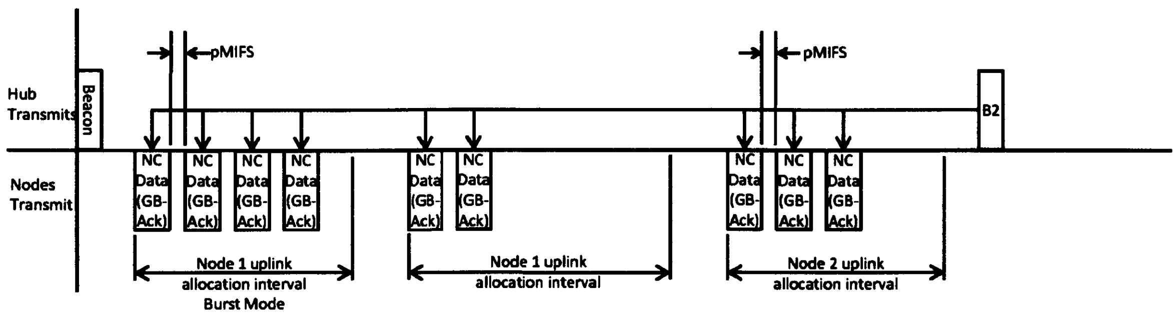


Figure 2-16: Example of coded data blocks with Group-Block Acknowledgments (GB-Ack).

earlier in this chapter. Figure 2-16 shows one possible such use of GB-Acks in scheduled access mode, for data blocks uploaded from all nodes to the hub. At each node to be included in group-block acknowledgment, data frames in a block can be coded with random or deterministic coefficients. In the MAC frame control header of the transmitted frame, a reserved bit can be used to indicate if the current packet is coded or not. At the hub, recall that a B2 frame is broadcasted at a pre-determined allocation slot when G-Acks are requests by multiple nodes. A B2 frame contains a set of NIDs, each representing a node to be group acknowledged. Since there can be a maximum of 64 nodes within a WBAN, each with a single-octet NID, but the B2 frame payload can be of 255 octets long, it is possible to append to the current B2 frame payload the numbers of degrees of freedom successfully received from each node, for its respective current generation. For example, a one-octet field can be added to the B2 frame payload for each node from which the hub has received coded frames in the current block. The first 4 bits can represent the generation number, while the last 4 bits can represent the number of degrees of freedom received from the node during the current block.

There are several possible advantages of using coded transmissions with block acknowledgemnts or group-block acknowledgments (GB-Ack) instead of sending uncoded payloads with I-Ack, G-Acks or L-Ack/B-Acks.

1. As we have discussed previously in this chapter, independently coded packets are individual degrees of freedoms. Block acknowledgment of coded packets does not need to specify which individual packet has been received successfully. With coded

payloads, more than 9 frames can be included in a block transmission.

2. From a throughput perspective, in cases where an acknowledgment is lost and retransmission occurs without the loss of the actual data frame, the retransmitted data frame is wasteful in the uncoded case. On the other hand, if the retransmitted frame is coded with a set of coefficients linearly independent from the previous frame, the retransmission is innovative, carrying one more degree of freedom to the hub.

   In addition, in burst mode block transmissions, interframe spacing between successive frames is smaller than the immediate acknowledgement case. Thus with less frequent transmissions, the overall throughput of individual nodes can be higher. As a simple example, let us compare the number of frames that can be transmitted in an allocation interval with different acknowledgement modes, when a data frame is transmitted with an medium-sized payload, at a fast information data rate of 485.7kbps. Figure 2-17 shows an example timing diagram of I-Ack, B-Ack (coded or uncoded) and GB-Ack schemes. First, we compute the time duration for transmitting a PHY frame using Eq. (2.11). As shown in Figure 2-13, the MAC frame body can be up to 255 octets long. Assume we have a data frame with a payload of 100 octets. Taking into account the MAC header and the FCS, number N_{PSDU} of bits in a PSDU is 872. After encoding with a rate 51/63 systematic BCH code and adding appropriate paddings to align symbol boundaries, the total number of bits in the PSDU is $N_{total} = N_{PSDU} + N_{CW} \times (n-k) + N_{pad} = 872 + \lceil \frac{872}{51} \rceil (63-51) + 0 = 1088$bits (Eq. (68) from [6]). Thus

$$t_{\text{data packet}} = \frac{1}{600 \times 10^3} \left[90 + 31 \times 4 + \frac{1088}{\log_2(2)} \times 1 \right] = 2.17ms\,. \qquad (2.12)$$

   On the other hand, the frame payload for an I-Ack frame is 3 octets long, while the frame payload for a B-Ack frame is 6 octets long. Assuming acknowledgements are sent at the same rate as data, the time duration of I-Ack and B-Ack frames are $t_{\text{I-Ack}} = 0.567ms$ and $t_{\text{B-Ack}} = 0.617ms$. The duration of an allocation interval is

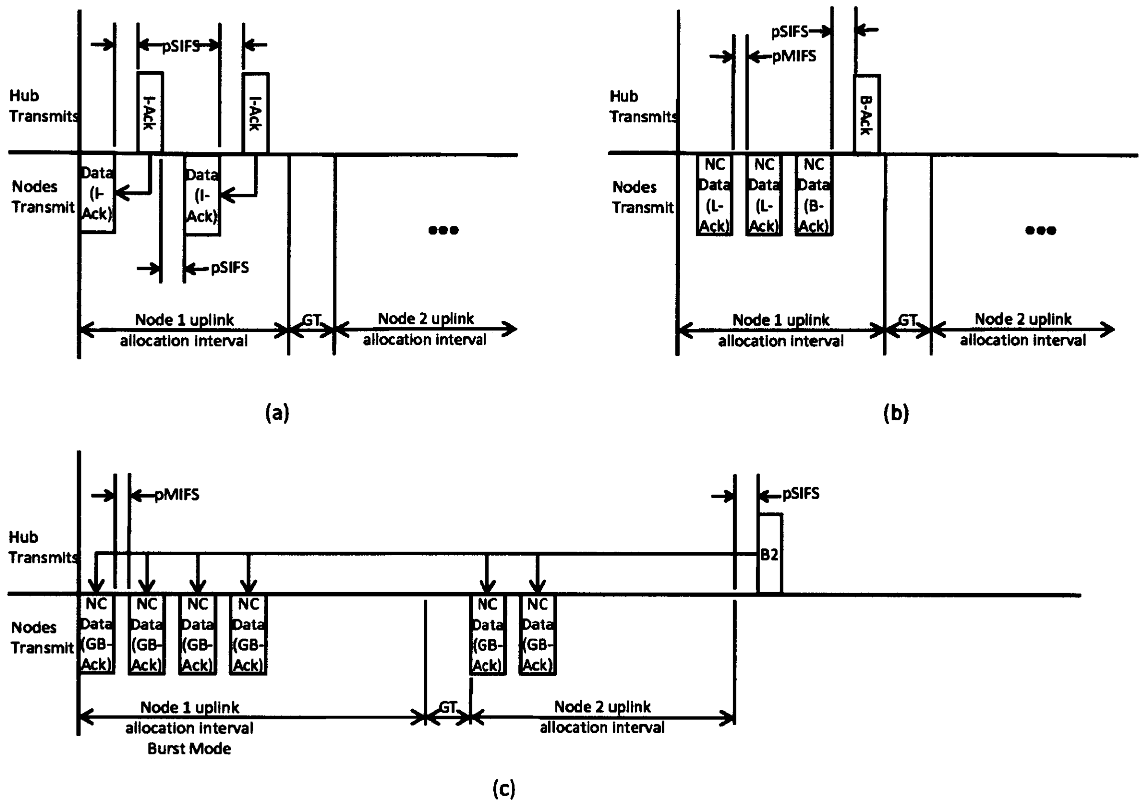


Figure 2-17: Timing example (a) uncoded transmission with I-Acks (b) coded data transmission with B-Acks and L-Acks (c) coded data transmission with GB-Acks.

given by pAllocationSlotMin + $L \times$ pAllocationSlotResolution, where $0 \leq L \leq 255$. Let us assume nodes are operating in the non-beacon mode with superframes. The entire superframe can therefore be used for scheduled allocation. Suppose there are two nodes connected to the hub, each with $L = 100$, or an allocation interval of duration $50.5ms$. We assume that guard intervals are centrally added in between allocation intervals instead of distributively included in each allocation interval. With I-Acks, each node can transmit $\lfloor 50.5ms/(t_{\text{data packet}} + 2 \times mSIFS + t_{\text{I-Ack}}) \rfloor = 17$ frames. With blocked transmissions in burst mode, however, each node can transmit $\lfloor (50.5ms - (pSIFS - pMIFS) - t_{\text{B-Ack}})/(t_{\text{data packet}} + mMIFS) \rfloor = 22$ frames. Note that with coded payloads, all 22 frames can be acknowledged with a single B-Ack frame. With GB-Acks, assuming that the B2 frame is transmitted outside of the allocation intervals of each node, the number of frames that can be sent in the allocation interval is $\lfloor (50.5ms)/(t_{\text{data packet}} + mMIFS) \rfloor = 23$ frames. Note that, if the allocation interval is shorter, with high probability, the same number of frames can fit into the interval regardless if B-Ack or GB-Ack is used. Nonetheless, it is not hard to see that if the number of nodes connected to the hub is large, using GB-Ack is beneficial in terms of reducing the total number of acknowledgment frames transmitted in one superframe.

3. From an energy perspective, given a fixed number of data frames to transmit by a node, block transmission requires the node to be active for a shorter period of time than when I-Acks are used. Such energy savings become more significant when switching between transmission and reception modes take up a non-negligible amount of power. Moreover, with the coded scheme, it is always possible to transmit more redundancies when listening to ack packets costs more energy. Our analysis in this chapter provides a framework for investigating the energy tradeoffs between transmission and reception in such scenarios. Also keep in mind that the operation of encoding requires additional energy, and a tradeoff exists between encoding energy and transmission/reception energy.

4. From a reliability perspective, using coded transmission allow redundant coded frames to be transmitted a priori to compensate for possible losses. Network coding thus provide another reliability mechanism, allowing less stringent requirements on the physical layer design. In addition, in cases where channel quality is low and scheduled allocations for a node is aborted because none of the transmitted frames is received successfully at the hub, such provision for losses may prolong the duration of the connection between a node to the hub.

Sliding Window Approach

An alternative to the generation-based network coding scheme is to use a sliding window approach, as proposed by by J. K. Sundararajan [76] for interfacing network coding with TCP/IP. In this case, at the transmitter, data frame payloads are put into a coding window of a fixed size as they arrive from the upper layer, as long as the coding window is not full. A frame is removed from the coding window when it is acknowledged by the receiver. At the sink, coding coefficients from received packets are put into a matrix and passed through Gaussian elimination into a reduced row echelon form. Frames on the pivot columns shall be taken as *seen*, and shall be acknowledged in the next available acknowledgment frame. In other words, a receiver is said to have seen a packet $\mathbf{p}_k$ if it can compute a linear combination of the form $\mathbf{p}_k + \sum_{l>k} \alpha_l \mathbf{p}_l$, where α_l are coding coefficients, and k represents frame sequence number. It can be shown that if all frames involved in the coding operations have been seen, then all coded packets can be decoded. Seen packets are therefore dropped from the coding window at the transmitter side. The notion of seen packets defines an ordering for the degrees of freedom conveyed by coded packets: for every additional degree of freedom received at the hub, the next data frame in the coding window is acknowledged and dropped from subsequent coding operations.

The sliding window approach can be used in a WBAN similar to the generation-based approach. At the transmitter, data frame payloads are coded, and sent with redundancies. At the receiver, either block or group-block acknowledgments can be used. The Oldest Expected Frame field of a B-Ack frame can be modified to represent the oldest expected seen frame. Frames which have been seen are subsequently removed from the coding window at the transmitter side. For GB-Ack frames, an oldest expected seen frame field can be added to indicate which frames are to be removed from the coding window of each transmitting node. In addition, a systematic code can be used with the sliding-window approach, and the transmission scheme can operate in a rateless fashion if needed.

An advantage of the sliding window approach is that it allows new data to be added to the transmission queue incrementally as they arrive at the MAC sublayer and removed as soon as they are seen at the receiver. It does not require a pre-defined generation size, although the size of the coding window can vary, depending on the amount of computation and memory available at sensor nodes. When data rate is not very high, fast removal of data packets from the coding window is beneficial, since fewer coding operations are then performed for subsequent transmission rounds. Another advantage of the sliding window approach is that its retransmission requests are very short, since it only needs to acknowledge the last seen packet, or the last seen degree of freedom.

A disadvantage of the sliding window approach is that even though memory and energy are assumed to be abundant at the base station, decoding delay is larger. Seen packets may have to stay in the receiver queue for longer periods than the generation-based case, before they can be decoded. This problem can be mitigated if the transmitter flushes its coding window periodically.

2.5 Challenges in System Implementation and Evaluation with SDRs

The theoretical analysis and numerical simulations presented earlier in this chapter characterize network performance bounds that can provide good guidance for practical implementations of a network coded WBAN system. However, the proposed model is only a simplified abstraction of a real-world wireless network, thus its accuracy cannot be validated without implementation and evaluation on actual wireless radio testbeds. Such testbeds can be constructed with off-the-shelf commercial transceivers and a custom-made VLSI network coding accelerator [9], but they generally do not provide MAC layer protocol implementations. Software-defined radios (SDR), on the other hand, allow access to signal processing and control algorithms within the protocol stack, thus seem to be a good choice for developing and testing our wireless networking protocol. In this section, we provide a very short description of a SDR-based platform, and discuss challenges we encountered in our attempt to implement and evaluate the proposed coded transmission scheme. Although we do not provide any experimental results, we believe the issues we faced would be common for any similar implementations using SDRs.

A software-defined radio [66, 85], as its name implies, is a radio implemented in software, with most of its signal processing functions performed in a general purpose computer. A typical SDR involves a dedicated external radio frontend (RF), and collections of software signal processing blocks. The RF contains appropriate digital to analog (DAC) and analog to digital converters (ADC). It can also contain down converters and up converters so that baseband digital data can be communicated with a computer system through a USB or ethernet connection. The advantages of an SDR lie in its flexibility and versatility.

It can provide quick development cycles to researchers and engineers who are interested in coding, signal processing, protocol, and radio architecture design, offering easy reconfigurability to the system without needing to modify a dedicated hardware design. It can provide a compact and low-cost platform for the co-existence of various wireless communication protocols, enabling devices to be universal and easily upgradable. It can also lead to more efficient spectrum use by allowing radio devices to be cognitive, automatically adapting to the environment.

Given its flexibility and low cost, SDR-based testbeds have been used in the verification of novel network coding schemes [30,38,40]. Our choice of using a SDR is partly based on its easy configurability, and partly because we have easy access to the Universal Software Radio Peripheral (USRP) platforms, a low-cost RF hardware designed by Ettus Research [5]. On the software front, we use GNURadio [70], an open source software development toolkit that provides signal processing blocks for SDR implementations. In GNURadio, performance-critical signal processing blocks are implemented in C++, then glued together and accessed through the Python programming language [57]. Python provides packages and modules to connect blocks together, establishes the signal path, schedules and controls signal flow, and provide tools for generating graphical user interfaces.

We encountered two major issues in our SDR implementation. One is timing synchronization, the other is channel estimation. First, as signal processing is performed in a general purpose computer where delays are added with multithread handling, timing synchronization is difficult to achieve across multiple units, especially when the time-division multiple access scheme is employed. External hardware clocks can be engaged, but data packets can be lost easily if the receiving unit does not have the receiver flow graph running when packets arrive from the transmitting unit. Instead of trying to achieve synchronization in our setup, we can allow very large guard intervals around each time period during which the wireless medium is accessed by a node or the base station, thus ensuring whenever transmission occurs, the receiving end is already in listening mode. Moreover, with the inserted network encoder and decoder, additional packet queues need to be managed to handle decoding delays.

A second and more significant challenge in implementing the proposed network coded transmission scheme is to estimate the packet erasure rates on channels from individual nodes to the base station. In the analytical model, we assume such packet loss rates are static over the entire transmission process, and known at the base station. However, depending on radio signal propagation effects around the human body and user body movements, actual packet erasure rates may need to be dynamically estimated frequently. In general wireless systems, packet loss rates are often derived from PHY layer achievable bit error rates. Since sensor nodes in WBANs are energy-limited, repeated transmissions of pre-determined signals for channel estimation are not desirable. In addition, quality of the channel from a node to the base station can only be estimated at the base station, but with the heuristic proposed in Section 2.3, this packet loss rate or the corresponding computed heuristic value needs to be communicated to the sensor node itself. It is not immediately clear how channel state can change during the time this information is communicated by the base station to individual sensor nodes.

2.6 Conclusions

In this chapter, we propose a network coded scheme to help improve energy efficiency of wireless body area networks. Assuming that the different channel conditions experienced by individual nodes are known at the base station, the base station can request each sensor node to send an optimal number of coded packets, taking into account anticipated packet losses during transmission, and energy needed for receiving control signals. We establish a Markov chain model to analyze the evolution of innovative packets at different nodes during the transmission process. Even though we do not provide a closed-form solution to the combinatorial optimization problem thus formulated, we show with numerical examples that in a two-node star, when transmitting a data packet and receiving an ack packet cost approximately the same amount of energy, the network coded scheme can achieve up to 29% percent reduction in expected completion energy per accepted data packet compared to the uncoded scheme. When receiving costs a lot more than transmitting, network coding can reduce energy use by up to 87%. We also show with numerical examples that the amount of energy gain achievable through coding increases as more nodes are added to the network, and when nodes see more asymmetric channel conditions.

Moreover, we provide a simple heuristic for computing the optimal number of packets to transmit. This heuristic determines the transmission schedule for each node separately. It provides a good approximation to the optimal solution found through exhaustive search in terms of expected completion energy, although it does not take into account interdependencies among sensor nodes through wakeups to listen to acknowledgment and scheduling information.

In summary, using a network coded scheme in a WBAN has three advantages over uncoded schemes. First, retransmission requests are simpler since they do not need to be packet specific. Second, redundant packets can be transmitted a priori, in expectation of packet losses. Such redundancies can be seen as a reliability mechanism in addition to forward error correction schemes in the physical layer. Third, allowing more coded packets to be sent in each round also allows the overall transmission process to terminate in fewer rounds. The total number of turnarounds between transmission and receptions modes, and the total number of times nodes wake up to listen to ack signals are thus reduced, leading to savings in energy use. This chapter provides an analytical framework for studying the energy performance of this coded scheme.

In this chapter, we also provide a brief description of the IEEE 802.15.6 WBAN standard, and consider the incorporation of network coded transmission schemes to the standard. Both generation-based and sliding window approaches are possible, and a group-block acknowledgment scheme can be easily implemented by modifying block acknowledgment control type frames. Since the standard only provides general rules and guidelines, actual implementations of the standard can differ in terms of hardware and signal processing algorithm design. Whether coding helps in terms of energy or throughput thus depends on how the standard is realized. In our study of the star network, we have tried to implement the proposed coded scheme with a software-defined radio (SDR) platform based on USRP and GNURadio. Timing synchronization and channel estimation are two problems that need to be resolved in this implementation process.

Chapter 3

Packet Erasure Relay Network

This chapter studies the strategic use of network coding in a three-node wireless packet erasure relay channel, with an emphasis on whether and where to code when the relay operates in half-duplex mode. We propose Markov chain models to characterize the system performance in terms of throughput and packet delivery energy, thus providing a way to find the optimal fraction of time for which the relay should participate in the transmission, and where linear network coding should be performed. We show, through numerical analysis, that when transmission, reception, and coding takes the same amount of energy per packet, coding at the relay alone while operating in a rateless fashion is neither throughput nor energy efficient, while coding at the source alone has performances close to the case where coding is performed at both nodes. The decision to code is also dependent on packet erasure probabilities. Although our numerical results are evaluated under one set of energy parameters, the analytical framework we provide in this chapter can be used in more general settings, where one or more energy terms can dominate, to determine the appropriate coded relaying strategy.

Although seemingly simple, the analysis of the three-node network can offer insights to more complicated systems with more source or sink nodes. One example is the WBAN system, another is advanced LTE cellular networks [3]. In a WBAN, the topology is almost always star-shaped: data are uploaded in a converge-cast sense to a central base station. Depending on the relative location of a sensor on the human body, it may be useful in terms of energy efficiency to deploy a relay around the shoulder, in direct line of sight with both the front and back of the body. To the best of our knowledge, the throughput and energy tradeoffs in this case have not been studied before. On the other hand, in advanced LTE systems, data are transmitted in both directions, with an emphasis on download in a broadcast sense from a central base station to individual user equipments. Currently relaying is being considered as an improvement tool, for example, for coverage of high data rates, and for temporary network deployment [3]. Here the relay is to wirelessly connect to the radio-access network, and may either function as a smart repeater, or have control of its own cell. For both WBANs and advanced LTE systems, the introduction of network coding and the insertion of a relay may bring energy or throughput gains. As a starting point, we consider the three-node packet erasure relay channel, assuming physical layer designs are readily available on point-to-point links. Further extensions of this setup can involve additional source nodes, as in a WBAN, additional sink nodes, as in an LTE system, and additional relay nodes, a scenario applicable to both examples.

The remaining of this chapter is organized as follows. Section 3.1 summarizes some of the previous work related to the wireless relay channel and the application of network coding in such settings. Differently from previous works that focus on joint channel and network coding for optimal throughput analysis, in this chapter, we assume a physical layer design is available, and network coding is inserted into the network layer, independently of the source and channel codes employed. Our goal is to determine where linear network coding should be performed, and whether using coding is beneficial. Section 3.2 details the assumptions made and introduces necessary definitions and notations. Section 3.3 then discusses three separate cases, depending on the coding locations. For each case, we try to characterize the expected completion time and the expected completion energy of transmitting a given number of packets from the source to the destination. Through Markov chain analysis, we provide a framework for evaluating bounds on the system performance when coding is conducted at the relay only, and for determining the system performance when coding is conducted at the source only. We also use a flow model to determine system performance when coding is performed at both the source and the relay. Section 3.4 then evaluates throughput and energy performances numerically. Section 3.5 concludes the chapter.

3.1 Background and Related Work

A relay channel models the problem where two nodes communicate through the help of one or more relays. This setup is common in multihop wireless networks such as sensor networks, where transmission power is limited, or in decentralized ad hoc networks, where nodes can communicate only with their immediate neighbors. Relays can overhear transmissions to the destination, owing to the broadcast advantage of the wireless medium. Recently there has been a renewed interest in the classical relay channel [18, 87], motivated by the potential to achieve cooperative diversity, and thus better capacity bounds [20,45,47,48]. Schemes such as amplify-and-forward, decode-and-forward, and compress-and-forward have been proposed and studied extensively in terms of capacity, outage, energy efficiency, and optimal power allocation schemes [90, 97]. Much of the analysis has focused on the fundamental performance limits at the physical layer and on the transmission of a single data packet. The introduction of network coding into the relay channel has also focused mostly on joint channel-network code design, with or without limited processing complexity constraints, to achieve cooperative diversity and capacity [32,81,82,88,89,94]. In larger networks, however, network coding typically resides in higher layers of the protocol stack, independent of physical layer implementations.

In this chapter, we assume a physical layer design is already available and network coding is inserted into the network layer, independent of source and channel codes employed. Such an assumption on the separation of channel and network codes may not necessarily be capacity achieving, but allows the introduction of network coding into existing systems.

The use of random linear network codes (RLNC) in wireless erasure networks under packetized operations is first studied by Lun *et al.* [56, 61], and extended to a scheduling framework by Traskov *et al.* [80]. Other schemes that employ network coding in a relay setup includes the MORE protocol [13], which performs RLNC at the source only to reduce the amount of coordination required by multiple relay nodes, and the COPE protocol, which employs RLNC at the relay only, in a 2-way relay channel to improve reliability, taking advantage of opportunistic listening and coding [40]. Fan *et al.* also proposed a network coding based cooperative multicast scheme to show that significant throughput gains can

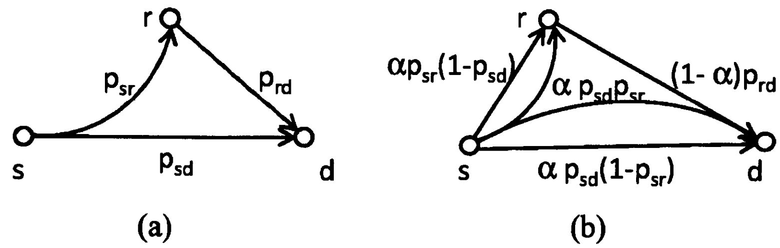


Figure 3-1: Single relay unicast network, with corresponding flow hypergraph. p_{sr}, p_{rd}, and p_{sd} represent the packet transmission success probabilities between s and r, r and d, and s and d respectively. The source s transmits for α fraction of the total time, while the relay r transmits during the remaining $1-\alpha$ fraction of the total time, $0 \leq \alpha \leq 1$.

be achieved when network coding is performed at the relay only [25]; one assumption in this work is that feedback is available from both the destination and the relay to the source after each packet reception. In practical systems, feedback can be costly in terms of both throughput and energy, depending on the underlying hardware architecture [72]. In this chapter, we explore rateless transmissions, where the acknowledgment for successful reception is sent only once by the destination when the transmission of all available data is completed. Similar to the star network case, we also take into account the energy spent on reception and packet processing in addition to the energy spent on transmission.

3.2 System Model and Problem Formulation

We represent the data flow through the relay channel using a hypergraph, as shown in Figure 3-1(a). A hypergraph is a generalization of a graph: a broadcast link is represented by a hyperarc between a single start node and a set of end nodes, and a multiple access link is represented by a hyperarc between a set of start nodes and a single end node [56]. A wireless relay channel consists of a source node s, a relay node r, and a sink node d. Source s has n packets of the same length to transmit to d. It broadcasts to both r and d. The relay r assists the transmission by either forwarding the original packet, or computing linear combinations of received packets before forwarding the ensuing mixtures.

We assume transmissions occur in a rateless fashion, with minimalistic feedback: s and r take turns to transmit, until d acknowledges that it has received enough degrees of freedom to recover the original n data packets. Such rateless operations are often desirable in systems where feedback can be costly in terms of energy or delay. Here we use degrees of freedom to represent linearly independent packets.

Similar to the case of the star network, our model considers packetized operations, independent of the physical layer implementation of the system. As such, erroneous packets are dropped, and channel losses are measured by a time-averaged erasure rate. This separation of channel and network coding follows from the assumption that physical layer designs are already available for the underlying point-to-point link, with a relay being inserted for performance improvements. In a WBAN, this setup corresponds to allowing a sensor node to serve as a relay for another sensor node. The transmission success rates are assumed to be p_{sr} between s and r, p_{rd} between r and d, and p_{sd} between s and d. Note that, in the previous chapter, we had used p to present packet erasure rates. Nodes operate in half-duplex mode, where a node cannot transmit and receive at the same time. To avoid interferences and collisions in a contention based scheme, we consider a time-division framework, where s and r share the use of the wireless medium. A genie scheduler allocates the wireless medium to the source α fraction of the total time, $0 \leq \alpha \leq 1$, and allocates the wireless medium to the relay the remaining $1 - \alpha$ fraction of time. One possible implementation of such a genie-aided scheduler is to share the same randomness at s and r. Figure 3-1(b) illustrates the maximum flow on each possible link in this network model, computed directly from the transmission success rates and the time-sharing constant α.

In terms of memory, let both s and d contain n units of memory, but assume r contains x units only, where $1 \leq x \leq n$. r uses its memory as a queue: arriving packets are stored; if r is already full, newly arrived packets are discarded. If r does not perform coding, it sends to d a packet from its memory directly and drops this packet from the queue; if r performs random linear network coding before forwarding, it sends to d a linear combination of stored packets, where each is multiplied by a random number chosen uniformly from a finite field $\mathbb{F}_q$ before being added together. In this chapter, $\mathbb{F}_q$ is assumed to be sufficiently large, such that linear combinations resulting from the coding process are linearly independent from each other with high probability.

To evaluate the amount of energy spent to successfully deliver a packet from s to d, we define four different energy terms. E_{tr} represents the transmission energy per packet, where transmission occurs at either s or r. E_{rx} represents the reception energy per packet at r. The relay therefore pays for being on and listening to the broadcast from the source. E_{nc} represents energy for generating a coded packet; it should be a function of n, since the complexity of network coding operations depends on field size and generation size, which is the number of packets coded together. Nonetheless, in this chapter, E_{nc} is assumed to be constant, representing a maximum allowable value. Lastly, E_{ack} represents the amount of energy spent by s to listen to the final acknowledgment from d. Note that all energy terms are defined relative to s or r. It is assumed that the destination d represents a base station without power or energy constraints.

3.3 Network Coding in the Wireless Relay Channel

In this section, three different cases are examined: RLNC at both the relay r and the source s, RLNC at the relay r alone, and RLNC at the source s alone. For the first case, a fluid flow model is considered to analyze the achievable rate, packet delivery energy, and the ratio of these two metrics. For the latter cases, we propose Markov chain models to characterize the expected completion time and the expected completion energy of transmitting n packets from s to d. We also offer a brief discussion on the use of systematic codes, the performance of which will be studied in future works.

3.3.1 Coding at Both the Source s and the Relay r

When RLNC is performed at both the source s and the relay r, we can use the fluid flow model by Lun *et al.* [56, 79] to study the rateless transmission of network coded packets through the relay channel. A packet is considered innovative when it carries a new degree of freedom to a node. In the relay channel, s injects innovative packets into the hyperarc $(s, \{r, d\})$ in α fraction of the total transmission time, while r injects innovative packets into the arc (r, d) in $1 - \alpha$ fraction of the total transmission time. Packet transmissions form innovative flows in this setup because both s and r perform RLNC over a large number of packets. Each mixture is an additional degree of freedom relative to d. The amount of innovative flow is limited by the packet erasure probabilities. Assuming flow conservation at r, the maximum achievable rate R from s to d can be derived by solving the following mathematical programming problem analytically using Fourier-Motzkin elimination [80].

$$\begin{aligned} \min \quad & c(R, \alpha) && (3.1) \\ \text{s.t.} \quad & R \leq \alpha(p_{sr} + p_{sd} - p_{sr}p_{sd}) && (3.2) \\ & R \leq \alpha p_{sd} + (1-\alpha)p_{rd} \\ & 0 < \alpha \leq 1 \end{aligned}$$

Packet-Level Capacity Bound, $c(R, \alpha) = 1/R$

This case is equivalent to maximizing R, making the optimization linear. A closed-form solution can be found:

- Case 1: $p_{sd} \leq p_{rd}$, then

$$R^* = \frac{p_{rd}(p_{sr} + p_{sd} - p_{sd}p_{sr})}{p_{rd} + p_{sr}(1 - p_{sd})}, \quad \alpha^* = \frac{p_{rd}}{p_{rd} + p_{sr}(1 - p_{sd})}. \tag{3.3}$$

- Case 2: $p_{sd} > p_{rd}$, then the relay is not used, and

$$R^* = p_{sd}, \quad \alpha^* = 1. \tag{3.4}$$

Packet-Delivery Energy, $c(R, \alpha) = [E_{tx} + E_{nc} + \alpha(1 - \mathbf{I}_{\alpha=1})E_{rx}]/R$

We define the packet delivery energy as the energy consumed per successfully transmitted packet and state it explicitly in terms of packet transmission, reception, and coding energies. $\mathbf{I}_{\alpha=1}$ represents an indicator function that equals to 1 if $\alpha = 1$, and 0 otherwise. In other words, if $\alpha = 1$, the relay is not used thus should not consume any energy. Because the scheme considered is rateless in the limit of infinitely large payloads, no energy is spent on listening to the acknowledgment at s. Also observe that if $E_{rx} = 0$, this problem reduces to the previous case, and the optimal α^* which achieves the highest throughput also leads to a minimal packet deliver energy. If E_{rx} is non-zero, we can rewrite the optimization in terms of α alone and solve numerically.

Packet-Delivery Energy per Throughput Rate, $c(R, \alpha) = [E_{tx}+E_{nc}+\alpha(1-\mathbf{I}_{\alpha=1})E_{rx}])/R^2$

The ratio of the two metrics described above is the expected packet delivery energy per throughput rate, which can be used to evaluate changes in both energy and throughput.

3.3.2 RLNC at the Relay r Only

First assume that s is limited to sending the original packets only, while r performs RLNC over all packets it has received and stored in memory. Since transmission is rateless with minimalistic feedbacks, s is unaware of the knowledge (degrees of freedom) at r or d. We assume that when allowed to transmit, s chooses one packet uniformly at random from the n available uncoded packets it has in meory.

One remark here is that more exhaustive or frequent feedbacks would allow retransmissions from s to be more intelligent. For example, if per-packet acknowledgment is available, s can choose to repeat only those that have not been successfully received by r or d. On the other hand, if d can acknowledge the exact number of degrees of freedom it received, r can adjust the number of coded packets it sends out to maximize throughput to d while minimizing its own energy use. We consider transmission with minimal feedback here, assuming that the cost of feedback is high.

We describe the states of the system by a three-tuple (m, k, l): m is the number of unique degrees of freedom at d, k is the number of degrees of freedom shared by d and r, and l is the number of degrees of freedom at r only. Since s does not code and r stores only uncoded packets, m and l here represent the numbers of distinct packets at r and d respectively. For example, if $n = 3$, assume r has successfully received packets 1 and 2, while d has received packets 2 and 3, then $(m, k, l) = (1, 1, 1)$. Once a packet at r has been mixed into a coded packet, if it is received at d, it becomes part of the shared degrees of freedom between r and d. For example, if r has received packets 1 and 2, while d has packet 3 and a mixture of packets 1 and 2, then $(m, k, l) = (1, 1, 1)$, since the mixture represents a shared degree of freedom between r and d. With such state definitions, any three-tuple satisfying $m + k + l \leq n$ is a valid state. Transmission initiates in state $(0, 0, 0)$, and terminates in states $\{(m^*, k^*, l^*) | \, m^* + k^* = n\}$.

For transmissions to be free of collisions, recall from the system model that we assume there is a genie-aided scheduler, such that s and r do not access the wireless medium at the same time. In each time slot, s transmits with probability α, while r transmits with probability $1-\alpha$. By randomizing the transmitter at each time slot, the state transition process becomes memoryless, and the numbers of degrees of freedom at each node can be tracked through a Markov chain. The memoryless property holds because the probability of the next transmitted packet being innovative relative to both or either of r and d can be expressed in terms of α and the current state (m, k, l), independent of past state evolutions. With probability α, s chooses one packet uniformly at random from its n uncoded packets, and with probability $(n-m-k-l)/n$, this packet is innovative to both r and d. Similarly, with probability $1-\alpha$, r computes a linear combination of the content of its memory, and sends the resulting mixture to d. Here we assume r has enough memory to store all n distinct packets.

An alternative to the genie-aided randomized transmissions is a collision-free, deterministic schedule, where s and r take turns to transmit for a fixed amount of time, determined by n, α, and the channel conditions. Without feedback, the average system throughput should be the same as the randomized case. However, with this deterministic schedule, since packets are uncoded at s, counting the numbers of degrees of freedom at r and d requires knowledge of the exact packets at these nodes. Even with a small n, it is hard to track the evolution of packets in the system.

Figure 3-2 gives a sample Markov chain when the source has a total of $n = 2$ packets to send to the destination. We have drawn the Markov chain as a tetrahedron, with the starting state on top, and the terminating states on the left vertical edge. State transitions occur after the transmission of a single packet, either from the source s or from the relay r. In the case where r has not received any packet successfully but is chosen to transmit, we assume the allocation slot is wasted. At state (m, k, l), the state transition probabilities can be computed by considering all possible outcomes of the transmission, assuming independent packet losses.

When $m + k < n$, s broadcasts with probability α, while r transmits with probability $1-\alpha$. Let the indicator function $\mathbf{I}_{l>0}$ be 1 when $l \neq 0$, and 0 otherwise, then

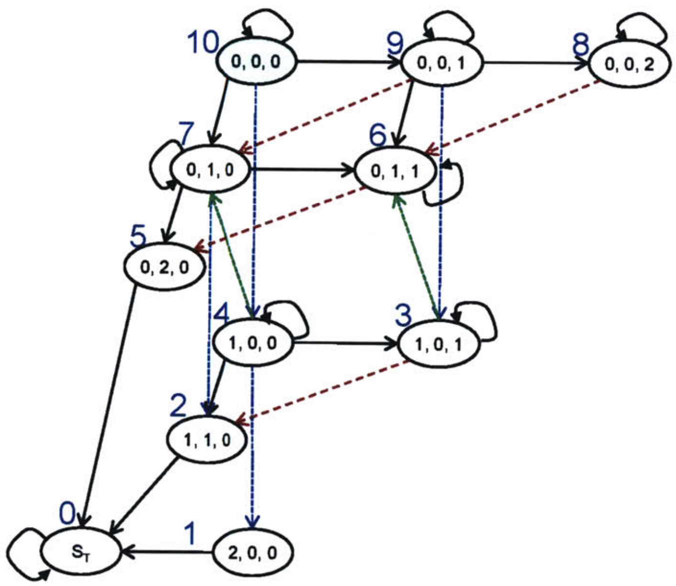


Figure 3-2: Markov chain model, with added terminating state S_T. The number of packets to send at s is $n = 2$. Any three-tuple (m, k, l) satisfying $m + k + l \leq n$ is a valid state.

$$P_{\{m,k,l\}\to\{m+1,k,l\}} = \frac{n-m-k-l}{n} p_{sd}(1-p_{sr})\alpha \tag{3.5}$$

$$P_{\{m,k,l\}\to\{m,k+1,l\}} = \frac{n-m-k-l}{n} p_{sd}p_{sr}\alpha \tag{3.6}$$

$$P_{\{m,k,l\}\to\{m,k,l+1\}} = \frac{n-m-k-l}{n} p_{sr}(1-p_{sd})\alpha \tag{3.7}$$

$$P_{\{m,k,l\}\to\{m-1,k+1,l\}} = \frac{m}{n} p_{sr}\alpha \tag{3.8}$$

$$P_{\{m,k,l\}\to\{m,k+1,l-1\}} = \frac{l}{n} p_{sd}\alpha + \mathbf{I}_{l>0}p_{rd}(1-\alpha) \tag{3.9}$$

$$\begin{aligned} P_{\{m,k,l\}\to\{m,k,l\}} = & \left[\frac{m}{n}(1-p_{sr}) + \frac{l}{n}(1-p_{sd}) + \frac{k}{n} + \frac{n-m-k-l}{n}(1-p_{sd})(1-p_{sr})\right]\alpha \\ & + (1-\mathbf{I}_{l>0}p_{rd})(1-\alpha) \end{aligned} \tag{3.10}$$

When s broadcasts, this transmission may succeed in any of the three hyperarcs originating from s, as shown in Figure 3-1(b). Depending on which packets r and d already have, a successful transmission may or may not lead to a transition to a different state. For example, if the transmitted packet has already been received by both nodes, a state transition does not occur regardless of whether the transmission is successful. In particular, the following state transitions are possible.

1. If d receives one more degree of freedom, while r does not,

   $P_{\{m,k,l\}\to\{m+1,k,l\}} = \frac{n-m-k-l}{n} p_{sd}(1-p_{sr})$;

2. if both d and r receive one more degree of freedom,

   $P_{\{m,k,l\}\to\{m,k+1,l\}} = \frac{n-m-k-l}{n} p_{sd} p_{sr}$;

3. if r receives one more degree of freedom, while d does not,

   $P_{\{m,k,l\}\to\{m,k,l+1\}} = \frac{n-m-k-l}{n} p_{sr}(1-p_{sd})$;

4. if the broadcasted packet has been previously received by d, and is now received at r, $P_{\{m,k,l\}\to\{m-1,k+1,l\}} = \frac{m}{n} p_{sr}$; note that if $m = 0$, this transition probability is 0;

5. if the broadcasted packet has been received by r previously, and is now received at d, then $P_{\{m,k,l\}\to\{m,k+1,l-1\}} = \frac{l}{n} p_{sd}$; if $l = 0$, this transition probability is 0;

6. a self transition occurs if the degree of freedom being sent has previously been received by d but not r, if it has previously been received by r but not d, if it is already shared, or if it has not been shared previously, yet it is not received successfully by r nor d during this transmission; correspondingly, $P_{\{m,k,l\}\to\{m,k,l\}}$ is the sum of four terms: $\frac{m}{n}(1-p_{sr})$, $\frac{l}{n}(1-p_{sd})$, $\frac{k}{n}$, and $\frac{n-m-k-l}{n}(1-p_{sd})(1-p_{sr})$.

Relay r transmits coded packets with probability $1-\alpha$. Observe that, when r has received only a small number of packets, a mixture it generates may not be innovative with respect to d, because it could have been received by d itself already. For example, if n is equal to 3, and d has already received packets 2 and 3, then a coded packet from r containing the sum of packets 2 and 3 is not innovative even if it is successfully received at d, but a coded packet containing the algebraic sum of packets 1 and 2 is innovative. Again, explicitly tracking the contents of coded packets is a difficult task. Instead, we assume that all packets transmitted from the relay to the sink are innovative. The computed expected completion time under this assumption is a lower bound on the actual expected completion time, and the corresponding throughput is an upper bound on the actual system throughput. When discussing numerical results in Section 3.4, we will show that even this upper bound on throughput is not efficient compared with schemes where coding is performed at s, or at both s and r. Also observe that if $n = 1$, an uncoded packet is always transmitted; the innovative packet assumption is then always true, with the computed expected completion time being exact. Under such assumptions, the following state transitions can occur

1. if r has no unique degree of freedom to share, $l = 0$, $P_{\{m,k,0\} \to \{m,k,0\}} = 1$;

2. if r has a unique degree of freedom to share, $l > 0$, and d receives successfully, $P_{\{m,k,l\} \to \{m,k+1,l-1\}} = p_{rd}$;

3. if r has a unique degree of freedom to share, $l > 0$, but d does not receive successfully, $P_{\{m,k,l\} \to \{m,k,l\}} = 1 - p_{rd}$.

The transmission process terminates when $m + k = n$, and $P_{\{m,n-m,0\} \to \{m,n-m,0\}} = 1$. In this Markov chain, all states are transient except the absorbing states, each of which is a recurrent class. When $n > 1$, multiple recurrent classes exist. Since there is a single starting state, there exists a unique steady state distribution. To simplify the computation of the absorbing time, we append a virtual terminating state S_T, such that $P_{\{m,n-m,0\} \to \{S_T\}} = 1$, and $P_{\{S_T\} \to \{S_T\}} = 1$. Thus, the new Markov chain has a single recurrent state only.

To explicitly compute the transition state matrix, we can index the states linearly starting from S_T to $(0,0,0)$. Let S_T be state 0 under this counting notation. In Figure 3-2, we give one possible set of linear indices, starting from the bottom to the top of the tetrahedron.

Let T_i be the expected first passage time to state 0, i.e., the expected number of steps to reach state 0, starting from state i. Also let T be the system expected transmission completion time. Hence $T = T_{10}$ in this example. Since there are no cycles in this Markov chain and the expectation operation is linear, T_i can be solved recursively using the equation

$$T_i = \frac{1}{1 - P_{ii}} \left\{ 1 + \sum_{j \neq i} P_{ij} T_j \right\}, \quad T_0 = 0\,, i \neq 0. \tag{3.11}$$

Alternatively, if we express the expected first passage times to state 0 in vector form $\bar{T} = \begin{bmatrix} T_1 & T_2 & \ldots & T_{10} \end{bmatrix}^T$, then $\bar{T} = 1 + P_{\backslash 0} \bar{T}$, where 1 is the vector of ones, and $P_{\backslash 0}$ is the submatrix of P with the first row and first column removed. Solving this linear system of equations gives $\bar{T} = (\mathbf{I} - P_{\setminus 0})^{-1}1$. Since the original Markov chain does not contain cycles other than loops, the states have a topological order, such that P is a lower-triangular matrix. In addition, since self-transition probabilities are non-zero for all states except those next to S_T, $I - P_{\setminus 0}$ is strictly lower-triangular, thus invertible in the real field.

To determine the expected amount of energy consumed by s and r, let E_i be the expected energy to be computed, starting from state i, and E be the expected transmission completion energy, i.e., $E = E_{10} + E_{ack}$. Here E_{ack} is included since a single acknowledgment is sent by r at the end to signal the end of transmission. A similar argument holds as in the derivation of $\bar{T}$, and a system of linear equation can be solved to find $\bar{E} = \begin{bmatrix} E_1 & E_2 & \dots & E_{10} \end{bmatrix}^T = (\mathbf{I} - P_{\setminus 0})^{-1}(1E_{use}) = \bar{T}E_{use}$, where $E_{use} = E_{tx} + \alpha(1 - \mathbf{I}_{\alpha=1})E_{rx} + (1 - \alpha)E_{nc}$.

One last observation is that it is possible to constrain the amount of memory at r to less than n, and have the relay r function as an accumulator, such that whenever an innovative packet is received, it is multiplied by new random coefficients and added to each of the memory units. When given a transmission opportunity, r uniformly randomly chooses one mixture from its memory. The achievable rate region of the limited memory case should be outer-bounded by the full memory case. We will show in Section 3.4 that even with full memory, coding at r alone is not efficient.

3.3.3 RLNC at the Source s Only

When RLNC is performed at the source s only, the analysis of the system performance is similar to the previous case, where RLNC is performed at the relay r only. First, every mixture sent by s is innovative with respect to r and d under the large field size assumption. Let state (m, k, l) represent m unique mixtures at destination d, k mixtures shared by relay r and destination d, and l unique mixtures at relay r. Assume r acts as a queue with x finite units of memory, where $1 \leq x \leq n$. Here the relay is allowed to have less than n units of memory, since it does not need to store distinct packets for explicit coding operations. The relay r receives linear mixtures directly from the source, functions as a queue, and drops any mixtures received after it is full. Any mixture transmitted from r is also dropped from the queue. A state (m, k, l) is valid as long as $m + k \leq n$ and $k + l \leq x$. Again, the source transmits α fraction of the time. We assume that $P_{\{m,k,l\} \to \{m',k',l'\}}$ is non-zero only if both (m, k, l) and (m', k', l') are valid states. Let the indicator function $\mathbf{I}_{f(\cdot)}$ be 1 if the logic function $f(\cdot)$ is true, and 0 otherwise. State transitions occur after the transmission of a single packet, either from s or from r. The state transition probabilities are as follows,

$$P_{\{m,k,l\} \to \{m+1,k,l\}} = \alpha p_{sd}(1 - p_{sr}) + \alpha p_{sd} p_{sr} \mathbf{I}_{k+l=x} \tag{3.12}$$

$$P_{\{m,k,l\} \to \{m,k+1,l\}} = \alpha p_{sd} p_{sr} \mathbf{I}_{k+l<x} \tag{3.13}$$

$$P_{\{m,k,l\} \to \{m,k,l+1\}} = \alpha p_{sr}(1 - p_{sd}) \mathbf{I}_{k+l<x} \tag{3.14}$$

$$P_{\{m,k,l\} \to \{m,k,l\}} = \alpha p_{sr}(1 - p_{sd}) \mathbf{I}_{k+l=x} + \alpha(1 - p_{sr})(1 - p_{sd}) + (1 - \alpha) \mathbf{I}_{k+l=0} \tag{3.15}$$

$$P_{\{m,k,l\} \to \{m+1,k-1,l\}} = (1 - \alpha) \frac{k}{l + k} \mathbf{I}_{k>0} \tag{3.16}$$

$$P_{\{m,k,l\} \to \{m+1,k,l-1\}} = (1 - \alpha) \frac{l}{l + k} p_{rd} \mathbf{I}_{l>0} \tag{3.17}$$

$$P_{\{m,k,l\} \to \{m,k,l-1\}} = (1 - \alpha) \frac{l}{l + k} (1 - p_{rd}) \mathbf{I}_{l>0} \tag{3.18}$$

In the case where the relay has not received any packet successfully but is chosen to transmit, the allocation slot is assumed to be wasted. Assuming independent packet losses, the state transition probabilities are computed as described below.

First, when $m + k < n$, s broadcasts with probability α, and the following can occur.

1. If d receives the transmitted mixture, but r does not, $P_{\{m,k,l\} \to \{m+1,k,l\}} = p_{sd}(1-p_{sr})$;

2. if both d and r receive the transmitted mixture, and $k + l < x$,

   $P_{\{m,k,l\} \to \{m,k+1,l\}} = p_{sd} p_{sr}$;

3. if both d and r receive the transmitted mixture, and $k + l = x$,

   $P_{\{m,k,l\} \to \{m+1,k,l\}} = p_{sd} p_{sr}$;

4. if r receives the transmitted mixture, but d does not, and $k + l < x$, the mixture is stored in memory, $P_{\{m,k,l\} \to \{m,k,l+1\}} = p_{sr}(1 - p_{sd})$;

5. if r receives the transmitted mixture, but d does not, and $k + l = x$, the mixture is

dropped, $P_{\{m,k,l\}\to\{m,k,l\}} = p_{sr}(1 - p_{sd})$;

6. if neither r nor d receives the packet, $P_{\{m,k,l\}\to\{m,k,l\}} = (1 - p_{sr})(1 - p_{sd})$.

The relay transmits coded packets with probability $1 - \alpha$, and the following state transitions can occur.

1. If r has no unique mixture to share, $l = 0$, and $k = 0$, $P_{\{m,k,l\}\to\{m,k,l\}} = 1$;

2. if r has no unique mixture to share, $l = 0$, and $k > 0$, $P_{\{m,k,l\}\to\{m+1,k-1,l\}} = 1$; observe that since a packet is dropped from r's memory after being sent, k decrements by 1 since the degree of freedom is no longer shared, while m increments by 1 since this degree of freedom becomes unique to d;

3. if r has a unique mixture to share, $l > 0$, and
   - a unique mixture is sent, d receives successfully, $P_{\{m,k,l\}\to\{m+1,k,l-1\}} = \frac{l}{l+k}p_{rd}$;
   - a unique mixture is sent, transmission is unsuccessful, $P_{\{m,k,l\}\to\{m,k,l-1\}} = \frac{l}{l+k}(1 - p_{rd})$;
   - $k > 0$, a shared mixture is sent, d receives successfully, $P_{\{m,k,l\}\to\{m+1,k-1,l\}} = \frac{k}{l+k}p_{rd}$;
   - $k > 0$, a shared mixture is sent, transmission is unsuccessful, $P_{\{m,k,l\}\to\{m+1,k-1,l\}} = \frac{k}{l+k}(1 - p_{rd})$.

Transmission terminates when $m+k = n$, and $P_{\{m,n-m,0\}\to\{m,n-m,0\}} = 1$. Again, a virtual terminating state S_T can be appended, such that $P_{\{m,n-m,l\}\to\{S_T\}} = 1$, and $P_{\{S_T\}\to\{S_T\}} = 1$. With this addition, the Markov chain has one recurrent state only. Figure 3-3 gives a sample Markov chain when the source has $n = 2$ packets to send to the destination. The states can be indexed linearly starting from S_T as state 0, to $(0, 0, 0)$ as the last state, which corresponds to state number 14 in this example.

Our goal is to find the value of α that minimizes either the expected completion time $T = T_{14}$, or the expected completion energy $E = E_{14}$. Unlike the coding at the relay only case, this Markov chain contains cyclic paths in addition to loops. The expected first

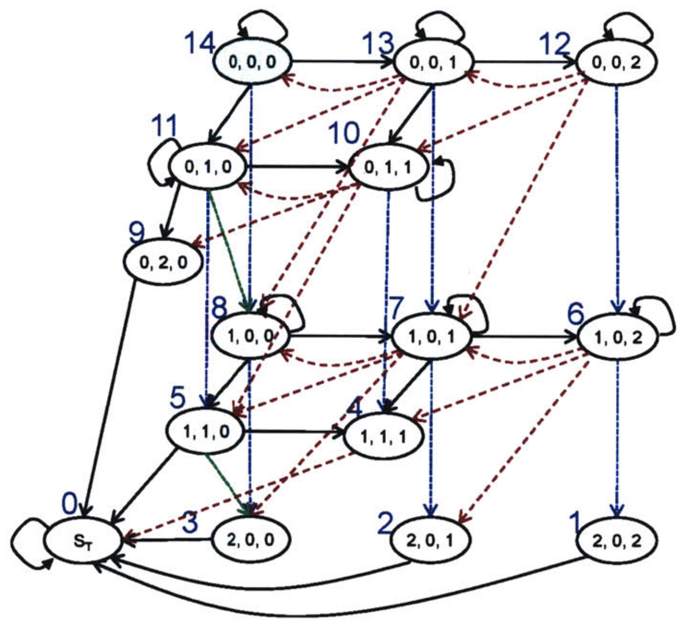


Figure 3-3: Markov chain model, coding at the source s only, with added terminating state S_T. The number of packets to send at s is $n = 2$; the amount of memory at the relay r is $x = 2$.

passage time starting from different states is $\bar{T} = (\mathbf{I} - P_{\backslash 0})^{-1}\mathbf{1}$. Here the invertibility of $\mathbf{I} - P_{\backslash 0}$ is guaranteed because $P_{\backslash 0}$ has entries less than 1 on the main diagonal, and $\mathbf{I} - P_{\backslash 0}$ is a lower Hessenberg matrix with non-zero entries on the main diagonal. Once the value of α that minimizes T is found, we can compute the associated E, where $E = TE_{use} + E_{ack}$. The only difference is the value of E_{use}, since coding is now performed at the source s: $E_{use} = E_{tx} + \alpha(1 - \mathbf{I}_{\alpha=1})E_{rx} + (1 - \alpha)E_{nc}$.

3.3.4 Use of Systematic Codes

Systematic network codes are an attractive alternative to non-systematic random linear network codes, since they can reduce computation complexity and energy use, while maintaining the innovation of independent flows. With a systematic code at the source s only, s can first broadcast uncoded packets one by one in order, then compute random linear mixtures for all remaining transmission from s. The relay r performs the store-and-forward function always. In a sufficiently large finite field, since every packet sent by s is innovative with respect to r and d, if we view each uncoded packet as an innovative mixture, the state evolution under this setup should be the same as the case where nonsystematic random linear network coding is performed at s only. If systematic coding is performed at s and nonsystematic coding is performed at r, the system should give the same performance as the case where nonsystematic coding is performed at both nodes. Another possibility is to perform systematic coding at both s and r. s first broadcasts uncoded packets one by one in order. It then computes a random linear mixture of all n packets whenever a transmission opportunity becomes available. The relay r acts as a size n queue. When it has the opportunity to transmit, it examines the next packet in the queue. If this packet is uncoded, r transmits the uncoded packet directly. If this packet is coded, r linearly combines all data it has in memory before sending out the mixture to d. The system performance under this setup should be upper-bounded by the nonsystematic coding case, and lower-bounded by the coding at s only case. The analysis of this additional systematic phase is non-trivial, so we leave its description and discussion to a later time.

3.4 Numerical Results

This section compares the performance of the three schemes discussed in Section 3.3 under different channel conditions. We first consider the coding at relay only and coding at source only cases and examine the expected transmission completion times per data packet. We then compare the three cases in terms of achievable throughput, computed as the inverse of average completion time. We also compare packet delivery energy for a given set of energy parameters.

3.4.1 RLNC at the Relay r Only

Figure 3-4 plots the expected completion time per transmitted data packet as a function of α for different values of n, when $p_{sd} = 0.5$, $p_{sr} = 0.8$, and $p_{rd} = 0.8$. The optimal α^* value that achieves the lowest T^*/n is indicated by a large dot on each curve.

In this figure, when $\alpha = 1$, the relay listens but does not transmit. If $n = 1$, the expected number of transmissions per data packet is 2. This is the solution to the ARQ scheme when $p_{sd} = 0.5$, where each packet is retransmitted until successfully received at the destination.

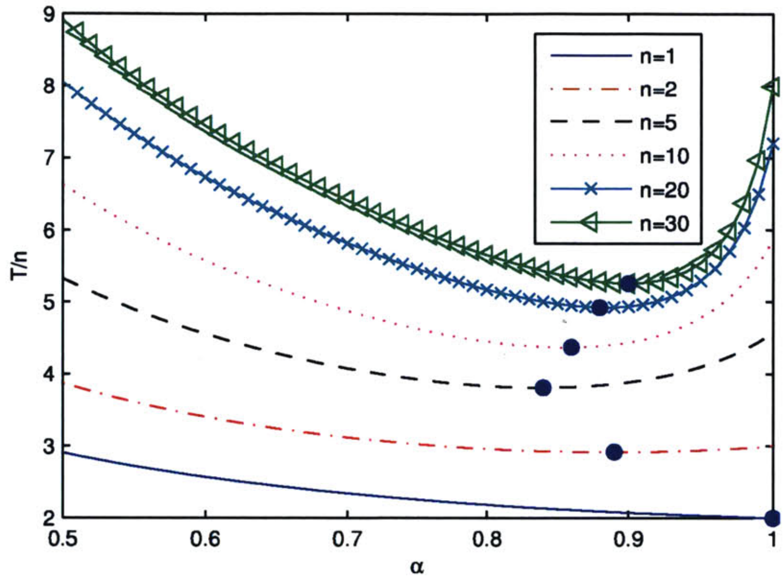


Figure 3-4: Coding at the relay r only, expected completion time per packet T/n vs. α, as n changes in value; packet transmission success rates are $p_{sd} = 0.5$, $p_{sr} = 0.8$, and $p_{rd} = 0.8$. The optimal T^*/n is labeled with a large dot on each curve.

When $n = 2$, the expected number of transmissions per data packet is 3. Observe that since the relay is unused and the source s does not code, s simply retransmits one of the two uncoded data packets each round, until both are received at the destination. This scenario is similar to the coupon collector's problem when there are 2 coupons available, except packet erasures need to be taken into account. When 2 coupons are to be collected, the expected number of trials until success is $2 \times (1 + \frac{1}{2}) = 3$. When divided by p_{sd} and normalized by the number of packets, this solution leads to the value of 3, which is the value on the curve $n = 2$, at $\alpha = 1$, in Figure 3-4.

Another observation from this figure is that as n increases, the expected completion time T/n increases as well. This increase comes from transmissions by the source s. Since s randomly chooses one from n packets to transmit, a packet to be transmitted would have been received by r or d already with non-zero probability. This effect is especially significant towards the end of the transmission, when d has collected most of the necessary degrees of freedom for decoding. In addition, the optimal α values, which correspond to the horizontal coordinates of the large dots, first decrease in value as n goes from 1 to 5, then increase in value as n increases to 20. This effect indicates that a tradeoff exists

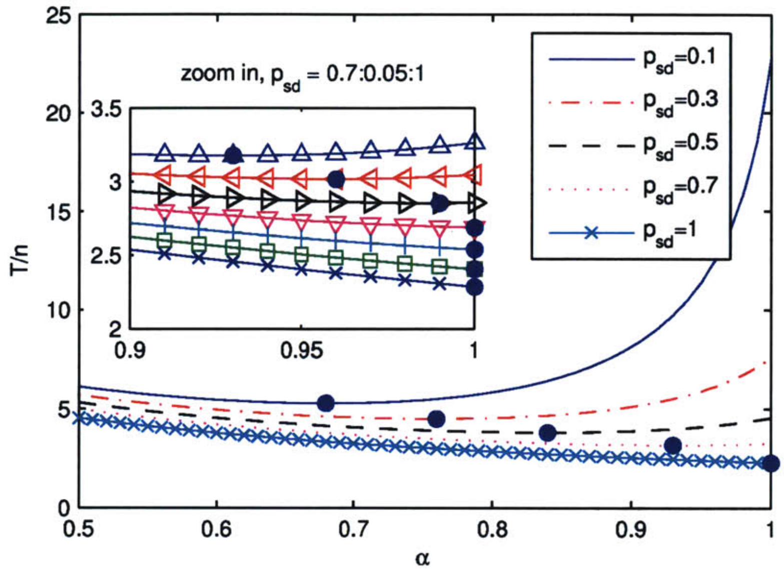


Figure 3-5: Coding at the relay r only, normalized expected completion time T/n vs. α, as p_{sd} changes in value; transmission success rates are $p_{sr} = 0.8$, and $p_{rd} = 0.8$. $n = 5$.

between the use of the relay and the amount of wasted retransmissions by the source.

Figure 3-5 compares the expected completion time per data packet when the channel between s and d varies. Here n is set to be 5. As p_{sd} increases, s is used a larger fraction of the time. When we zoom into the region where p_{sd} increases from 0.7 to 1, it can be observed that r is not used as long as p_{sd} is larger than 0.8. Note that this is similar to the condition $p_{sd} > p_{rd}$ as discussed Section 3.3.1.

From the above numerical evaluations, we can conclude that RLNC at the relay r only while operating in a rateless fashion is not an efficient transmission scheme in terms of throughput. Given $n > 1$, the decision of whether to code at r depends on the channel conditions. Figure 3-4 shows that using ARQ without coding ($n = 1$, $\alpha = 1$) achieves the best expected completion time, or the best throughput. One issue with the $n = 1$ case is that each data packet, when transmitted successfully, requires an acknowledgement from d, i.e., $E/n = TE_{use} + E_{ack}$. Such frequent feedbacks are not energy efficient. If $n > 1$, even though the effect of the E_{ack} term is mitigated by amortization over a larger n, the large increase in the value of T/n shown in Figure 3-4 indicates that coding at the relay only is the most energy efficient when $n = 2$. In Section 3.4.3, we will compare the energy use of this particular case with other schemes.

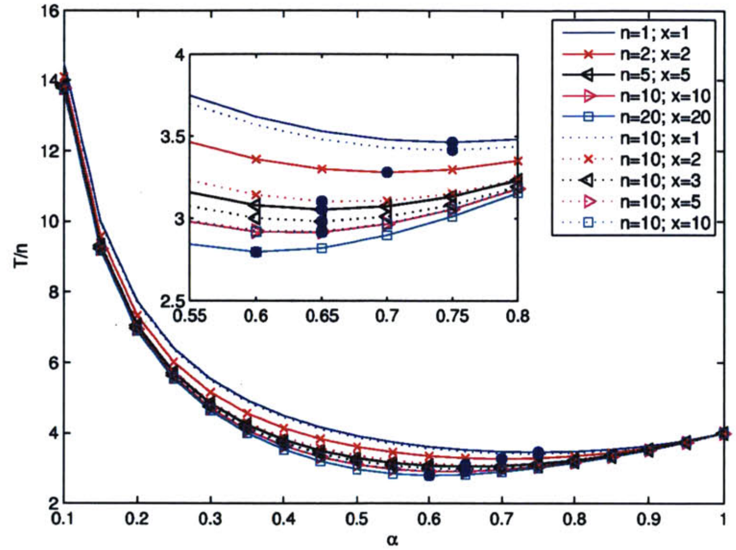


Figure 3-6: Coding at the source s only, expected completion time per packet T/n vs. α, as n and x changes in value; transmission success rates are $p_{sd} = 0.25$, $p_{sr} = 0.8$, and $p_{rd} = 0.8$. The optimal T^*/n is labeled with a large dot on each curve.

3.4.2 RLNC at the Source s Only

Figure 3-6 plots the expected completion time per data packet as a function of α, when n and x vary. n is the number of data packets to be transmitted by the source s, and x is the amount of memory available at the relay r to store received mixtures. Unlike the coding at r only case, here T/n decreases as n becomes larger, because each packet sent by s is innovative relative to both r and d, and as more packets are combined, the probability that a mixture sent by r is innovative to d becomes larger. In addition to reducing T/n, another advantage of coding n packets together at s is that the cost for feedback can be amortized over a large number of data packets. Also observe from this figure that as low as $x = 3$ units of memory suffices to achieve the expected completion time of the full memory case (i.e., $x = n$).

Figure 3-7 plots the expected completion time per data packet as a function of α, for different packet transmission success rates of p_{sd}, p_{sr}, and p_{rd}. Comparison among curves (1), (4) and (5) show that r should be given more time to transmit when the tandem links from s to d through r is more reliable than the direct link between s and d. Comparison between (3) and (4), however, show that r should not be used if the channel between r and

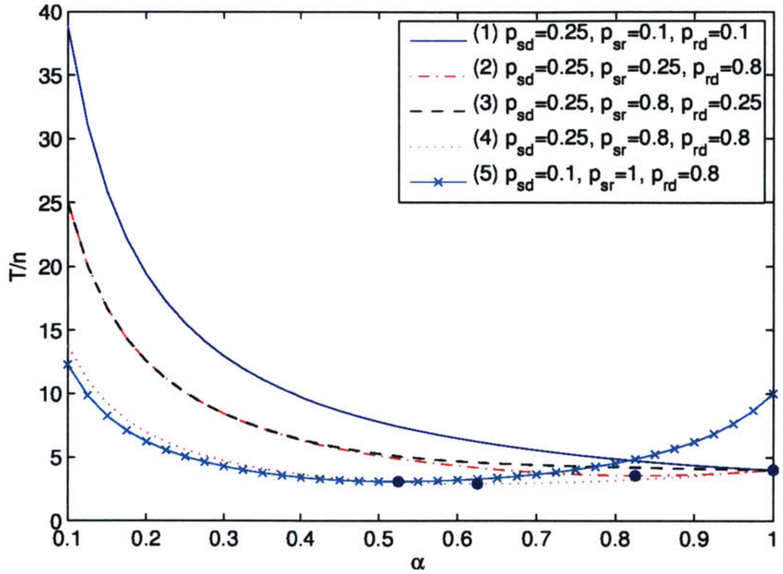


Figure 3-7: Coding at the source s only, expected completion time per packet T/n vs. α, as p_{sr} and p_{rd} change; $p_{sd} = 0.5$, $n = 10$, $x = n$.

d sees large packet erasure probabilities, even if the channel between s and r is relatively reliable. This observation echoes the decision in the full coding case, as discussed in Section 3.3.1. Moreover, comparison among curves (2), (3), and (4) show that the optimal value of α is a function of channel conditions.

3.4.3 Comparisons

Figure 3-8 compares the maximum achievable rates of three cases: coding at the relay r only as discussed in Section 3.3.2, coding at the source s only as discussed in Section 3.3.2, and coding at both s and r as discussed in Section 3.3.1. Figure 3-9 plots the corresponding α^* values that achieve these rates. For the coding at r and coding at s cases, the metric being plotted is the inverse of the optimal expected transmission completion time per data packet (T^*/n). This inverse corresponds to the throughput R^* of the systems under discussion. For the case where coding is performed at both s and r, the achievable rate is computed using Equations (3.3) and (3.4).

When RLNC is performed at the relay r only, as previous discussions have suggested, it is more desirable to mix fewer number of packets; since packets retransmitted from the source are uncoded, a larger fraction of the repetitions are wasted. In Figure 3-8, the

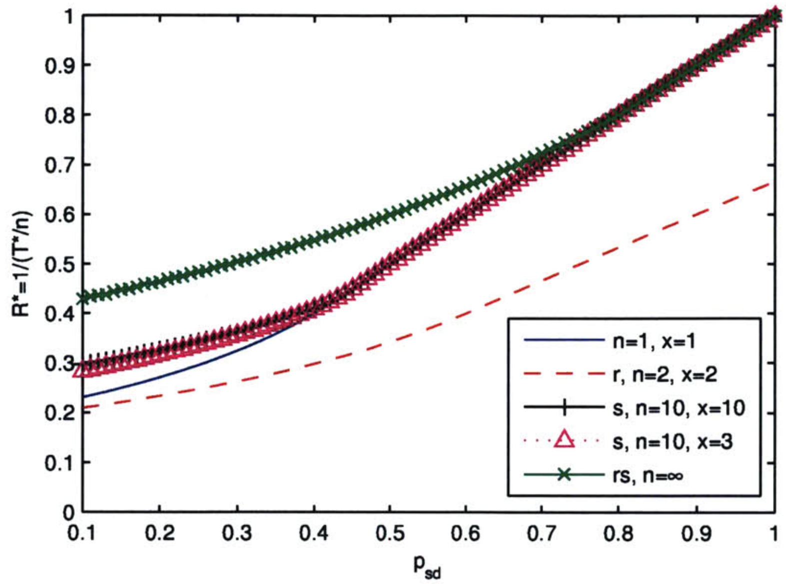


Figure 3-8: Comparison of achievable throughput as a function of p_{sd}, $R^* = \frac{1}{T^*/n}$; packet transmission success rates are $p_{sr} = 0.8$ and $p_{rd} = 0.8$.

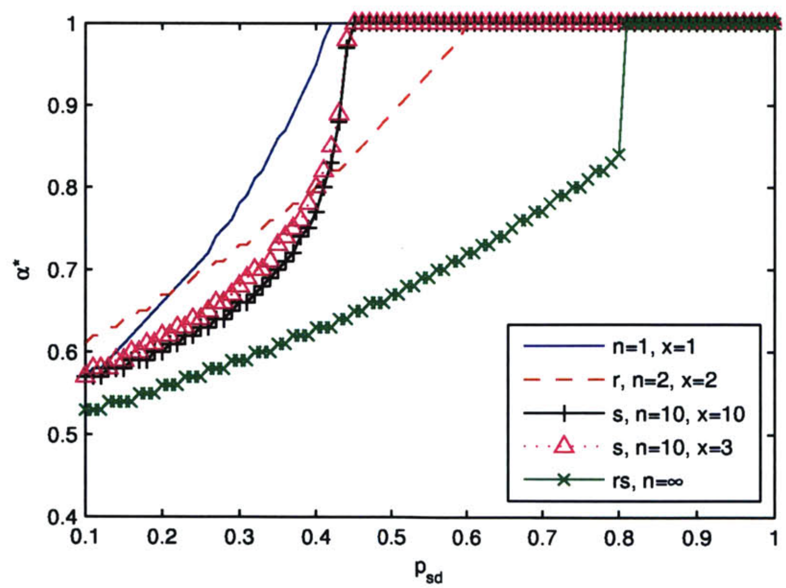


Figure 3-9: Optimal α^* corresponding to throughput values in Figure 3-8; packet transmission success rates are $p_{sr} = 0.8$ and $p_{rd} = 0.8$.

achievable rates are given for two different values of n. When $n = 1$, coding is not performed, hence the transmission degenerates into a routing scheme: s and r retransmit a single packet until an acknowledgment is received from the destination. Observe from Figure 3-9 that when the channel between s and d is poor (eg. $p_{sd} = 0.2$), the route through r is preferred ($\alpha^* \sim 0.65$), otherwise r is not used ($\alpha^* = 1$). When $n = 2$, r still performs network coding, but only as the sum of two packets. Recall the assumption that all mixed packets transmitted by the relay are innovative relative to the destination; the second curve ('$r, n = 2, x = 2$') in Figure 3-8 is therefore an upper bound on the actual system throughput, reconfirming that coding at the relay only is not throughput efficient.

When RLNC is performed at the source s only, Figure 3-8 shows that more than 69% of the rate attained by the coding at both nodes scheme can be achieved. Here the achievable rates are plotted for only one set of channel realizations, with $p_{sr} = 0.8$ and $p_{rd} = 0.8$. The exact amount of coding gain depends on the reliability of all three links in the relay channel. Also observe that the performance gap decreases as the channel between s and d becomes more reliable. Moreover, Figure 3-9 shows that when coding performed at s only, transmissions from r are not required after p_{sd} becomes reasonably good (eg. $p_{sd} > 0.442$). This is because transmissions from the relay r follow a randomized scheme, leading to redundant repetitions that do not contribute additional degrees of freedom to the destination.

In maximizing throughput, coding as much as possible while fully utilizing the relay r seems to be the optimal strategy, followed by coding at s alone. However, assuming that both coding and listening cost power, such approaches may pay more in terms of energy. For example, if E_{tx}, E_{rx}, E_{nc}, E_{ack} are identically 1, Figure 3-10 plots the packet delivery energy corresponding to the optimal α^* in Figure 3-9, while Figure 3-11 plots this energy consumption scaled by the maximum achievable rate. The different energy terms have been chosen assuming that coding and listening consumes energy on the same scale as transmission. Such assumptions are valid in systems where just having the circuitry turned on constitutes the most significant portion of energy use. Other ranges of values are also possible, as we have discussed in the previous chapter, depending on the underlying physical layer hardware implementations.

It is easy to see from these figures that when p_{sd} is low, coding at both the source s and

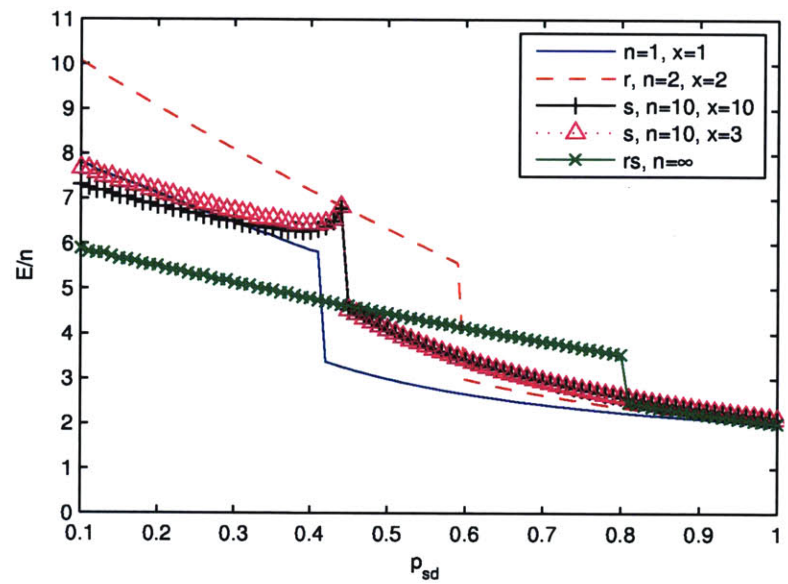


Figure 3-10: Packet delivery energy E/n as a function of p_{sd}, corresponding to the optimal α^* in Figure 3-9; packet transmission success rates are $p_{sr} = 0.8$ and $p_{rd} = 0.8$. Per packet energy use are assumed to be $E_{tx} = 1$, $E_{rx} = 1$, $E_{nc} = 1$, and $E_{ack} = 1$.

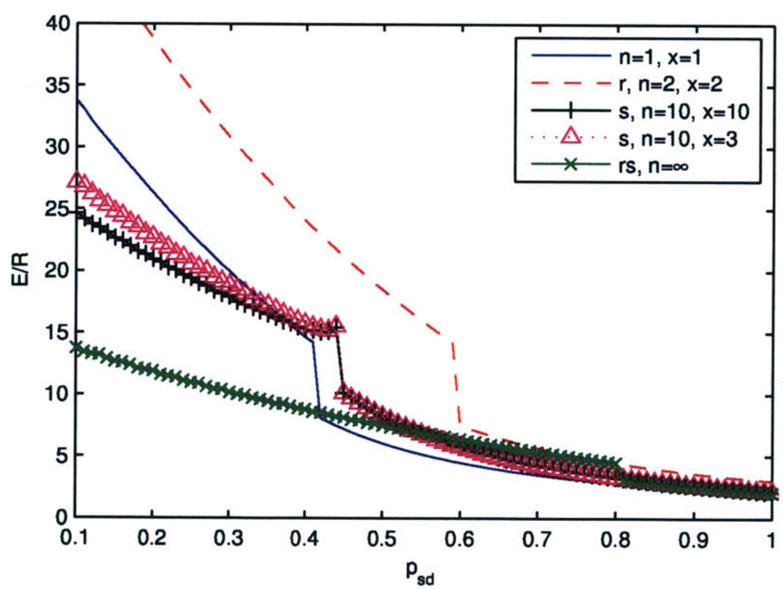


Figure 3-11: Packet delivery energy per throughput rate $E/(nR)$ as a function of p_{sd}, corresponding to the optimal α^* in Figure 3-9; packet transmission success rates are $p_{sr} = 0.8$ and $p_{rd} = 0.8$. Per packet energy use are assumed to be $E_{tx} = 1$, $E_{rx} = 1$, $E_{nc} = 1$, and $E_{ack} = 1$.

the relay r is the most throughput and energy efficient, while coding at s alone provides a compromise between throughput and energy use; under better channel conditions, however, not coding ($n = 1$) and not using the relay ($\alpha = 1$) require less energy, while achieving equally good throughputs. At $p_{sd} = 1$, the energy cost for the successful delivery of one data packet is 2 when coding is conducted at both s and r: one unit is spent on transmission, and one on coding; the energy cost for coding at s only, assuming $n = 10$, is 2.1: one for transmission, one for coding, and $1/10$ for listening to transmission termination acknowledgement; the energy cost for coding at r only, assuming $n = 2$, is 2: according the coupon collector's problem, on average three units of energy are spent on transmitting the 2 packets, and one unit of energy on receiving the acknowledgement; lastly with simple ARQ ($n = 1$), two units of energy are spent on each successfully delivered packet.

Under the same channel conditions and system parameters as given in Figure 3-10, optimizing for energy use leads to a very different set of α values, plotted in Figure 3-13. The corresponding optimal packet delivery energies are shown in Figure 3-12. Observe that the decision to turn off the relay r entirely comes at smaller p_{sd} values. This is because r consumes energy in listening to incoming packets from s as well as sending outgoing packets to the destination. The energy cost of using the relay is the same as retransmitting twice from the source. In addition, since r shares the use of the wireless medium with s, having r turned on reduces the rate at which packets can be transmitted from s. With these two effects combined, r is used only at small p_{sd} values. Another result of the energy tradeoff observable from these two figures is that even though the optimal packet delivery energy curve is continuous, α^* sees a jump for each of the coding strategies.

Note at, in this section, we had plotted packet delivery energy curves for a single set of parameters, namely, when $E_{tx} = 1$, $E_{rx} = 1$, $E_{nc} = 1$, and $E_{ack} = 1$. Similar energy and throughput curves can be evaluated when energy parameters E_{tx}, E_{rx}, E_{nc} and E_{ack} take on different ranges. In practical systems, depending on the underlying circuit implementation, one or more of these energy terms can dominate over the others, and the optimal transmission schedule could be very different from the ones shown above. Nonetheless, our analysis enables robust decision making to determine when and where to code in a wireless packet erasure relay channel.

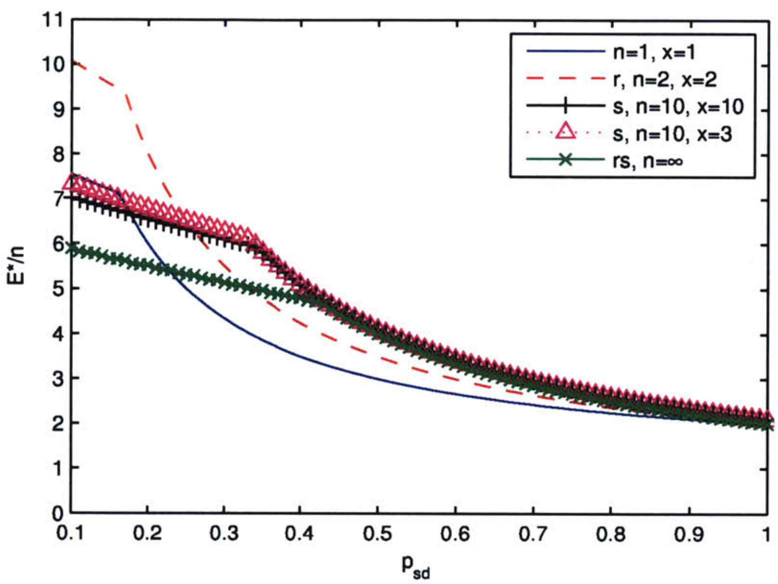


Figure 3-12: Minimum packet delivery energy E^*/n as a function of p_{sd}, $p_{sr} = 0.8$, $p_{rd} = 0.8$, $E_{tx} = 1$, $E_{rx} = 1$, $E_{nc} = 1$, $E_{ack} = 1$.

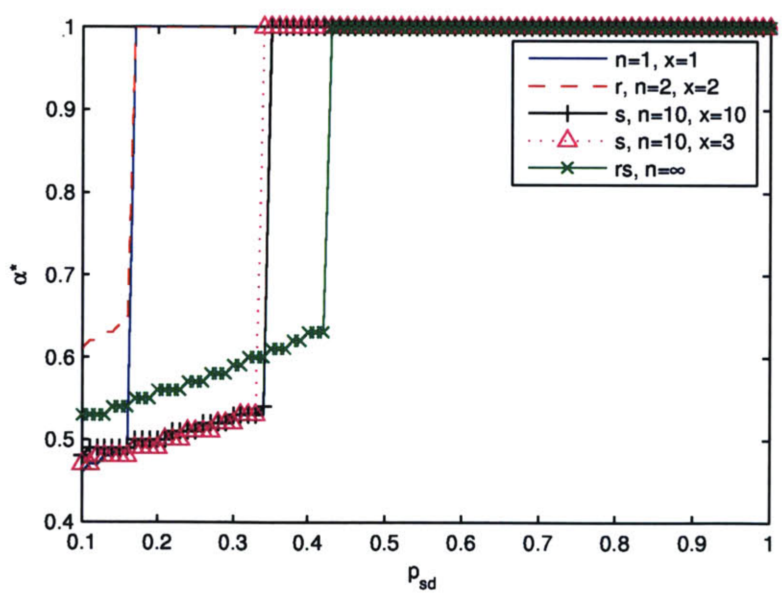


Figure 3-13: Optimal α^* corresponding to packet delivery energy values in Figure 3-12; $p_{sr} = 0.8$, $p_{rd} = 0.8$, $E_{tx} = 1$, $E_{rx} = 1$, $E_{nc} = 1$, $E_{ack} = 1$.

3.5 Conclusions

In this Chapter, we consider a network coded scheme for rateless transmission of data in a three-node, packet erasure relay network. Although the topology is very simple, such relay networks are often found in more general wireless systems. In WBANs, a relaying scheme can be employed if a sensor node is capable of forwarding data for other sensor nodes.

We assume the relay operates in half-duplex mode, either transmitting or receiving, but not both at the same time. Linear network coding can be performed at the source only, at the relay only, or at both nodes. For the first two cases, we establish Markov chain models to track the evolution of innovative packets in the network, and analyze their throughput and packet delivery energy performances. We show through numerical evaluations that using a random code at the relay alone is not throughput efficient, when compared with the coding at both nodes case. We also show that when the per-packet energy cost of transmission, reception and coding are the same, performing network coding at the relay alone is not energy efficient either. Coding at the source alone, on the other hand, can provide a good tradeoff between throughput and energy use. Moreover, we show that only a very small amount of memory is required at the relay when coding is performed at the source only.

Although we do not attempt to explicitly categorize the optimal network coding strategy in the relay channel under different system parameters, we provide, in this chapter, a framework for deciding whether and where to code, taking into account throughput maximization and energy depletion constraints. In practical systems, depending on the underlying circuit implementation, different energy terms may dominate. For example, developments in network coding accelerators may reduce the energy cost of performing linear coding, such that only transmission and reception energies are significant in characterizing energy efficiency of the system. The analytical model we have provided can be used in such cases to evaluate where coding should be performed, and when the relay should be involved in the transmission process. A direct generalization of our given framework to more general networks with arbitrary topologies does not seem tractable. Nonetheless, it is clear from our short analysis that the problem of choosing an optimal coding subgraph is very important when practical constraints, such as energy, are taken into account.

Chapter 4

Conclusions and Future Work

Continuous scaling of integrated circuits and advancements in network protocol design have enabled recent developments in low-power personal wireless communication systems. Wireless body area networks (WBANs), in particular, utilize bio-sensors, and allow inexpensive, continuous collection of physiological data for health monitoring and other medical purposes. A typical WBAN is composed of small sensors attached to the human body, and a central control base-station that collects data wirelessly from these sensors. The central communication problem in a WBAN is to ensure reliable and energy-efficient transmission of measured data from sensors to the base station in the presence of frequent link failures. In this thesis, we explore the use of linear network coding in WBANs, with the goal of determining coding and transmission schemes that are optimal in terms of throughput or energy use. We consider two network topologies. The first is a star-shaped WBAN, where each sensor uploads data directly to the base station. The second is a three-node wireless relay network, where a relay assists in the delivery of data from a sensor to the base station.

In a star-shaped WBAN, we propose a network coded transmission scheme. Assuming that different channel conditions experienced by individual nodes are known at the base station, each sensor node can upload an optimal number of coded packets in each round of the transmission process, taking into account anticipated packet losses, and energy needed for receiving acknowledgment signals. We establish a Markov chain model to analyze the evolution of innovative packets in the network, taking into account interdependencies among sensor nodes through wakeups to listen to feedbacks and scheduling information broadcasted by the base station. Although we do not provide a closed-form solution to the combinatorial optimization problem thus formulated, we show through numerical examples that when receiving costs a lot more than transmitting, using the coded scheme can reduce energy use by a significant factor. The amount of energy reduction achievable through coding depends on the number of nodes in the network, and the asymmetry among channel conditions seen by different sensor nodes.

In addition, we provide a simple heuristic for computing the optimal number of packets to transmit by each node. This heuristic ignores the interdepndencies among sensor nodes, nonetheless provides a good approximation to the optimal scheme in terms of expected transmission completion energy per packet. Furthermore, we examine the IEEE 802.15.6 WBAN standard. Our proposed modifications to the block acknowledgment mode allows network coding to be incorporated into compliant implementations of the standard. In an attempt to verify the practicality of the proposed coded transmission scheme, we have also tried to implement this system using a software-defined radio platform. Timing synchronization and channel estimation seem to be the two most challenging issues in this process. In summary, the work done on a star-shaped WBAN in this thesis helps determine the benefit of network coding in a practical setting, and calls attention to design considerations relevant for implementations that are compatible with current WBAN standards.

In a three-node wireless relay network, we consider a network coded scheme for rateless transmission of data when the relay operates in half-duplex mode. The source node and the relay are assumed to share the wireless medium in the time domain. Linear network coding can be performed at the source alone, at the relay alone, or at both nodes. Again, we propose Markov chain models to track the evolution of innovative packets in the network, thus analyzing the throughput and packet delivery energy performances of the coded scheme when coding is performed at different locations. We show through numerical evaluations that using a random code at the relay alone is not throughput efficient. We also show that when the per-packet energy cost of transmission, reception, and coding are the same, performing coding at the relay alone is not energy efficient either. On the other hand, coding at the source alone, even with limited amount of memory at the relay, can provide a good tradeoff between throughput and energy use. Even though our numerical studies use a single set of energy parameters only, under more general settings where one or more energy terms can dominate over others, the analytical framework we provide can be used to determine whether and where coding should be performed in the relay network, and the fraction of time the relay should participate in the transmission process. A direct generalization of our given framework to networks with arbitrary topologies does not seem tractable. Nonetheless, our study highlights the importance of the network coding subgraph selection problem when practical constraints such as energy are taken into account.

One possible extension of the current study is to consider the use of systematic codes to reduce the amount of energy overheads associated with coding. It is also important to consider the effects of finite field sizes in terms of expected number of additional transmissions to ensure decodability of accepted packets at the base station. Moreover, although we did not mention explicitly in this thesis, an important issue in WBANs is security. On the one hand, security is much desired because of the sensitive nature of medical data; on the other hand, traditional cryptosystems often require intensive computations which impose further energy constraints on sensor nodes. The use of network coding with proper signatures computed from simple hash functions, and partial encryption of network coding coefficients, seem to offer promising solutions to security and privacy issues in WBANs.

Bibliography

[1] "IEEE Standard for Information technology – Local and metropolitan area networks– Specific requirements– Part 15.1a: Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications for Wireless Personal Area Networks (WPAN) ," 2005. [Online]. Available: http://standards.ieee.org/findstds/standard/802.15.1-2005.html (Cited on page 17.)

[2] "IEEE 802.15 WPAN Task Group 6 (TG6) Body Area Networks," 2009. [Online]. Available: http://www.ieee802.org/15/pub/TG6.html (Cited on pages 15, 17, and 44.)

[3] "3GPP TR 36.814 V9.0.0, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9)," 2010. (Cited on pages 65 and 66.)

[4] "IEEE Standard for Local and metropolitan area networks – Part 15.4: Low-Rate Wireless Personal Area Networks (LR-WPANs)," 2011. [Online]. Available: http://standards.ieee.org/findstds/standard/802.15.4-2011.html (Cited on pages 16 and 17.)

[5] "Ettus Research Homepage," 2012. [Online]. Available: http://ettus.com (Cited on page 61.)

[6] "IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks," 2012. [Online]. Available: http://standards.ieee.org/findstds/standard/802.15.6-2012.html (Cited on pages 17, 44, 46, and 56.)

[7] R. Ahlswede, N. Cai, S. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000. (Cited on page 17.)

[8] G. Angelopoulos, M. Médard, and A. Chandrakasan, "Energy-aware hardware implementation of network coding," in *NETWORKING 2011 Workshops*. Springer, 2011, pp. 137–144. (Cited on pages 29 and 30.)

[9] G. Angelopoulos, A. Paidimarri, A. P. Chandrakasan, and M. Médard, "Design, Implementation and Evaluation of Network Coding for Ultra Low Power Sensor Applications," August 2012, private communications. (Cited on pages 42 and 60.)

[10] D. Barry, J. Parlange, L. Li, H. Prommer, C. Cunningham, and F. Stagnitti, "Analytical approximations for real values of the Lambert W-function," *Mathematics and computers in simulation*, vol. 53, no. 1, pp. 95–103, 2000. (Cited on page 33.)

[11] H. Cao, V. Leung, C. Chow, and H. Chan, "Enabling technologies for wireless body area networks: A survey and outlook," *Communications Magazine, IEEE*, vol. 47, no. 12, pp. 84–93, 2009. (Cited on page 16.)

[12] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "MORE: A network coding approach to opportunistic routing," 2006. (Cited on page 19.)

[13] ——, "Trading structure for randomness in wireless opportunistic routing," *ACM SIGCOMM Computer Communication Review*, vol. 37, no. 4, pp. 169–180, 2007. (Cited on page 67.)

[14] F. Chapeau-Blondeau and A. Monir, "Numerical evaluation of the Lambert W function and application to generation of generalized Gaussian noise with exponent 1/2," *IEEE transactions on signal processing*, vol. 50, no. 9, 2002. (Cited on page 33.)

[15] M. Chen, S. Gonzalez, A. Vasilakos, H. Cao, and V. Leung, "Body area networks: a survey," *Mobile Networks and Applications*, vol. 16, no. 2, pp. 171–193, 2011. (Cited on page 16.)

[16] P. Chou, Y. Wu, and K. Jain, "Practical network coding," in *Proceedings of the Annual Allerton Conference on Communication Control and Computing*, vol. 41, no. 1. The University; 1998, 2003, pp. 40–49. (Cited on pages 19 and 54.)

[17] R. Corless, G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, "On the Lambert W function," *Advances in Computational mathematics*, vol. 5, no. 1, pp. 329–359, 1996. (Cited on page 33.)

[18] T. Cover and A. Gamal, "Capacity theorems for the relay channel," *Information Theory, IEEE Transactions on*, vol. 25, no. 5, pp. 572–584, 1979. (Cited on page 67.)

[19] D. Daly, P. Mercier, M. Bhardwaj, A. Stone, Z. Aldworth, T. Daniel, J. Voldman, J. Hildebrand, and A. Chandrakasan, "A pulsed UWB receiver SoC for insect motion control," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 1, pp. 153–166, 2010. (Cited on pages 30 and 41.)

[20] A. Dana, R. Gowaikar, R. Palanki, B. Hassibi, and M. Effros, "Capacity of wireless erasure networks," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 789–804, 2006. (Cited on pages 18 and 67.)

[21] D. Davenport, N. Seidl, J. Moss, M. Patel, A. Batra, J.-M. Ho, S. Hosur, J. Roh, T. Schmidl, O. Omeni, and A. Wong, "MedWiN MAC and Security Proposal," May 2009. (Cited on page 17.)

[22] ——, "MedWiN Physical Layer Proposal," May 2009. (Cited on page 17.)

[23] A. El-Hoiydi and J. Decotignie, "WiseMAC: an ultra low power MAC protocol for the downlink of infrastructure wireless sensor networks," in *Computers and Communications, 2004. Proceedings. ISCC 2004. Ninth International Symposium on*, vol. 1. Ieee, 2004, pp. 244–251. (Cited on page 16.)

[24] E. Erez and M. Feder, "Convolutional network codes," in *Proc. of IEEE Int. Sym. on Info. Theory (ISIT)*, 2005, p. 146. (Cited on page 18.)

[25] P. Fan, C. Zhi, C. Wei, and K. Ben Letaief, "Reliable relay assisted wireless multicast using network coding," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 5, pp. 749–762, 2009. (Cited on page 68.)

[26] G. Fang and E. Dutkiewicz, "BodyMAC: Energy efficient TDMA-based MAC protocol for wireless body area networks," in *ISCIT 2009*. IEEE, pp. 1455–1459. (Cited on page 16.)

[27] C. Fragouli and E. Soljanin, "A connection between network coding and convolutional codes," in *Communications, 2004 IEEE International Conference on*, vol. 2. IEEE, 2004, pp. 661–666. (Cited on page 18.)

[28] ——, "Network coding applications," *Foundations and Trends® in Networking*, vol. 2, no. 2, pp. 135–269, 2007. (Cited on page 17.)

[29] ——, "Network coding fundamentals," *Foundations and Trends® in Networking*, vol. 2, no. 1, pp. 1–133, 2007. (Cited on pages 13 and 17.)

[30] S. Gollakota and D. Katabi, "Zigzag decoding: Combating hidden terminals in wireless networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 159–170, 2008. (Cited on pages 20 and 61.)

[31] W. Guo, N. Cai, X. Shi, and M. Medard, "Localized dimension growth in random network coding: A convolutional approach," in *ISIT 2011*. IEEE, 2011, pp. 1156–1160. (Cited on pages 18 and 22.)

[32] C. Hausl and J. Hagenauer, "Iterative network and channel decoding for the two-way relay channel," in *Communications, 2006. ICC'06. IEEE International Conference on*, vol. 4. IEEE, 2006, pp. 1568–1573. (Cited on page 67.)

[33] T. Ho, M. Médard, R. Koetter, D. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, 2006. (Cited on pages 17 and 54.)

[34] S. Jaggi, P. Chou, and K. Jain, "Low complexity algebraic multicast network codes," in *Information Theory, 2003. Proceedings. IEEE International Symposium on*. IEEE, pp. 368–368. (Cited on page 18.)

[35] S. Jaggi, P. Sanders, P. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *Information Theory, IEEE Transactions on*, vol. 51, no. 6, pp. 1973–1982, 2005. (Cited on page 18.)

[36] C. Jones, K. Sivalingam, P. Agrawal, and J. Chen, "A survey of energy efficient network protocols for wireless networks," *Wireless Networks*, vol. 7, no. 4, pp. 343–358, 2001. (Cited on page 14.)

[37] S. Katti, S. Gollakota, and D. Katabi, "Embracing wireless interference: Analog network coding," in *Proceedings of the 2007 conference on Applications, technologies, architectures, and protocols for computer communications*. ACM, 2007, p. 408. (Cited on page 20.)

[38] S. Katti, D. Katabi, H. Balakrishnan, and M. Médard, "Symbol-level network coding for wireless mesh networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 401–412, 2008. (Cited on pages 20 and 61.)

[39] S. Katti, I. Maric, A. Goldsmith, D. Katabi, and M. Médard, "Joint relaying and network coding in wireless networks," in *IEEE International Symposium on Information Theory, 2007. ISIT 2007*, 2007, pp. 1101–1105. (Cited on page 20.)

[40] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *IEEE/ACM Transactions on Networking (TON)*, vol. 16, no. 3, pp. 497–510, 2008. (Cited on pages 19, 61, and 67.)

[41] M. Kim, D. Lucani, X. Shi, F. Zhao, and M. Médard, "Network coding for multiresolution multicast," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9. (Cited on page 22.)

[42] M. Kim, M. Médard, and J. Barros, "Modeling network coded TCP throughput: A simple model and its validation," *Arxiv preprint arXiv:1008.0420*, 2010. (Cited on page 19.)

[43] R. Koetter and F. Kschischang, "Coding for errors and erasures in random network coding," *Information Theory, IEEE Transactions on*, vol. 54, no. 8, pp. 3579–3591, 2008. (Cited on page 18.)

[44] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking (TON)*, vol. 11, no. 5, pp. 782–795, 2003. (Cited on page 17.)

[45] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *Information Theory, IEEE Transactions on*, vol. 51, no. 9, pp. 3037–3063, 2005. (Cited on page 67.)

[46] K. Kwak, M. Ameen, D. Kwak, C. Lee, and H. Lee, "A study on proposed IEEE 802.15 WBAN MAC protocols," in *ISCIT 2009*. IEEE, pp. 834–840. (Cited on page 17.)

[47] L. Lai, K. Liu, and H. El Gamal, "The three-node wireless network: Achievable rates and cooperation strategies," *Information Theory, IEEE Transactions on*, vol. 52, no. 3, pp. 805–828, 2006. (Cited on page 67.)

[48] J. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3062–3080, 2004. (Cited on page 67.)

[49] C. Li, H. Li, and R. Kohno, "Performance evaluation of IEEE 802.15. 4 for wireless body area network (WBAN)," in *Communications Workshops, 2009. ICC Workshops 2009. IEEE International Conference on*. IEEE, 2009, pp. 1–5. (Cited on page 16.)

[50] H. Li and J. Tan, "An ultra-low-power medium access control protocol for body sensor network," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*. IEEE, 2006, pp. 2451–2454. (Cited on page 16.)

[51] S. Li and R. Yeung, "On convolutional network coding," in *Proc. of IEEE Int. Sym. on Info. Theory*, 2006, pp. 1743–1747. (Cited on page 18.)

[52] S. Li, R. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, 2003. (Cited on page 17.)

[53] D. Lucani, M. Médard, and M. Stojanovic, "Broadcasting in Time-Division Duplexing: A Random Linear Network Coding Approach," in *Proc. NetCod*, vol. 9, 2009, pp. 62–67. (Cited on page 36.)

[54] D. Lucani, M. Stojanovic, and M. Médard, "Random Linear Network Coding For Time Division Duplexing: Energy Analysis," in *Proc. ICC*, vol. 9, 2009. (Cited on page 19.)

[55] ——, "Random linear network coding for time division duplexing: When to stop talking and start listening," in *Proc. INFOCOM*, vol. 9, 2009, pp. 1800–1808. (Cited on pages 19 and 26.)

[56] D. Lun, M. Médard, R. Koetter, and M. Effros, "On coding for reliable communication over packet networks," *Physical Communication*, vol. 1, no. 1, pp. 3–20, 2008. (Cited on pages 18, 67, 68, and 70.)

[57] M. Lutz, *Programming python*. O'Reilly Media, Inc., 2006. (Cited on page 61.)

[58] I. Maric, A. Goldsmith, and M. Médard, "Analog Network Coding in the High SNR Regime," in *invited paper, ITA workshop*, 2010. (Cited on page 20.)

[59] S. Marinkovic, E. Popovici, C. Spagnol, S. Faul, and W. Marnane, "Energy-efficient low duty cycle MAC protocol for wireless body area networks," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 6, pp. 915–925, 2009. (Cited on page 16.)

[60] M. Médard, M. Effros, D. Karger, and T. Ho, "On coding for non-multicast networks," in *Proc. of the 41st Allerton Conference*, vol. 41, no. 1, 2003, pp. 21–29. (Cited on page 18.)

[61] M. Médard, D. Lun *et al.*, "Efficient operation of coded packet networks," Ph.D. dissertation, Massachusetts Institute of Technology, 2006. (Cited on pages 18 and 67.)

[62] M. Médard, D. Shah, J. Sundararajan *et al.*, "On the role of feedback in network coding," Ph.D. dissertation, Massachusetts Institute of Technology, 2009. (Cited on page 19.)

[63] M. Médard and A. Sprintson, *Network Coding: Fundamentals and Applications*. Academic Press, 2011. (Cited on pages 13 and 17.)

[64] P. Mercier and A. Chandrakasan, "A Supply-Rail-Coupled eTextiles Transceiver for Body-Area Networks," *Solid-State Circuits, IEEE Journal of*, vol. 46, no. 6, pp. 1284–1295, 2011. (Cited on page 15.)

[65] P. Mercier, D. Daly, and A. Chandrakasan, "An Energy-Efficient All-Digital UWB Transmitter Employing Dual Capacitively-Coupled Pulse-Shaping Drivers," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 6, pp. 1679–1688, 2009. (Cited on pages 30 and 41.)

[66] J. Mitola, "The software radio architecture," *Communications Magazine, IEEE*, vol. 33, no. 5, pp. 26–38, 1995. (Cited on page 60.)

[67] O. Omeni, A. Wong, A. Burdett, and C. Toumazou, "Energy efficient medium access protocol for wireless medical body area sensor networks," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 2, no. 4, pp. 251–259, 2008. (Cited on page 16.)

[68] B. Otal, L. Alonso, and C. Verikoukis, "Highly reliable energy-saving MAC for wireless body sensor networks in healthcare systems," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 4, pp. 553–565, 2009. (Cited on page 16.)

[69] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *Wireless Communications, IEEE*, vol. 17, no. 1, pp. 80–88, 2010. (Cited on page 15.)

[70] T. Rondeau, E. Blossom, J. Corgan, M. Ettus *et al.*, "GNU Radio Wiki," 2011. [Online]. Available: http://gnuradio.org/redmine/projects/gnuradio/wiki (Cited on page 61.)

[71] J. Ryckaert, C. Desset, A. Fort, M. Badaroglu, V. De Heyn, P. Wambacq, G. Van der Plas, S. Donnay, B. Van Poucke, and B. Gyselinckx, "Ultra-wide-band transmitter for low-power wireless body area networks: Design and evaluation," *Circuits and Systems I: Regular Papers, IEEE Trans. on*, vol. 52, no. 12, pp. 2515–2525, 2005. (Cited on pages 30 and 41.)

[72] X. Shi, M. Médard, and D. Lucani, "When both transmitting and receiving energies matter: an application of network coding in wireless body area networks," in *NETWORKING 2011 Workshops*. Springer, 2011, pp. 119–128. (Cited on pages 21 and 68.)

[73] ——, "Whether and Where to Code in the Wireless Packet Erasure Relay Channel," *accepted to the IEEE Journal on Selected Areas in Communications: Special Issue on Theories and Methods for Advanced Wireless Relays*, 2012. (Cited on page 21.)

[74] E. Shih, S. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan, "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proceedings of the 7th annual international conference on Mobile computing and networking*. ACM New York, NY, USA, 2001, pp. 272–287. (Cited on page 19.)

[75] D. Silva, F. Kschischang, and R. Koetter, "A rank-metric approach to error control in random network coding," *Information Theory, IEEE Transactions on*, vol. 54, no. 9, pp. 3951–3967, 2008. (Cited on page 18.)

[76] J. Sundararajan, D. Shah, M. Médard, M. Mitzenmacher, and J. Barros, "Network coding meets TCP," in *Proceedings of IEEE INFOCOM*, 2009, pp. 280–288. (Cited on pages 19 and 58.)

[77] C. Tachtatzis, F. Di Franco, D. Tracey, N. Timmons, and J. Morrison, "An energy analysis of IEEE 802.15.6 scheduled access modes," in *GLOBECOM Workshops (GC Wkshps), 2010 IEEE*. IEEE, 2010, pp. 1270–1275. (Cited on page 17.)

[78] S. Teerapittayanon, K. Fouli, M. Médard, M. M.-J., X. Shi, I. Seskar, and A. Gosain, "Network Coding as a WiMAX Link Reliability Mechanism," in *5th International Workshop on Multiple Access Communications (MACOM)*, 2012. (Cited on page 22.)

[79] D. Traskov, M. Heindlmaier, M. Médard, R. Koetter, and D. Lun, "Scheduling for network coded multicast: A conflict graph formulation," in *GLOBECOM Workshops, 2008 IEEE*. IEEE, 2008, pp. 1–5. (Cited on pages 18 and 70.)

[80] D. Traskov, "Network Coding for the Multiple Access Layer," Ph.D. dissertation, Technische Universität München, 2010. (Cited on pages 67 and 71.)

[81] D. Tuninetti and C. Fragouli, "Processing along the way: forwarding vs. coding," *ISITA 2005*, 2004. (Cited on page 67.)

[82] ——, "On the throughput improvement due to limited complexity processing at relay nodes," in *Information Theory, 2005. ISIT 2005. Proceedings. International Symposium on*. IEEE, 2005, pp. 1081–1085. (Cited on page 67.)

[83] S. Ullah, H. Higgins, B. Braem, B. Latre, C. Blondia, I. Moerman, S. Saleem, Z. Rahman, and K. Kwak, "A comprehensive survey of wireless body area networks," *Journal of medical systems*, pp. 1–30, 2010. (Cited on page 16.)

[84] S. Ullah and K. Kwak, "Throughput and delay limits of IEEE 802.15.6," in *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. IEEE, 2011, pp. 174–178. (Cited on page 17.)

[85] T. Ulversoy, "Software defined radio: Challenges and opportunities," *Communications Surveys & Tutorials, IEEE*, vol. 12, no. 4, pp. 531–550, 2010. (Cited on page 60.)

[86] T. Van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*. ACM, 2003, pp. 171–180. (Cited on page 16.)

[87] E. van der Meulen, "Transmission of information in a T-terminal discrete memoryless channel," Ph.D. dissertation, University of California, 1968. (Cited on page 67.)

[88] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "A network coding approach to cooperative diversity," *Information Theory, IEEE Transactions on*, vol. 53, no. 10, pp. 3714–3722, 2007. (Cited on page 67.)

[89] S. Yang and R. Koetter, "Network coding over a noisy relay: a belief propagation approach," in *Information Theory, 2007. ISIT 2007. IEEE International Symposium on*. IEEE, 2007, pp. 801–804. (Cited on page 67.)

[90] Y. Yao, X. Cai, and G. Giannakis, "On energy efficiency and optimum resource allocation of relay transmissions in the low-power regime," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 6, pp. 2917–2927, 2005. (Cited on page 67.)

[91] K. Yazdandoost and K. Sayrafian-Pour, "Channel model for body area network (BAN)," *IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs) (IEEE P802.15-08-0033-00-0006)*, 2008. (Cited on page 15.)

[92] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *Networking, IEEE/ACM Transactions on*, vol. 12, no. 3, pp. 493–506, 2004. (Cited on page 16.)

[93] R. Yeung, S. Li, and N. Cai, *Network coding theory*. Now Pub, 2006. (Cited on page 17.)

[94] S. Zhang and S. Liew, "Channel coding and decoding in a relay system operated with physical-layer network coding," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 5, pp. 788–796, 2009. (Cited on page 67.)

[95] S. Zhang, S. Liew, and P. Lam, "Hot topic: Physical-layer network coding," in *Proceedings of the 12th annual international conference on Mobile computing and networking*. ACM, 2006, p. 365. (Cited on page 20.)

[96] F. Zhao and M. Médard, "On analyzing and improving COPE performance," *invited paper, ITA Workshop*, 2010. (Cited on page 19.)

[97] Y. Zhao, R. Adve, and T. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," in *Information Theory, 2006 IEEE International Symposium on*. IEEE, 2007, pp. 1234–1238. (Cited on page 67.)