RFID: A Communication System Perspective

Colby Scott Boyer

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Sumit Roy, Chair

Joshua R Smith

Pavel Nikitin

Program Authorized to Offer Degree:
Electrical Engineering

University of Washington

Abstract

RFID: A Communication System Perspective

Colby Scott Boyer

Chair of the Supervisory Committee:
Professor Sumit Roy
Electrical Engineering

Backscatter modulation is the keystone behind the technical and commercial success of Radio Frequency IDentification (RFID) in applications that require fast, efficient, and automatic identification of tagged objects. Owing to RFID's ultra low cost and low power attributes, wireless sensor networks (WSNs) and other biological/physiological sensing applications now use backscatter modulation as a communication link. These systems build upon existing RFID industry protocols such as ISO 18000-6C (or EPC Gen2).

RFID tags do not use an active RF transceiver architecture, but instead operate on a principle similar to RADAR. A nearby RFID reader transmits a continuous wave that the tag's antenna partially scatters back to the reader's antenna array. By switching its antenna impedance loading between different values, the tag modulates the antenna scattering observed by the reader. Therefore reverse engineering RFID at the systems level will provide new insights and design recommendations. This dissertation focuses on three key areas relating to backscatter modulation and RFID systems: 1) analyze the co-design of backscatter modulation and error correction coding, 2) characterize space-time coding and MIMO performance limits of the dyadic backscatter channel, and 3) hybrid-ARQ (HARQ) and random access improvements to the Gen2 protocol.

Passive and semi-passive RFID tags depend on scavenged energy to power their IC. While backscatter modulation itself consumes a negligible amount of energy, the modulator creates an impedance mismatch between the tag's antenna and power harvester, thereby decreasing

the antenna to the tag power transfer efficiency. This required impedance mismatch couples the link performance to the power harvester's performance, so to quantify this tradeoff, we introduce a new metric: backscatter power efficiency loss per bit. Higher order constellations improve the link's spectral efficiency, but have lower power efficiency as compared to binary modulation schemes. We propose new coded modulation schemes based on unequal error protection, which improves both the spectral efficiency and the backscatter power efficiency loss metric.

MIMO processing is a canonical technique to improve wireless link capacity and reliability, which will require future tags and readers to have multiple antennas. The dyadic backscatter channel (DBC) models the behavior of small-scale fading in RFID MIMO systems, however, its statistics differ from those of the classic Rayleigh fading MIMO channel. We analytical characterize the performance of space-time trellis codes and orthogonal space-time block codes, derive an upper bound to the pairwise error probability (PEP), and derive the maximum diversity order of the DBC. Unlike Rayleigh fading, the diversity order only depends on the number of tag antennas but not the number of reader receive antennas.

In fading channels, MIMO techniques offer two opposing performance gains: diversity (reliability against outage events) or multiplexing gain (spectral efficiency). The diversity multiplexing tradeoff (DMT) is an asymptotic measure that quantifies the achievable diversity for a given multiplexing gain. Starting from the definition of the DBC and the DMT of the double scattering channel, the corresponding DMT of the DBC is derived. The statistics of the DBC limit the amount diversity when compared to the Rayleigh MIMO channel, although the available multiplexing gain is unchanged. Increasing the number of receive antennas improves both diversity and multiplexing gain until the receive antenna count equals the number of tag antennas, otherwise additional receive antennas offer no gains with respect to the DMT.

The current EPC Gen2 standard does not use any form of error correction and does not allow for fast link adaption between reading separate tags. We consider a protocol that uses HARQ algorithms without requiring major changes to the Gen2 protocol. In addition,

we develop theoretical models that capture EPC Gen2's baseline performance and capacity in terms SNR and tag read rate. Existing HARQ algorithms, such as Chase combining (CC) and incremental redundancy (IR), are studied via simulations and the performance quantified in terms of tag read rate. The simulation results show that CC allows for graceful system degradation and IR achieves read rates close to EPC Gen2's capacity limit.

Random access plays a critical role in RFID tag singulation. EPC Gen2 uses frame slotted ALOHA (FSA) to arbitrate channel resources between tags, but FSA has low efficiency due to empty slots and collisions. To aid in tag collision resolution, we consider multiuser detection (MUD) and incremental redundancy enhancements to the FSA protocol. The theoretical performance of the MUD receiver is analyzed from a compressive sensing viewpoint. As an example of a practical code construction, we evaluate the performance of punctured second order Reed Muller codes. These enhancements improve FSA's throughput and its response to high system loads.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

ARQ: Automatic Repeat Request

ASK: Amplitude Shift Keying

AWGN: Additive White Gaussian Noise

BER: Bit Error Rate

BPSK: Binary Phase Shift Keying

CC: Chase Combining

CRC: Cyclic Redundancy Check

CS: Compressive Sensing

CSI: Channel State Information

CW: Continuous Wave

DBC: Dyadic Backscatter Channel

DMT: Diversity Multiplexing Tradeoff

E-FIELD: Electric field

EEP: Equal Error Protection

EM: Electro-Magnetic

EPC: Electronic Product Code

FEC: Forward Error Correction

FSA: Frame Slotted ALOHA

HARQ: Hybrid ARQ

IC: Integrated Circuit

IR: Incremental Redundancy

IR-FSA: Incremental Redundancy for Frame Slotted ALOHA

MAC: Media Access Control or Multiple Access Channel

MIMO: Multiple Input Multiple Output

MISO: Multiple Input Single Output

ML: Maximum Likelihood

MLSE: Maximum Likelihood Sequence Estimation

MUD: Multiuser Detection

OMP: Orthogonal Matching Pursuit

OSTBC: Orthogonal Space Time Block Code

PDF: Probability Density Function

PEP: Pairwise Error Probability

PIE: Pulse Interval Code

PSK: Phase Shift Keying

QAM: Quadrature Amplitude Modulation

QPSK: Quadrature Phase Shift Keying

RCPC: Rate Compatible Punctured Convolutional

RF: Radio Frequency

RFID: Radio Frequency Identification

RM: Reed Muller

RN16: Random 16 bit

SA: Slotted ALOHA

SEQOMP: Sequential OMP

SER: Symbol Error Rate

SISO: Single Input Single Output

SNR: Signal to Noise Ratio

STBC: Space Time Block Code

STTC: Space Time Trellis Code

UEP: Unequal Error Protection

UWB: Ultra Wide Band

WISP: Wireless Integrated Sensing Platform

WSN: Wireless Sensor Network

ACKNOWLEDGMENTS

This dissertation would have not been possible without the help and support of my colleagues, friends, and family.

I thank my PhD adviser, Prof. Sumit Roy, for his guidance throughout my time in graduate school. His commitment to both research and excellence serves as inspiration for my own work.

I thank my committee, Prof. Joshua Smith and Dr. Pavel Nikitin, for their advice and suggestions, which have undoubtedly improved the quality of this dissertation. Additionally, I would like to thank Prof. Peter Smith of the University of Canterbury, NZ for his help with random matrices.

I would like to thank the current and past members of the Fundamentals of Networking Laboratory and fellow UW graduate students for their insightful discussions and advice throughout my stay at UW. Fei Ye, Ling Luo, Chittabrata Ghosh, Rohit Gupta, Hamed Firooz, Linda Bai, Trevor Bosaw, Yoshikazu Nishida, Joshua Stahl, Abdulmohsen Mutairi, Hossein Safavi, Farzad Hessar, Shwan Ashrafi, Yue Yang, Chang Wook Kim, Benjamin Morgan, Shae Hurst, Eldridge Alcantara, Xuhang Ying, Ritochit Chakraborty, Jason Silver, Anthony Smith, Jessica Tran, Nicole Thomas, Tamara Vurin, Amin Jalali, Dennis De Meng, Dan Tidwell, Andrew Lingley, Fredrik Ryden, Allan Ecker, Xiang Zou, and Zixia Hu. Without their contribution and company, my graduate school experience would not have been as nearly as fruitful or enjoyable.

Lastly, I cannot express the amount of gratitude and thanks I owe my parents for their support and encouragement throughout my endeavors. Their words have always carried both hope and inspiration.

DEDICATION

to my parents, Scott and Kathy

Chapter 1

INTRODUCTION

1.1 Background & Motivation

Radio Frequency IDentification (RFID) is a short-range wireless communication technology for reading data, such as an electronic product code (EPC), stored on a RFID *tag* affixed to objects [2]. The vision for RFID is the fast, efficient and automatic reading of numerous tagged items by a RFID *reader*. A typical RFID tag consists of a flexible antenna attached to an integrated circuit (IC), which is typically mounted on a flexible substrate (e.g. an adhesive plastic strip), although in contrast, a RFID reader is capable of sophisticated signal processing and has a dedicated power source. Today, RFID technologies have proliferated into multiple commercial sectors, such as logistic, retail (anti-theft), and automatic road tolling [3]. As an example of how people interact with RFID technology on a daily basis, the tolling system installed on the SR-520 bridge across Lake Washington, WA utilizes RFID tags to identify crossing cars. These applications are enabled by battery-free or *passive* RFID tags and low-cost fabrication of the tag's integrated circuit (IC) [4].

Unlike other wireless transmitters, a class of RFID tags are *entirely passive from an energy point of view*. To read a passive tag, the RFID reader 1) illuminates the tag with a RF signal, 2) sends an interrogation command to tag for it to respond, and 3) listens for the tag's response. A passive RFID tag utilizes the RF signal generated by the RFID reader both as an energy source and for uplink communication from the tag to the reader. A rectifier circuit converts the RF signal impinging on the tag to usable direct current (DC) voltage, which is sufficient to power the tag's IC at close range. For uplink communication, the tag does not generate its own RF signal, but instead modulates the RF signal reflected from its antenna, a method otherwise known as *backscatter modulation*. The concept of using reflected RF signals as a communication link was first published [4] in 1948 by Stockman [5], and in a similar time frame demonstrated in practice by Theremin with the 'Great Seal Bug'

(a Cold-War era espionage device) [6]. The latter is an example of an analog backscatter device that transmits an analog microphone signal, as opposed to modern RFID tags that transmit exclusively digital symbols.

Backscatter modulation operates on a principle akin to RADAR [5], where a typical RADAR scenario involves the transmission of RF signal pulses to locate targets of interest (objects that scatter RF signals). During the tag reading process, the RF reader transmits a continuous wave (CW) at some center frequency ω_c, i.e. $\cos(\omega_c t)$. The tag's antenna scatters back a portion of the reader's CW, and the remaining portion is absorbed by the load attached to the antenna. In a static environment, the RFID tag appears as any other scattering object, however, modulating the antenna's load in turn modulates the scattered CW [7]. The RFID reader now observes a time varying signal superimposed with relatively constant signals (the reader's CW and static scatters), so the tag can encode information onto digital symbols by switching the antenna's load between two or more discrete states.

It is important to emphasize that a backscatter modulator uses none of the active RF circuits found in nearly all transmission circuits, such as a power amplifier or a RF oscillator. The only active component is the switching circuit that modulates the antenna's impedance loading. These impedance elements consist of resistive and/or capacitive/inductive components, which are energized by reader's CW and not by a signal generated on the tag. If the reader does not generate a CW, then a tag cannot initiate any communication with the reader. Additionally, the RFID reader has the added complexity of generating the CW and detecting the tag's response simultaneously. Fig. 1.1 illustrates the architectural differences between an active transceiver and backscatter. This architecture was first developed [4] at Los Alamos National Laboratory in 1975 [8], but unlike previous analog backscatter modulators, it uses *digital modulation* to continuously transmit a three digit ID code when energized by a reader. Its backscatter modulator consists of two switched resistive elements controlled by the ID generation circuit, which generates a simple binary amplitude shift keying (ASK) modulation scheme [8]. Conceptually, this reader and passive tag system architecture has remained virtually unchanged to the modern day [9].

In general, RFID tag types may be classified into one of the following: *active*, *semi-passive*, or *passive*. Active RFID tags utilize powered transceivers for the uplink; see for

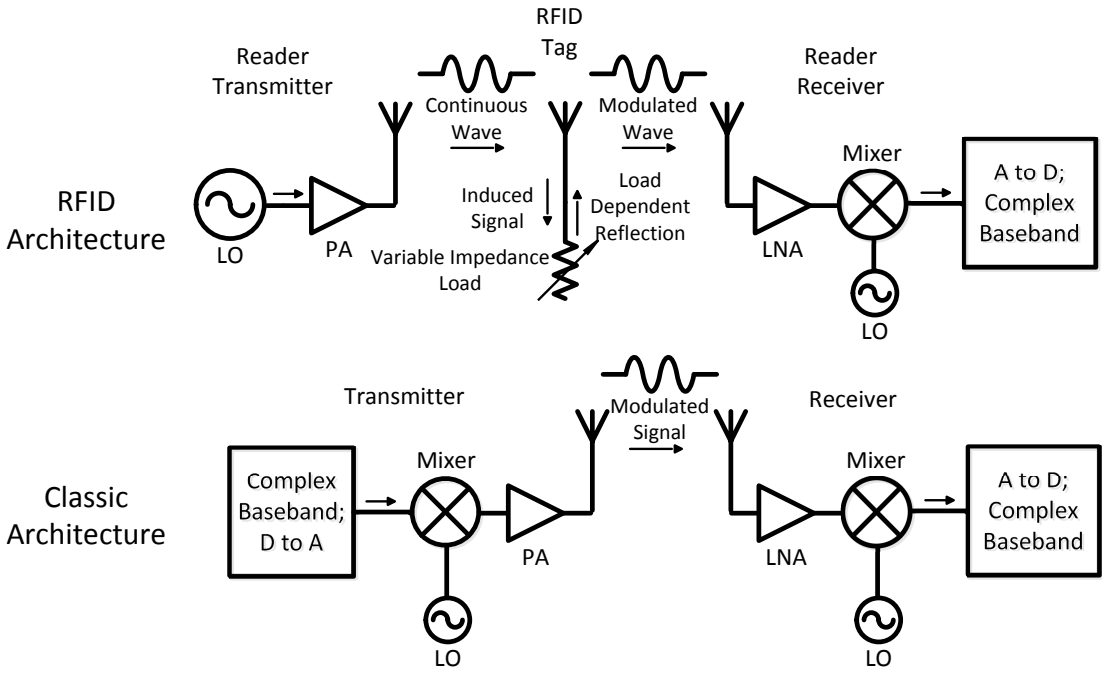


Figure 1.1: Comparison between the active transceiver architecture and the backscatter based transceiver architecture.

reference IEEE 802.15.4f [10] regarding standardization efforts. The standard details two different physical layer designs, either a narrow band minimum shift keying modulation or an ultra-wideband (UWB) architecture. Despite its name, active RFID has more similarity to other low power radios, such as Bluetooth, than backscatter based tags. Semi-passive tags consist of a low cost/power integrated IC capable of limited processing, along with RF circuitry to support backscatter modulation and downlink demodulation, a rectifier to harvest RF energy, and long term energy storage (e.g. battery or capacitor). Passive tags follow the same general architecture as semi-passive tags, except they cannot store energy and are completely reliant on harvested RF energy.

Recently, a new class of passive RFID tags known as *chipless* tags [11, 12] have emerged as an alternative to reduce the cost of RFID tags. Unlike other tag architectures, chipless tags have no computational capability due to their lack of an IC, so they cannot execute the simplest of communication protocols. Lastly, the RFID systems described thus far operate in the far-field region of electromagnetic propagation as opposed to the near-field, and while

near-field based RFID is technical and commercial success, its range is severely limited as opposed to far-field systems [3]. This dissertation only considers RFID systems based on far-field backscatter modulation, i.e. passive and semi-passive tags.

An example architecture of a passive or semi-passive RFID tag is given in Fig 1.2 [13–15]. $\vec{\mathcal{E}}_{\text{inc}}$ represents the incident electrical field (E-field) generated by the reader's CW at the tags antenna, and $\vec{\mathcal{E}}_{\text{scat}}$ represents the backscatter E-field at the reader's receiver. To modulate $\vec{\mathcal{E}}_{\text{scat}}$, the on-tag microcontroller logic (or a digital circuit finite state-machine) switches the backscatter modulator between two impedance states Z_1 and Z_2. Additionally, the microcontroller executes and manages higher level protocol operations. The demodulator (demod.) block decodes the reader to tag communication, and the power harvester captures part of the energy in the incident RF $\vec{\mathcal{E}}_{\text{inc}}$ for use by the tag. Within this example, the interaction between the incident and scatter E-fields is defined by [7]

$$\vec{\mathcal{E}}^{i}_{\text{scat}} = \vec{\mathcal{E}}^{\text{M}}_{\text{scat}} - \Gamma_i I_{\text{M}} \vec{K}_a, \tag{1.1}$$

where $\vec{\mathcal{E}}^{\text{M}}_{\text{scat}}$ is the scattered E-field and I_{M} is the antenna current when the tag impedance impedance is matched to the antenna impedance Z_{ant} and the tag's incident E-field is $\vec{\mathcal{E}}_{\text{inc}}$. The power reflection coefficient is defined to be between the tag's antenna and the antenna's load

$$\Gamma_i = \frac{Z_i - Z^*_{\text{ant}}}{Z_i - Z_{\text{ant}}}, \tag{1.2}$$

which has the property of $|\Gamma_i| \leq 1$ for any impedance value Z_i $(i = 1, 2)$ and Z_{ant} [16]. The remaining term $\vec{K}_a$ is determined by the tag's antenna properties and geometry with respect to the reader's antenna.

Physical properties unique to backscatter modulation differentiate it from the classic wireless transceiver architecture. The analytical expression (1.1) of the reader's received signal provides insight on the fundamental characteristics of backscatter as a communication channel. Namely, the two primary distinctions are the trade-off between energy cost and backscatter modulator design, and the dependence of the reader's received signal $\vec{\mathcal{E}}_{\text{scat}}$ on the tag's incident signal $\vec{\mathcal{E}}_{\text{inc}}$.

Since the RFID tag does not actively generate a RF signal, the energy cost to transmit (backscatter) needs to be reevaluated. Varying the impedance mismatch between the tag

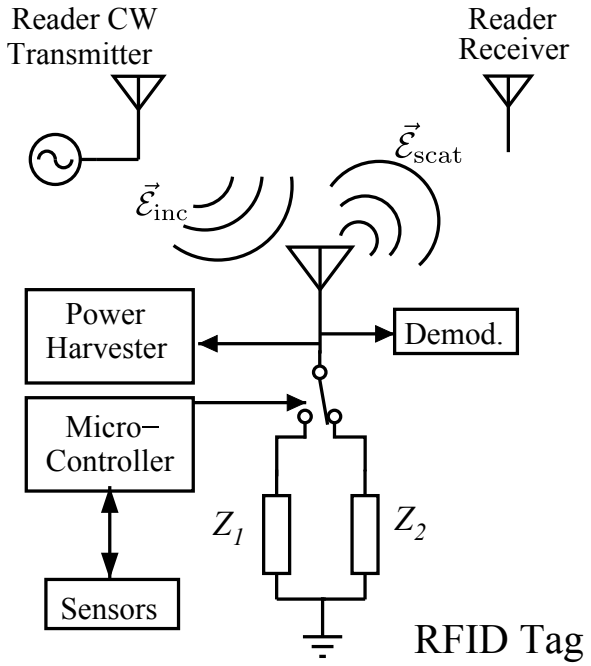


Figure 1.2: System level diagram of an RFID tag with two modulator states and an RFID reader energizing the tag.

antenna and its load is necessary for backscatter modulation, otherwise the tag cannot modulate its E-field signature, i.e. $\Gamma_1 \neq \Gamma_2$. In general, increasing the mismatch between two modulation states (as measured by $|\Gamma_1 - \Gamma_2|$) increases the receive signal to noise ratio (SNR) [7]. A consequence of increasing the impedance mismatch is the decrease in power transfer efficiency from the tag's antenna to the tag's power harvesting circuit [17], thereby reducing the amount of power available to the tag from the reader's CW. The modulation design choice directly impacts the tag's power budget despite the absence of active RF components in backscatter modulators.

The scattered signal $\vec{\mathcal{E}}_{\text{scat}}$ is a function of the impinging signal on the tag's antenna $\vec{\mathcal{E}}_{\text{inc}}$, as quantified in (1.1) by the term I_{M}. So the reader's CW signal, as received by the tag, has a *direct impact* on the backscattered signal received by the reader $\vec{\mathcal{E}}_{\text{scat}}$. Common to most wireless signals, both $\vec{\mathcal{E}}_{\text{inc}}$ and $\vec{\mathcal{E}}_{\text{scat}}$ are subject to small-scale fading caused by environmental multipath effects [18]. The coupling between the tag's backscatter and the reader's CW creates unique fading statistics that differ from the canonical Rayleigh fading

channel [19]. To counter fading's negative effects, [19] proposes the use of multi-antenna tags and readers, otherwise known as multiple-input multiple-output (MIMO).

While there has been a tremendous focus on designing RFID systems for traditional commercial applications (e.g. logistics and road tolling), there is now an expanding interest in using RFID in non-traditional applications. The simplicity and low power consumption of passive or semi-passive RFID tags has made backscatter based communications an attractive candidate for wireless sensor networks (WSNs) [20, 21]. RFID tags are envisioned to act as sensor nodes collecting data, and RFID readers acting as fusion nodes collecting data from the tags. Target applications include scenarios where traditional *active* RF systems cannot achieve required link efficiency vs. power consumption trade-offs, while meeting weight/size constraints [9, 22]. Of particular note are low-power, constrained form-factor scenarios from biological/physiological sensing applications, for which specialized RFID tags with sensors have been fabricated [23–25]. An example of a RFID based sensor platform is the Wireless Integrated Sensing Platform (WISP) [15], which integrates various sensors and a programmable microcontroller onto a custom designed RFID tag [26]. The WISP uses an existing industry developed protocol stack, notably ISO 18000-6C or EPC Class 1 Generation 2 [27] (EPC Gen2), to transport any dynamic sensory data collected by a tag. In part, the WISP's use of an industry standard protocol stack has contributed to its success, because commercially available RFID readers can interact with WISPs, and therefore the further development of EPC Gen2 is crucial for the future backscatter based WSNs.

The EPC Gen2 protocol uses a simple uncoded binary modulation for the physical layer, and the link layer uses frame slotted ALOHA for random access and a stop-and-wait ARQ [28] for data reliability [27]. Compared to other modern communication protocols, the Gen2 physical and link layers are quite rudimentary, which often lead to poor performance [29]. As sensor systems grow in the number of tags, data volume, and range, passive backscatter based networks will require improved link and physical layer efficiencies. Likewise, traditional RFID applications can benefit from these improvements as they will result in higher tag read rates and improved tag reading reliability.

1.2 Contribution & Organization

RFID's widespread commercial success and its adoption by the WSN research community suggests that RFID is a mature technology, however, backscatter modulation exhibits unique characteristics with respect to power efficiency and multi-antenna systems. Furthermore, future WSNs and RFID systems will likely demand improvements to the uplink efficiency. This dissertation focuses on three key areas relating to backscatter modulation and RFID systems: *1) analyze the co-design of backscatter modulation and error correction coding*, *2) characterize space-time coding and MIMO performance limits of the dyadic backscatter channel*, and *3) hybrid-ARQ (HARQ) and random access improvements to the EPC Gen2 protocol*. The contributions to these three areas are described in the sequel.

1.2.1 Coded Modulation for Backscatter

The rate a tag can harvest energy depends on if the tag is idle or backscattering, and during the idle state, the efficiency of the harvesting *circuitry* [13, 30], is the primary determinant. The harvesting circuity and antenna impedance are typically conjugate matched, when in the idle state, in order to maximize the power transfer from the antenna to the tag. During the active state, the tag backscatters by switching its modulator impedance between different states. As consequence, the antenna and power harvester are no longer matched and less power is transfered to the harvester. Clearly the choice of tag modulation and coding influences the amount harvested and affect the link reliability and rate. One fundamental measure of link performance is its spectral and energy efficiency, which is determined by the modulation and error correction code choice. Backscatter modulation has unique trade-offs in terms of modulation design and power harvesting efficiency, so it is necessary to reconsider how to design higher order constellations, such as 4 point quadrature amplitude modulation (4-QAM), in conjunction with error correction codes.

The contributions of Chapter 2 are the following.

- We develop a novel metric to quantify the spectral and energy efficiency of backscatter: the *backscatter power efficiency loss per bit*.

- We analyze the power harvesting efficiency of a general 4-QAM constellation. We develop an optimization framework to find the modulator design that maximizes power transfer to the tag.

- We propose two different error correction schemes for backscatter: equal error protection (EEP) and unequal error protection (UEP). The tag's power harvesting efficiency exhibits asymmetry with respect to a constellation's in-phase and quadrature components. We find that designs based on UEP outperform EEP.

1.2.2 Dyadic Backscatter Channel: Space-Time Coding & MIMO Performance Limits

The performance of MIMO systems is in part determined by the channel's fading statistics. Due to the unique physical properties of backscatter modulation, a RFID MIMO system has a novel fading characterization known as the dyadic backscatter channel (DBC) [19]. The hallmark of the DBC is the cascade of two independent fading terms. The first term is determined by the small-scale fading on the reader-to-tag downlink, and the second term is determined by the small-scale fading on the tag-to-reader uplink, which results in a fading distribution different from the Rayleigh fading channel. So the performance of space-time codes in RFID systems will exhibit new trade-offs, as opposed to the Rayleigh MIMO channel, in regards to system parameters.

The contributions of Chapter 3 are the following.

- A modified but equivalent DBC model is developed to simplify the analysis of space-time codes.

- We examine, for the DBC, two important performance metrics that quantify the reliability of space-time codes in fading channels: pair-wise error probability (PEP) and the maximum diversity order [31]. Unlike Rayleigh fading channels [18], the number of tag antennas solely determines the diversity order of orthogonal space-time block codes. Additional receive antennas at the reader provide a coding gain but no increase in diversity order.

- We verify our analytical results through numerical simulations.

The contributions of Chapter 4 are the following.

- A fundamental measure of the performance in MIMO fading channels is the diversity-multiplexing tradeoff (DMT) [32]. We derive the DMT corresponding to the DBC, which provides additional insight on behavior of the DBC.

- Increasing the number of reader receive antennas beyond the number of tag antennas does not increase the spatial multiplexing gain or diversity order. If the number of reader antennas is greater than or equal to the number of tag antennas, then it is possible to achieve the DMT with space only codes.

- Since the DMT is an asymptotic (in SNR) performance measure, we examine the finite-SNR DMT through numerical simulations.

1.2.3 HARQ & Random Access Enhancements

Two key functions of the EPC Gen2 (or ISO 18000-6C) protocol are the transfer of the EPC data from the tag to the reader and the singulation of RFID tags through random access. EPC Gen2's uplink does not incorporate any forward error correction or fine grained link adaptation; only simple cyclic redundancy checks are used for data integrity. The protocol's modulation scheme is based on binary line codes [33], where the only available variable to adjust the link's reliability and rate is the symbol period. Additionally, the same modulation scheme is fixed for all tags communicating with the reader. Consequently, the current design does not provide a robust uplink during poor link conditions, and it cannot adapt the link to varying conditions. For example, the link quality may change overtime for a specific tag, or the link quality will vary between different tags. We apply hybrid-ARQ (HARQ) methods to the Gen2 protocol in order to improve the link throughput and reliability.

Frame slotted ALOHA (FSA) arbitrates the channel between RFID tags, however, within EPC Gen2 FSA does not support true collision resolution, because any tag participating in a collision simply repeats their attempt in the following ALOHA frame. Transmissions between frames are independent, so any previous collisions do not help future retransmissions.

To improve the efficiency of FSA, we explore the concept of collision resolution through the application of error correction codes and multi-user detection (MUD). The performance of MUD and construction of practical codes are studied from a compressive sensing viewpoint.

The contributions of Chapter 5 are the following.

- We design a hybrid-ARQ protocol that supports fixed and variable length EPC payloads, and it requires only minimal changes to the EPC Gen2 protocol.

- The performance of different HARQ schemes (Chase combining and incremental redundancy) are evaluated through simulations.

The contributions of Chapter 6 are the following.

- We propose a random access protocol named incremental redundancy for frame slotted ALOHA (IR-FSA). Its performance is analyzed using methods derived from compressive sensing.

- Punctured Reed-Muller codes are used as a practical error correction code construction. Their performance is analyzed through simulations and compared against existing research on multi-tag detection.

In summary, this dissertation is organized as follows. Chapter 2 studies the co-design of higher order modulation and error correction codes. Chapter 3 analyzes the pair-wise error probability and diversity order of space-time codes in the dyadic backscatter channel, and Chapter 4 derives the diversity-multiplexing tradeoff of the dyadic backscatter channel. Chapter 5 evaluates the performance of Hybrid ARQ in the context of the EPC Gen2. In Chapter 6, a random access protocol based on incremental redundancy and multiuser detection is proposed and analyzed through simulations and compressive sensing methods. Lastly, Chapter 7 contains the dissertation's concluding remarks.

Chapter 2

CODED MODULATION FOR POWER AND SPECTRAL EFFICIENT RFID

2.1 Introduction

Presently, poor tag sensitivity and power harvesting in passive tags limit the downlink range [17] and achievable link throughput. On the other hand, semi-passive tags that incorporate some energy storage mechanism but still use passive backscatter are typically uplink limited, hence, uplink improvements (enhanced front-ends for power harvesting, circuit designs with lower sensitivity and more energy aware protocol stack) will be key to future RFID systems, as was explored in [17]. In this work, we continue to further explore different facets of the uplink-downlink coupling that relates achievable uplink rate/reliability with downlink power harvesting.

The current industry standard - EPC Global Gen-2 [27] - specifies two binary encoding schemes, FM0 and Miller sub-band coding, for the uplink, implemented with either two state amplitude shift keying (ASK) or phase shift keying (PSK). Both schemes use bi-orthogonal symbols for transmission and incorporate encoder memory in mapping past information bits onto the current transmitted waveform. This suggests the benefits of maximum likelihood sequence estimation (MLSE) at the reader, however, it was shown in [33] that a linear receiver for FM0 operating over two successive symbols achieves near-MLSE performance. The power harvesting efficiency of uncoded binary modulations is analyzed by [17].

It has recently been shown by [34] that it is feasible to build backscatter systems supporting higher order constellations, e.g. 4-QAM or 8-QAM. The work by [34] does not consider coding or power harvesting efficiency, but instead focus on device and circuit design. Higher dimensional constellations opens up the possibility for coded modulation or other advanced modulation and coding combinations for passive backscatter RFID systems, however, improved spectral efficiency implies an energy cost. During nominal (non-backscatter) opera-

tion, the tag IC equivalent impedance is set to allow for maximum power transfer from the antenna. During backscatter modulation, the tag's impedance is intentionally mismatched, and hence the power harvested is reduced (as compared to quiescent state).

Typically, E_b/N_0 or energy per bit to noise power spectral density serves as the standard metric to compare different modulation/coding schemes, because often link design aims to minimize the energy expenditure at the transmitter. Backscatter modulation by itself does not require power to operate, but the tag relies on harvested energy to power its tag IC and in the case of the WISP [15], to power on-tag sensors. Since the modulation choice affects the harvesting efficiency, the tag to reader link efficiency and the tag's power harvesting capabilities are explicitly coupled. We define a novel metric *backscatter power efficiency loss per bit*[1]

$$\mathcal{B}_{loss} = \frac{P_{match} - P_{avg}}{C\ P_{match}}, \tag{2.1}$$

where P_{match} is the power harvested under conjugate matched conditions, P_{avg} is the average power harvested during backscatter, and C is the information bits per symbol. Effectively, $\mathcal{B}_{loss}$ serves as a way to measure the efficiency loss a tag must pay in order to backscatter (communication). Sec. 2.5 will further develop the intuition behind this metric.

Our objective in this chapter is two-fold: to 1) investigate ways to improve uplink spectral efficiency and 2) to characterize the associated backscatter power efficiency loss per bit. Increasing the constellation order from binary signaling to 4-QAM doubles the spectral efficiency, but causes a decrease in P_{avg}. Our work investigates how the co-design of the 4-QAM constellation geometry and forward error correction (FEC) may be used to balance the backscatter power efficiency loss per bit and increased spectral efficiency. We first establish an equivalent circuit model in Sec. 2.2 for the RFID tag in terms of a backscatter modulation design parameter, which determines the trade-off between backscattered signal to the reader and the available power transferred to the tag. In Sec. 2.3 and Sec. 2.4, we explore uncoded 4-QAM constellation design such that the power loss during backscatter is minimized. Next in Sec. 2.5, we investigate FEC encoding schemes suitable for 4-QAM modulation - these include traditional equal error protection (EEP) as well as unequal error

[1]The published work that this chapter is largely derived from [35] uses a different name for this metric, 'normalized power loss per bit per information bit.' We believe this new name is more precise.

protection (UEP) [36] [2]. The latter exploits the advantage of asymmetric power harvesting in a 4-QAM constellation caused by backscatter modulation. Viewed as two independent binary modulations, different levels of protection are offered to the symbol states of a 4-QAM symbol, so as to achieve equal BER post error correction. In Sec. 2.6, we compare the different modulations and codes using a hypothetical link budget. This provides numerical results to quantify the performance of each scheme.

2.2 Passive RFID System Model

Passive and semi-passive RFID tags have desirable form factors with a low-power IC connected to external components such as the antenna. The IC integrates several different analog subsystems such as an AC-to-DC voltage converter (power harvester), impedance match network between the IC and antenna, the (data) modulator/demodulator, and digital subsections comprised of a low power microprocessor [13] [3]. An early tag implementation was demonstrated by [14], and a present-day industry benchmark is the Monza4 chip from Impinj [37].

During backscatter operation the tag IC is the only active component, which modulates the antenna's load impedance, Z_{IC}, via switching a variable impedance component within the tag IC as shown in Fig. 2.1(a). The mismatch between the antenna's impedance R_{ant} and the IC's impedance Z_{IC} decreases the power transfer from antenna to tag. We next introduce a simple equivalent circuit model to quantify backscatter modulation and power delivered to the tag IC.

2.2.1 Equivalent Circuit Model

A Thevenin equivalent circuit model of the tag IC and antenna is shown in Fig. 2.1(a) [7]. It consists of a voltage source V_{OC}, with the same frequency as the reader's RF carrier ω, in series with the antenna represented by complex impedance $Z_{ant} = R_{ant} + j\omega L_{ant}$; the tag

[2] UEP codes protect bits at different levels, whereby both bits are not equally significant in terms of their intrinsic information.

[3] In some designs that target ultra-low power consumption, a finite state machine is used instead of a microprocessor.

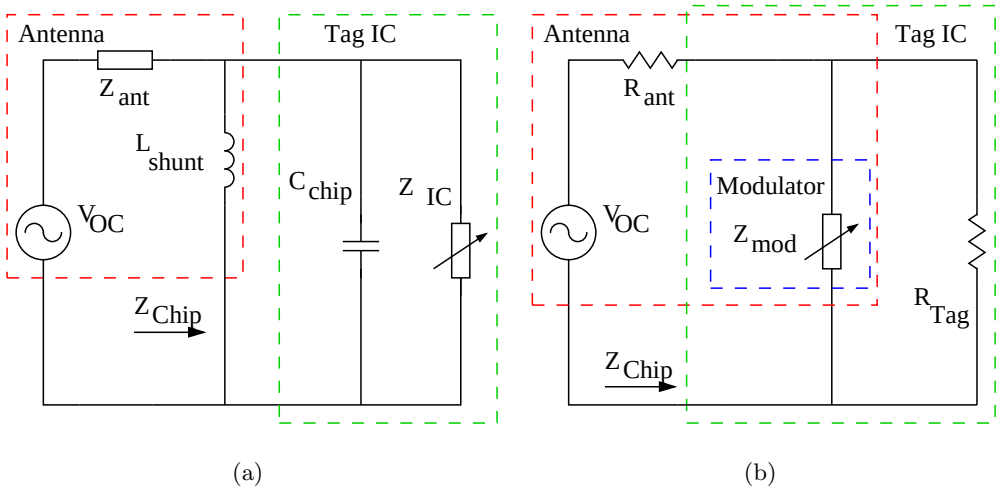


Figure 2.1: Tag circuit model before the matching network: (a) illustrates the model antenna inductance and chip capacitance. Whereas (b) has a matching network with the variable components of Z_{IC}, C_{chip}, and L_{shunt} composing Z_{mod}.

IC connects to this circuit via a matching network. V_{OC} is the antenna voltage generated by the reader's electric field at the tag. The RFID IC chip impedance is assumed to have a capacitive component in parallel, i.e., $Z_{chip} = Z_{IC} \parallel \frac{1}{j\omega C_{chip}} \parallel j\omega L_{shunt}$, and it is assumed there is a shunt inductance L_{shunt} from the antenna across the tag's IC [13]. The impedance matching network between the chip and antenna is designed such that L_{ant} term cancels at ω of choice, e.g. 915MHz, which leads to a simplified antenna equivalent circuit of R_{ant} [13].

Define Z_{IC} to be composed of two subcomponents Z_{mod} and R_{Tag} in parallel. R_{Tag} models the input impedance of all tag IC components other than the modulator circuit, such as the demodulator and power conversion circuits. Thus the power delivered to R_{Tag} is the power available for use on the tag. The value of R_{Tag} is chosen to equal R_{ant} so that during a non-backscatter state, the maximum amount of energy transfered from the antenna to the tag is given by

$$P_{match} = \frac{|V_{OC}|^2}{8R_{ant}}. \tag{2.2}$$

Since R_{Tag} is a static value and Z_{mod} is a variable impedance controlled by the tag IC logic, any power delivered to Z_{mod} cannot be utilized by the IC and is considered wasted energy.

While modeling a RFID tag as a static lumped circuit greatly simplifies analysis, this approaches ignores certain frequency and power dependent effects. First, the impedance of Z_{mod} varies with respect to the reader's carrier frequency ω, because it consists of frequency dependent components (capacitors and inductors). For RFID systems operating in the UHF ISM band within the US, a typical design assumes an operating frequency of 915 MHZ.[4] Prior work addresses this issue through RF and device level simulations, which shows negligible performance differences in both BER and error vector magnitude (EVM) due to frequency variations within UHF ISM frequency operating range [34].

Second, the tag's power harvesting circuitry exhibits non-linear properties with respect to incident RF power, which effect the circuit's AC to DC conversion efficiency [30] and equivalent input impedance [38, 39]. As further discussed in Sec 2.4, the optimality of the modulation/coding designs proposed within this chapter depend on the assumption of a fixed AC-to-DC conversion efficiency. Although some harvester circuit designs [30] do follow a nearly linear relationship between input power and rectified power once the input power reaches a critical threshold. Furthermore, varying either the operating frequency or input power affects the harvester's equivalent impedance [38, 39], so a lumped R_{Tag} model fails to capture second order effects. The discussion in Sec. 2.2.3 on the design of high order backscatter modulator circuits assumes a fixed R_{Tag} model, and any changes in R_{Tag} will distort the constellation geometry.

In the sequel, we assume a fixed harvesting efficiency and a constant R_{Tag} with respect to operating frequency and input power. Future work, which is further expanded upon in Chapter 7, should address the variability of R_{Tag} to these two parameters.

2.2.2 Choosing Z_{mod}

A variable impedance $Z_{mod} = R_{mod} + jX_{mod}$ consists of a variable capacitor (bank of varactors), chip capacitance, variable resistor (switched), and a shunt inductor in parallel. This modulator architecture is a variation on the PSK modulator proposed by [13]. Variable components are controlled by the on-tag logic, which can be either a micro-controller or a

[4]The EU uses a separate frequency band of $865 - 868$ MHz.

hardwired state machine. Each set of resistive and capacitive values map to a specific Z_{mod} value given by

$$Z_{mod} = R_{mod} \parallel j\omega L_{shunt} \parallel \frac{1}{j\omega(C_{mod} + C_{chip})}. \tag{2.3}$$

The values of C_{mod} are chosen to be greater and less than the design parameter C_{center}, such that C_{chip} and L_{shunt} are tuned out at the frequency of operation when $C_{mod} = C_{center}$, i.e. $L_{shunt} \parallel (C_{center} + C_{chip}) = j\infty$ [13]. Capacitance values greater (less) than C_{center} generate positive (negative) reactances, respectively.

In general, the combination of L_{shunt} and $C_{mod} + C_{chip}$ is a purely reactive component jX_{mod}, hence, Z_{mod} behaves a variable impedance with positive and negative reactance and positive resistance. So for any target Z_{mod}, the corresponding R_{mod} and X_{mod} values are mapped to a circuit level design.

2.2.3 Backscatter Modulation

During backscatter operation the reader transmits a constant carrier wave that illuminates the tag's antenna with an electro-magnetic (EM) field. Corresponding to the physical properties and impedance loading of the tag, a portion of the impinging EM wave is backscattered from the tag. The backscattered wave comprises of two components: a) structural mode and b) antenna mode scattering. Structural mode scattering depends on the physical properties of the antenna, and antenna mode scattering depends on the antenna's impedance load. The physical configuration of the antenna remains independent of any impedance load placed on the antenna [40]. It has been shown by [7] that the net electric field scattered from a tag is

$$\vec{\mathcal{E}}^{i}_{scat} = \vec{\mathcal{E}}_{struct} - \Gamma_i \; I_{match} \; \vec{K}_a \tag{2.4}$$

where

- $\Gamma_i = \frac{Z_{IC} - R_{ant}}{Z_{IC} + R_{ant}}$ is the power reflection coefficient for the i-th modulator state;
- $\vec{\mathcal{E}}_{struct}$ is the structural scattering term;
- $I_{match} = \frac{V_{OC}}{2R_{ant}}$ is the current through R_{ant} when the impedance of the chip and antenna are conjugate matched;

- $\vec{K}_a$ is the Volts per meter per Ampere; the electric field strength radiated by the antenna per unit Ampere of antenna current.

In (2.4), Γ_i is the only term the tag can modulate with its impedance; the remaining terms are effectively constants. Furthermore, the work of [7] proves that the effect of structural scattering equates to adding a constant offset to a modulated signal. A common approach by the literature assumes $\vec{\mathcal{E}}_{struct} = I_{match}\ \vec{K}_a$ [7, 17], thereby making $\vec{\mathcal{E}}_{scat} \propto I_{ant}$; I_{ant} as the current through the antenna.

For the subsequent discussion, we assume an additive white Gaussian noise (AWGN) model and perfect knowledge at the receiver of the constant offset, phase rotation, and symbol synchronization. [5] The net received signal at the reader $z_i \in \mathbb{C}$ is the sum of three components: a) the modulated antenna mode scattering x_i, b) L as the constant offset from transmit to receive leakage and structual scattering and c) WGN, i.e.,

$$z_i = h\ x_i + L + w \tag{2.5}$$

where $w \sim \mathcal{CN}(0, N_0)$ and h is a random phase rotation.

Consider a 2-QAM baseband constellation with symbols $x_0\ ,\ x_1 \in \mathbb{C}$ observed by the readers receiver, as modeled by (2.5). Illustrated by Fig. 2.2, the constellation has a Euclidean distance of $d = |x_1 - x_0|$ yielding an energy per symbol $E = \left(\frac{|x_1 - x_0|}{2}\right)^2$. Hence, the corresponding bit error probability over an equivalent lowpass AWGN channel with noise power spectral density of $\frac{N_0}{2}$ is [6]

$$p(e) = \mathcal{Q}\left(\sqrt{\frac{2E}{N_0}}\right) \tag{2.6}$$

We emphasize that in error analysis of digital communication methods such as the above, it is convenient to shift the constellation by an offset equal to the structural mode scattering and transmitter leakage, so as to center them around the origin, since it has no impact on the resulting error probability. Although, this requires a the reader with a monostatic[7]

[5] In practice, this is done through parameter estimation on a known preamble sequence.

[6] Where $\mathcal{Q}(x) = 1 - \Phi(x)$ and $\Phi(x)$ is the CDF of the standard normal distribution.

[7] Monostatic readers use the same local oscillator (LO) for the transmit and receive chains, and bistatic readers use a seperate LOs. If the LOs are not common, then there is a frequency offset observed at the receiver. [41]

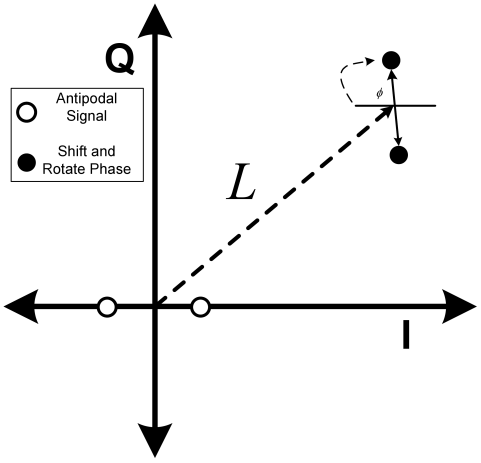


Figure 2.2: A binary antipodal constellation where $x_i = \pm 1$, L is a constant offset and $h = e^{-j\phi}$

architecture and that the channel state remains coherent over multiple symbol periods. Otherwise L and h cannot be shifted and de-rotated, so the receiver must use non-coherent detection techniques [41, 42]. For short range and low mobility operation with symbol rates on the order of 100s of Kbaud, it is reasonable to assume a coherent receiver is possible at the reader. We defer to [43] for additional discussion on coherent detection receivers for backscatter modulation.

To relate the value of (2.6) to a set of modulator impedances, consider that $|x_1 - x_0| \propto |\vec{\mathcal{E}}^1_{scat} - \vec{\mathcal{E}}^0_{scat}|$, and $|\vec{\mathcal{E}}^1_{scat} - \vec{\mathcal{E}}^0_{scat}| \propto |I^1_{ant} - I^0_{ant}|$. Thus choosing different I_{ant} via tag impedance modulation is sufficient to design a backscatter constellation. Under the model described by Fig. 2.1(b), the tag antenna current is given by

$$I_{ant} = \frac{V_{OC}}{Z_{total}} = \frac{V_{OC}}{R_{ant} + Z_{IC}} = \frac{V_{OC}}{R_{ant} + \frac{R_{Tag} Z_{mod}}{R_{Tag} + Z_{mod}}}. \tag{2.7}$$

Substituting $R_{Tag} = R_{ant}$, and replacing $Z_{mod} = \beta R_{ant}$, where $\beta \in \mathbb{C}$ is the variable design factor, yields

$$\begin{aligned} I_{ant} &= \frac{V_{OC}}{R_{ant} + \frac{R_{ant} Z_{mod}}{R_{ant} + Z_{mod}}} = \frac{V_{OC}(R_{ant} + Z_{mod})}{R_{ant}(R_{ant} + Z_{mod}) + R_{ant} Z_{mod}} \\ &= \frac{V_{OC}}{R_{ant}} \left(\frac{1+\beta}{1+2\beta} \right). \end{aligned} \tag{2.8}$$

Denoting backscatter coefficient

$$\eta_i = \frac{1+\beta_i}{1+2\beta_i}, \tag{2.9}$$

where i is the modulator state index, we get

$$I_{ant,i} = \frac{V_{OC}}{R_{ant}} \eta_i. \tag{2.10}$$

The relation of the tag antenna current and its induced voltage observed by the reader is further discussed in Appendix A.

Clearly, the choice of backscatter symbols η_i determines the BER. Any choice of $\eta \in \mathbb{C}$ is not feasible, because it must map to a valid modulator impedance. A modulator impedance is realizable if the resistive component is non-negative, i.e., $\mathrm{Re}(\beta) \geq 0$. Since

$$\beta = \frac{1-\eta}{2\eta - 1}, \tag{2.11}$$

this implies

$$\mathrm{Re}\left(\frac{1-\eta}{2\eta - 1} \right) \geq 0. \tag{2.12}$$

Which is simplified by the following

$$\begin{aligned} \mathrm{Re}((1-\eta)(2\eta^* - 1)) &\geq 0 \\ |\eta|^2 - 1.5\mathrm{Re}(\eta) + 0.5 &\leq 0 \\ |\eta|^2 - 1.5\mathrm{Re}(\eta) + 0.75^2 &\leq 0.25^2 \\ (\eta - 0.75)(\eta - 0.75)^* &\leq 0.25^2, \end{aligned}$$

where the first line multiplies out the denominator, the second line simplifies so the real operation affects only one term and divides out the leading coefficient, the third line uses

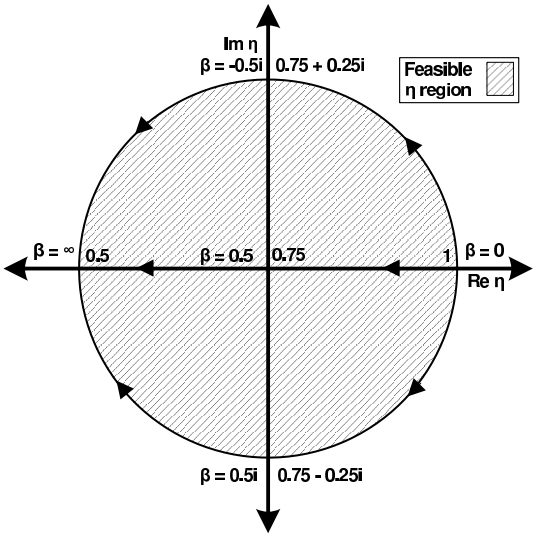


Figure 2.3: Visualization of β and $\eta \in W$, (2.11).

the fact $0.5 = 0.75^2 - 0.25^2$, and the final line is the factored form. Hence, the feasible set for η:

$$W = \{\eta : \ |\eta - 0.75| \leq 0.25\} . \tag{2.13}$$

Fig. 2.3 illustrates the region of feasible backscatter symbols in relation to modulator impedance.

In general the design of a backscatter constellation consists of the following steps.

1. Find a set of $\eta_i \in W$ such the constellation geometry achieves the specified bit error rate vs. tag-reader separation distance and possesses other desirable properties such as symmetry;

2. From (2.11), solve for β_i, and determine the corresponding impedance values for circuit level design.

2.3 4-QAM Constellation Design

It is well-known from communication theory that increasing the number of constellation points improves the bit rate but also requires an increase in SNR (or equivalently, near-

est neighbor symbol separation) at the receiver to preserve BER performance. While the proposition of higher order QAM constellations for the RFID uplink is not new [34], we investigate further the issue of QAM constellations from the perspective of power loss due to the coupling of uplink BER performance and power harvesting by the tag. The dependence warrants a new joint design optimization approach, where we describe the preliminaries in this section and formalize in the following section. We investigate a 4-QAM modulation scheme consisting of two independent antipodal encodings on the orthogonal (I) and quadrature (Q) components, respectively, as a case study.

2.3.1 Constellation Geometry and Bit Mapping

4-QAM encodes two bits - labeled as the pair (A, B) onto a constellation symbol. We represent a general 4-QAM constellation as four symbols on a circle of radius r centered at $y \in \mathbb{C}$, i.e.,

$$\eta_i = re^{j(\theta_i - \psi)} + y \quad \text{such that } \eta_i \in W, i \in \{0, 1, 2, 3\} \tag{2.14}$$

where θ_i and r are considered fixed and ψ represents a rotation for *all* symbols. By imposing symmetry, the symbol points are related to each other by

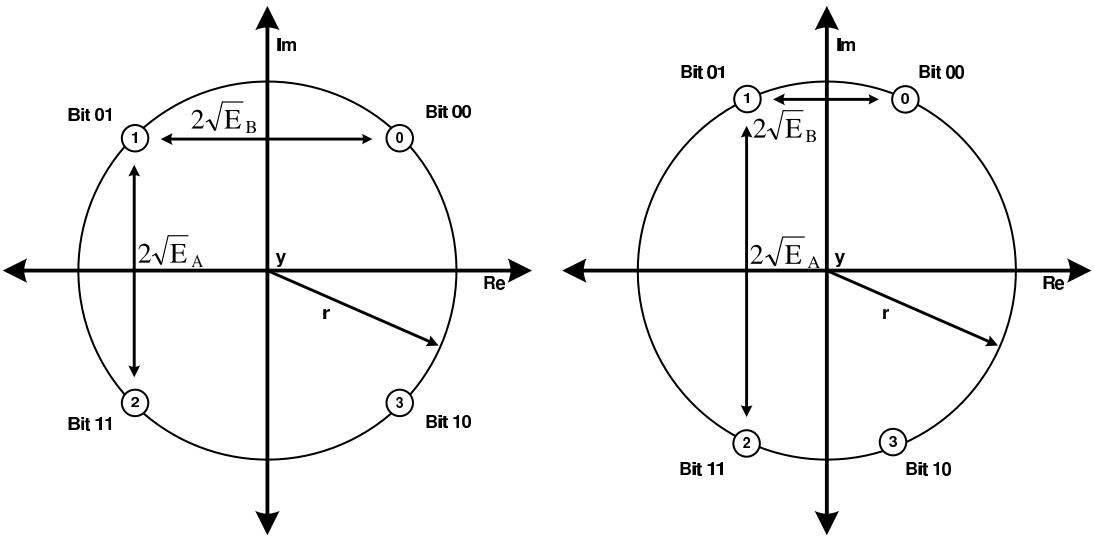
$$\begin{aligned} \theta_0 + \pi &= \theta_2 \\ \theta_1 + \pi &= \theta_3. \end{aligned} \tag{2.15}$$

From the geometry of a circle, the following is true

$$\begin{aligned} |\eta_0 - \eta_3| &= |\eta_1 - \eta_2| \\ |\eta_0 - \eta_1| &= |\eta_2 - \eta_3|. \end{aligned} \tag{2.16}$$

Assuming Gray code labeling, as per Fig. 2.4, the constellation points at the reader post demodulation are assumed to be $\{(\sqrt{E_B}, \sqrt{E_A}), (-\sqrt{E_B}, \sqrt{E_A}), (\sqrt{E_B}, -\sqrt{E_A}), (-\sqrt{E_B}, -\sqrt{E_A})\}$, where

$$\begin{aligned} E_A &\propto |\eta_0 - \eta_3|^2 = |\eta_1 - \eta_2|^2 \\ E_B &\propto |\eta_0 - \eta_1|^2 = |\eta_2 - \eta_3|^2. \end{aligned} \tag{2.17}$$



(a) 4-QAM constellation with $E_A = E_B$ (b) 4-QAM constellation with $E_A > E_B$

Figure 2.4: Example 4-QAM constellations described by (2.14) and (2.15)

The values of E_A and E_B directly map to the parameters r and θ_i.

The BER analysis follows that for quadrature phase shift keying (QPSK), where it is assumed that the AWGN in the I and Q components are independent and identically distributed. The maximum likelihood (ML) decision rule chooses the constellation point nearest to the received signal. It is straight forward to show that the optimal rule for decoding bit A and bit B depends only on $\mathrm{Im}(z_i)$ and $\mathrm{Re}(z_i)$, respectively, where z_i represents the complex sample post demodulation from (2.5). The corresponding probability of error can be written in terms of the respective symbol energy, E_A and E_B

$$\begin{aligned} p_A(e) &= \mathcal{Q}\left(\sqrt{\frac{2E_A}{N_0}}\right) \\ p_B(e) &= \mathcal{Q}\left(\sqrt{\frac{2E_B}{N_0}}\right) \end{aligned} \tag{2.18}$$

Fig. 2.4 illustrates a realization of bit mapping for two cases: $E_A = E_B$ and $E_A > E_B$.

2.3.2 Tag Power and Backscatter Modulation

During backscatter, the tag changes its IC impedance Z_{mod} to mismatched state, thereby lowering the power transfer from the antenna (P_{Tag}). Energy lost due to the impedance mismatch is captured by the effective cost, defined as the ratio of energy lost to maximum harvested energy: $\frac{P_{match}-P_{Tag}}{P_{match}}$. The power transfer between a source and a load is maximized if the source and load impedances are conjugate matched, i.e.,

$$Z_{source} = Z^*_{load}. \tag{2.19}$$

Under the equivalent circuit model in Fig. 2.1(b), the conjugate impedance match condition is met if and only if

$$Z_{IC} = R_{Tag} = R_{ant}. \tag{2.20}$$

The power available to the tag is the power available in R_{Tag}, i.e.,

$$P_{Tag} = \frac{1}{2}|I_{Tag}|^2 R_{Tag}. \tag{2.21}$$

The parallel configuration of Z_{mod} and R_{Tag} is equivalent to a current divider, and hence

$$I_{Tag} = I_{ant}\frac{Z_{mod}}{R_{Tag}+Z_{mod}} = \frac{V_{OC}}{R_{ant}+Z_{IC}}\frac{Z_{mod}}{R_{Tag}+Z_{mod}}. \tag{2.22}$$

Rearranging,

$$\frac{V_{OC}}{R_{ant}+Z_{IC}} = \frac{V_{OC}}{R+\frac{Z_{mod}R_{Tag}}{R_{Tag}+Z_{mod}}} = \frac{V_{OC}(R_{Tag}+Z_{mod})}{R_{ant}(R_{Tag}+Z_{mod})+Z_{mod}R_{Tag}}. \tag{2.23}$$

Substituting back into the (2.22) and simplifying

$$\begin{aligned} I_{Tag} &= \frac{V_{OC}(R_{Tag}+Z_{mod})}{R_{ant}(R_{Tag}+Z_{mod})+Z_{mod}R_{Tag}}\frac{Z_{mod}}{R_{Tag}+Z_{mod}} \\ &= \frac{V_{OC}Z_{mod}}{R_{ant}(R_{Tag}+Z_{mod})+Z_{mod}R_{Tag}}. \end{aligned} \tag{2.24}$$

Set $R_{Tag} = R_{ant}$ and $Z_{mod} = \beta R_{ant}$,

$$I_{Tag} = V_{OC}\frac{\beta R_{ant}}{R^2_{ant}+2\beta R^2_{ant}} = \frac{V_{OC}}{R_{ant}}\frac{\beta}{1+2\beta}. \tag{2.25}$$

Finally, the power transferred to the tag IC as a function of β,

$$P_{Tag}(\beta) = \frac{|V_{OC}|^2}{2R_{ant}}\frac{|\beta|^2}{|1+2\beta|^2} = 4P_{match}\frac{|\beta|^2}{|1+2\beta|^2}. \tag{2.26}$$

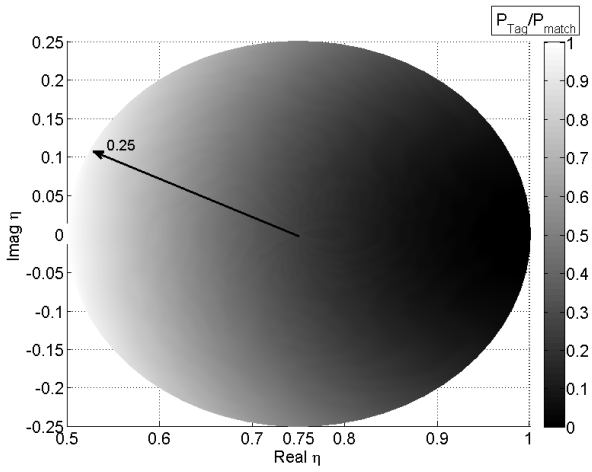


Figure 2.5: Graphical visualization of $\eta \in \mathbb{C}$ versus $\frac{P_{Tag}}{P_{match}}$ in gray scale, (2.27).

Substituting for β in terms of η yields

$$\begin{aligned} P_{Tag} &= P_{Tag}\left(\frac{1-\eta}{2\eta-1}\right) = 4P_{match}\left|\frac{\frac{1-\eta}{2\eta-1}}{1+2\frac{1-\eta}{2\eta-1}}\right|^2 \\ &= 4P_{match}\left|\frac{\frac{1-\eta}{2\eta-1}}{\frac{2\eta-1}{2\eta-1}+2\frac{1-\eta}{2\eta-1}}\right|^2 = 4P_{match}\left|\frac{\frac{1-\eta}{2\eta-1}}{\frac{1}{2\eta-1}}\right|^2 \\ &= 4P_{match}|1-\eta|^2. \end{aligned} \tag{2.27}$$

Fig. 2.5 plots the ratio $\frac{P_{Tag}}{P_{match}}$ as a function of $\eta \in \mathbb{C}$. η values near the feasible region boundary, such as 0.5 (impedance match condition), tend to maximize the power transfer. Intuitively, symbols should be placed as close as possible to the P_{match} ($\eta = 0.5$) state in order to maximize average power transfer. Next, we propose an optimization approach for the choice of η_i for a specific constellation such that power transfer is maximized.

2.4 Power Loss Minimization

Minimizing $\mathcal{B}_{loss}$ (2.1) is equivalent to maximizing P_{avg}, which is the mean P_{Tag} over the modulator states $\{\eta_0, \eta_1, \eta_2, \eta_3\}$ [8],

$$\begin{aligned} P_{avg}(\eta_0, \eta_1, \eta_2, \eta_3) &= \frac{1}{4}\ \left(P_{Tag}(\eta_0) + P_{Tag}(\eta_1) + P_{Tag}(\eta_2) + P_{Tag}(\eta_3)\right) \\ &= P_{match}\ \left(|1-\eta_0|^2 + |1-\eta_1|^2 + |1-\eta_2|^2 + |1-\eta_3|^2\right) \end{aligned} \tag{2.28}$$

We seek the set $\{\eta_i\}$ such that P_{avg} is maximized subject to the following constraints:

- the achievable BER is no greater than a predetermined threshold value, representing reliable link layer operation;
- the set $\{\eta_i\}$ satisfy the geometry described by (2.14) and (2.15);
- the set $\{\eta_i\}$ map to feasible modulator impedances (2.13).

The reliable link requirement determines the parameters $\{r, \theta_i\}$ of (2.14) to values defined as $\{\hat{r}, \hat{\theta}_i\}$. In this section, we solve for the remaining parameters $\{y, \psi\}$ such that P_{avg} is maximized subject to the constraints described above.

Substituting for η_i in terms of $\{\hat{r}, \hat{\theta}_i, y, \psi\}$, the objective function P_{avg} further simplifies due to the symmetry of 4-QAM; $P_{avg} = P_{match}(\sum_{i=0}^{3} |1 - \hat{r}e^{j(\hat{\theta}_0 - \psi)} - y|^2)$. From (2.15), the points η_0 and η_2 satisfy $\hat{r}e^{j(\hat{\theta}_2 - \psi)} + y = -\hat{r}e^{j(\hat{\theta}_0 - \psi)} + y$; likewise for η_1, η_3. Now, each term in P_{avg} can be written as

$$|1 - \hat{r}e^{j(\hat{\theta}_i - \psi)} - y|^2 = |1-y|^2 - 2\mathrm{Re}\{(1-y^*)\hat{r}e^{j(\hat{\theta}_i - \psi)}\} + \hat{r}^2 \tag{2.29}$$

Using the fact that $\hat{r}e^{j(\hat{\theta}_2 - \psi)} = -\hat{r}e^{j(\hat{\theta}_0 - \psi)}$, cancels out the Re{.} terms, leading to

$$|1 - \hat{r}e^{j(\hat{\theta}_0 - \psi)} - y|^2 + |1 - \hat{r}e^{j(\hat{\theta}_2 - \psi)} - y|^2 = 2(|1-y|^2 + \hat{r}^2). \tag{2.30}$$

and resulting in

$$P_{avg} = 4P_{match}(|1-y|^2 + \hat{r}^2). \tag{2.31}$$

[8] We assume that the symbols are equiprobable.

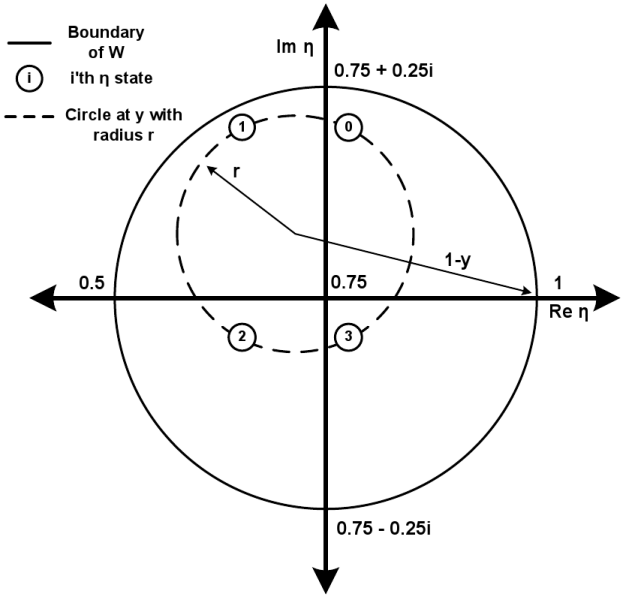


Figure 2.6: To maximize P_{avg} choose y and ψ (rotation of the η_i points) such that $|1-y|$ is maximized, and η_i remain confined to W (2.13).

Hence ignoring scalar multiplier and additive constants, the optimization problem seeks to

$$\begin{aligned} &\text{Maximize}_{y,\psi} \;\; |1-y|^2 \\ &\text{Subject to: } |\hat{r}e^{j(\hat{\theta}_i-\psi)} + y - 0.75| \leq 0.25 \;,\; 0 \leq i \leq 3 \\ &\qquad 0 \leq \psi \leq 2\pi \;,\; y \in \mathbb{C} \end{aligned} \tag{2.32}$$

The first set of constraints ensure that the η_i found by the optimization are feasible (2.13). Geometrically interpreted, the η_i correspond to four corners of a rectangle centered at y, which is confined to the region of feasible η_i; see Fig 2.6. The choice of ψ rotates the constellation points around the center y.

The solution to the optimal $\{y, \psi\}$, follows the geometric proof outlined by:

- the optimal solution ($y^\star$) occurs when at least two symbols, η_i, lay on the boundary of W (2.13);

- the set of feasible y can be reduced to $|y - 0.75| \leq d$ where d depends on E_A and E_B;
- for a constellation with $E_A > E_B$, $y^\star = 0.75 - \sqrt{0.25^2 - E_A} + \sqrt{E_B}$.

As described in Section 2.3 the choice of $\{\hat{r}, \hat{\theta}_i\}$ map directly to the values of E_A and E_B.

For any fixed ψ, the optimization problem described by (2.32) is the maximization of a convex function over a closed convex set [44]. Convex analysis [45] states that the maximum occurs somewhere on the boundary of the constrain region W, i.e. at least one constraint is met with equality. Since the value of ψ does not affect the link performance, it is necessary to consider all feasible $\{y, \psi\}$ pairs that meet the boundary conditions.

From the rectangular geometry of the four η_i points and the circle shape of constraint region W, it is possible to choose $\{y, \psi\}$ such that two symbols lie on the boundary of W; Fig 2.6 illustrates the relationship between η_i and W. The symmetry of W reduces the set of all feasible y (for some ψ) to a circle defined by $|y - 0.75| \leq d$. Following the argument given in Appendix B, d is given as

$$\{y : \ |y - 0.75| \leq \sqrt{0.25^2 - E_A} - \sqrt{E_B}, \ E_A > E_B\} \tag{2.33}$$

or

$$\{y : \ |y - 0.75| \leq \sqrt{0.25^2 - E_B} - \sqrt{E_A}, \ E_A \leq E_B\}. \tag{2.34}$$

Therefore the solution that maximizes $|1 - y|$ clearly occurs at the boundary of W where

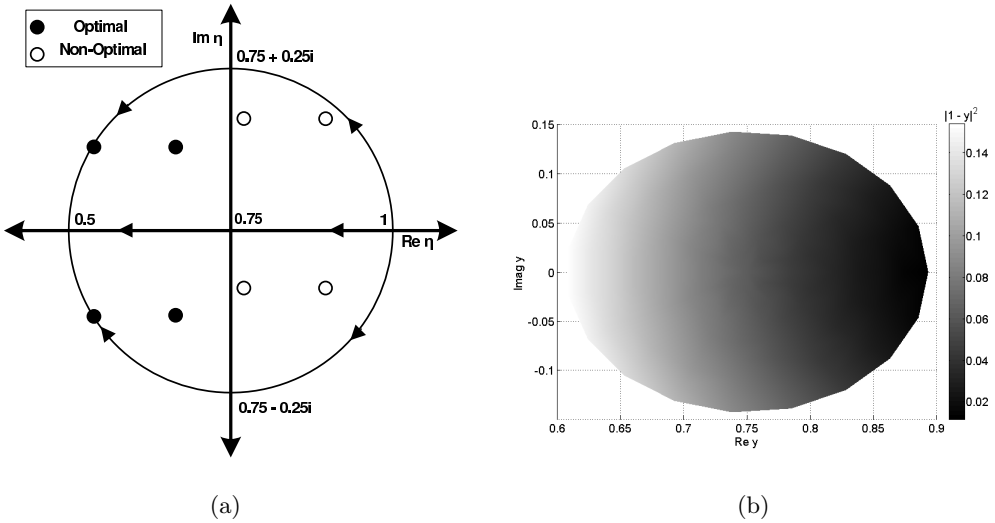
$$y^\star = 0.75 - \sqrt{0.25^2 - E_A} + \sqrt{E_B}, \ E_A > E_B \tag{2.35}$$

or

$$y^\star = 0.75 - \sqrt{0.25^2 - E_B} + \sqrt{E_A}, \ E_A \leq E_B. \tag{2.36}$$

Geometrically, the optimal constellation occurs at the far left side of W.

Fig. 2.7(a) illustrates an example of two 4-QAM constellations (optimal and sub-optimal) where the values for E_A and E_B are chosen arbitrarily. The backscatter power efficiency loss $\frac{P_{avg} - P_{match}}{P_{match}} = 0.27$ for the optimized case and $\frac{P_{match} - P_{avg}}{P_{match}} = 0.84$ for the non-optimized example. Fig. 2.7(b) plots the set of all feasible y vs $|1 - y|^2$, which shows that the optimal y occurs at the point closest to the matched state. However the optimal η_i does not include the matched state (which maximizes P_{Tag} for that state) as one of the modulator states.



(a) (b)

Figure 2.7: Figure (a) Example of backscatter symbol placement for a 4-QAM constellation geometry ($r = 0.15$,$\theta_0 = \frac{4\pi}{11}$,$\theta_1 = \frac{7\pi}{11}$,$\psi = 0$): optimal and non-optimal. Figure (b) plots y (2.14) vs. $|1-y|^2$ for the same constellation parameters as (a).

2.5 Coded Modulation Design

Increasing the modulation order improves spectral efficiency but decreases the power harvested P_{avg}. To fairly compare different modulations and FEC codes, we use the backscatter power efficiency loss per bit (2.1)

$$\mathcal{B}_{loss} = \frac{P_{match} - P_{avg}}{C\ P_{match}},$$

which is a strictly positive number. This metric arises from the notion that energy loss occurs when the modulator generates an impedance mismatch between the harvester circuitry and antenna. As consequence, the tag loses the energy it could have harvested if no backscatter occurred at all.

More formally, $\mathcal{B}_{loss}$ can be derived from first principles. The energy per bit lost due to backscatter modulation is

$$\frac{P_{match} - P_{avg}}{C/T_s}, \tag{2.37}$$

where T_s is the symbol length, hence, C/T_s represents the effective bit rate. If the tag did not backscatter the tag would have harvester $P_{match}T_s$ amount of energy, so by normalizing

(2.37) with this term results in our definition of $\mathcal{B}_{loss}$

$$\mathcal{B}_{loss} = \frac{P_{match} - P_{avg}}{C/T_s} \frac{1}{P_{match} T_s} = \frac{P_{match} - P_{avg}}{C\ P_{match}}. \tag{2.38}$$

The motivation for normalizing (2.37) stems from the fact that both P_{match} and P_{avg} vary based on the reader's transmit power and the parameters of the tag's antenna. Two systems with identical modulation and coding designs will exhibit different energy loss per bit values if the reader transmit power differs between them, which makes it difficult to compare different modulation and coding designs. If this term is normalized by P_{match}, then both will have identical $\mathcal{B}_{loss}$ values.

Additionally, we emphasize that lower $\mathcal{B}_{loss}$ only implies improved power harvesting during backscatter. So for a typical passive RFID tag reading scenario, where a tag spends most of its time idle and only transmits a short EPC, an improved $\mathcal{B}_{loss}$ has a smaller impact as opposed to a sensor network application. A sensor network moves a larger volume of information from each sensor, thereby increasing the system impact on the cost to send each bit. This section will explore the use of FEC coding to improve (lower) $\mathcal{B}_{loss}$ required for 4-QAM as compared to the uncoded case.

2.5.1 Uncoded 2-QAM and 4-QAM Modulation

The uncoded 2-QAM constellation follows directly from the modulations described in the Gen2 standard [27]. Two common binary modulator designs in current use - (binary) ASK and PSK - achieve the same BER given by (2.6), however, ASK uses only resistive modulator impedances while PSK uses only reactive modulator impedances. Even though both have identical link performance, PSK achieves a lower $\mathcal{B}_{loss}$ than ASK. The power efficiency difference between ASK and PSK has been recognized previously by [13] but not yet fully explored.

The 4-QAM constellation is designed by the method discussed in Sec. 2.3.2. Both bits have equal probability of error, which leads to a constellation with equal distance between symbols as illustrated by Fig. 2.4(a) ($E_A = E_B$) and BER given by (2.18).

2.5.2 Coded 4-QAM: UEP and EEP

The proposed 4-QAM constellation along with a suitable forward error correction code promises new options to balance desired uplink rates and power harvesting. The coding gain allows for reliable bit decoding at a lower SNR, allowing for a new modulator design with different impedances. Lower SNR requirements maps to constellations with higher P_{avg}, which could be used for a variety of purposes - e.g. increased uplink range or improved link reliability or support operation of additional sensing circuits. But FEC codes also decrease the amount of info. bits per symbol (C). Thus it is necessary to examine both P_{avg} and code choice.

The performance of FEC options with memory - such as convolutional coding - provides appropriate guidelines for the choice of coding to be used in conjunction with 4-QAM. Recall that our choice of 4-QAM constellation is equivalent to two orthogonal BPSK signals. The probability of bit error $p_b(e)$ of a coded BPSK signal at high SNR has a good approximation [46], and the corresponding BER for bit A and B is

$$p_A(e) \approx \mathcal{Q}\left(\sqrt{\frac{2E_{b,A}Rd_{free}}{N_0}}\right) \tag{2.39}$$

$$p_B(e) \approx \mathcal{Q}\left(\sqrt{\frac{2E_{b,B}Rd_{free}}{N_0}}\right) \tag{2.40}$$

where R is the code rate, d_{free} is the free distance of the code, and $(E_{b,A}\, E_{b,B})$ is the respective energy per information bit. This approximation to the BER is used to calculate the reader's BER for the UEP and EEP schemes.

In the case of EEP, the total symbol energy is distributed equally to both bits that are coded with identical FEC, resulting in equal error protection: $p_A(e) = p_B(e)$. Additionally, the FEC can decode bits with the same reliability as uncoded detection using a factor of d_{free} less energy per symbol, or Rd_{free} less energy per bit. Lowering the symbol energy maps to decreasing r of the constellation (2.14) Constellations with smaller r and constant θ_i result in backscatter modulator designs with a higher P_{avg}. The P_{avg} is calculated using two different FECs, each with different rates and d_{free}, specifically ($R = \frac{1}{2}$, $d_{free} = 8$), and ($R = \frac{3}{4}$, $d_{free} = 5$). The inclusion of the new P_{avg} and the new info. rate (because of FEC)

results in the new $\mathcal{B}_{loss}$.

The second approach uses a variation of UEP coding. Originally, the motivation for UEP was to design codes that provide unequal protection for each bit based on the differences in their respective significance [47]. However, in our work, we use UEP FEC to yield equal bit error rates post coding for a constellation with unequal BER when uncoded, i.e., for the 4-QAM case where $p_A(e) \neq p_B(e)$. From (2.18), unequal error probabilities requires asymmetry in the constellation, as illustrated by Fig. 2.4(b), that is then compensated by the UEP FEC.

In case of an asymmetry between the reactive (quadrature) and resistive (inphase) bits, PSK out performs ASK modulation in terms of power harvesting [2]. To minimize the power loss of the resistive modulated (B) bit, the symbol energy for bit B is decreased (by an amount determined by the FEC coding gain), i.e. $E_B < E_A$ as in Fig. 2.4(b). Decreasing only E_B and keeping E_A constant thereby decreases the constellation's r and changes θ_0 and θ_1. These changes in constellation geometry allow for a feasible backscatter modulator with a higher P_{avg}. To maintain reliability, this bit is then protected with a FEC. Bit A is otherwise left *uncoded*, which results in $1 + R$ information bits per symbol. Leaving a bit uncoded also reduces complexity at the reader, because only part of the constellation is required to be processed by the FEC decoder. Lastly, the same two codes used in EEP are used to encode bit B and in addition a code with $R = \frac{1}{4}$ and $d_{free} = 18$ is also used.

2.6 Numerical Results

Table 2.1 displays the backscatter power efficiency loss per bit $\frac{P_{match} - P_{avg}}{C\, P_{match}}$ for each of the modulations discussed previously. Each combination of constellation and FEC is designed to meet an uplink range of 25 m, a BER of 10^{-3} and 10^{-5}, for a fixed symbol rate of 640 KHz. The SNR and P_{avg} is calculated using the approach discussed in Sec. 2.2.3 and Sec. 2.3.2. Additionally, the values for the reader and tag parameters are listed in Appendix A. As expected, PSK has the superior power loss figure as compared to ASK, despite the fact that at the receiver, the two constellations are identical in BER and information rate. The next notable result is the difference in performance between uncoded 4-QAM and PSK. As discussed in Sec. 2.3.1, the uncoded 4-QAM constellation is equivalent to a mean shifted

Modulation Type	$\mathcal{B}_{loss}$ at $p(e) = 10^{-3}$	$\mathcal{B}_{loss}$ at $p(e) = 10^{-5}$	C
2-PSK	0.001	0.003	1
UEP $R = 1/4$ $d_{free} = 18$	0.017	0.024	1.25
UEP $R = 1/2$ $d_{free} = 8$	0.021	0.029	1.5
UEP $R = 3/4$ $d_{free} = 5$	0.022	0.031	1.75
EEP $R = 3/4$ $d_{free} = 5$	0.025	0.034	1.5
EEP $R = 1/2$ $d_{free} = 8$	0.03	0.041	1
4-QAM	0.04	0.057	2
2-ASK	0.08	0.11	1

Table 2.1: Backscatter power efficiency loss per bit for different modulations and FECs. The numerical result is calculated from a hypothetical tag and reader setup

QPSK constellation. It is well-known that BPSK and QPSK have identical power cost per bit for active transceiver architectures, however in this example, the 4-QAM constellation is considerably more expensive in terms of the backscatter power efficiency loss per bit than PSK.

UEP performs the best out of all the code choices, except for 2-PSK. Despite UEP and EEP using identical FECs, UEP performs better. For nearly equal $\mathcal{B}_{loss}$, EEP achieves better spectral efficiency at the same reliability. Indicating that the majority of the power is lost in the resistive modulated bit as opposed to reactive modulated bit. Hence in this example, coding the reactive bit (A) is less efficient than leaving the bit uncoded. Additionally, it explains that constellation design for backscatter modulation should be treated differently than traditional modulation by including the notion of (normalized) backscatter efficiency loss, which is not bit-wise identical. Finally, it demonstrates the importance of co-design of the constellation and FEC.

While 2-PSK out performs UEP in terms of backscatter power efficiency loss per bit, UEP offers better spectral efficiency than PSK. UEP also outperforms uncoded 4-QAM for both BER constraints. This suggests that a system designer should consider UEP coding

in conjunction with 4-QAM for RFID uplink, if that is a feasible option given the overall specifications.

2.7 Conclusion

This chapter presents a new analysis for RFID link that integrates the traditional communication metrics (BER and spectral efficiency) with the notion of backscatter power efficiency loss per bit. The latter is an important consideration in power starved RFID tags that must rely on harvested power for operation. Its inclusion couples the downlink with the uplink via the choice of the power harvesting circuitry parameters (downlink) and constellation design (uplink) - this distinguishes our work from traditional communication systems.

Our results show a significant increase in the backscatter power efficiency loss per bit from 2-QAM to 4-QAM in return for the increased spectral efficiency of the latter. Consequently, we propose novel coded 4-QAM schemes based on UEP that help bridge the gap in performance between 2-QAM and 4-QAM. These new physical layer options may provide guidance for future Class 1 Gen 2 enhancements that yield a better balance between the tag's two key attributes: harvested energy and uplink efficiency.

Chapter 3

SPACE TIME CODING FOR MIMO RFID SYSTEMS

3.1 Introduction

As RFID applications become more data intensive, the expectations for higher rates and greater reliability on the uplink (tag to reader) will escalate beyond the current limits (maximum 640 Kbps on uplink with demonstrated achieved ranges of a few meters). As seen in Wi-Fi and cellular technologies [18], multi-antenna techniques offer simple and effective solutions that improve the uplink rate or reliability. This work seeks to explore the potential application of such concepts to RFID systems. While adding multiple antennas to an RFID reader is a natural solution, the number of antennas allowed per tag is limited by the tag's form factor, hence, this chapter studies the performance gains from a limited number of antennas.

A full chain consists of two signal paths: forward unmodulated signal from reader to tag on the downlink, and the backscattered modulated signal from tag to reader on the uplink. In general, the backscattered signal is subject to environmental *multipath* en route reader that cause both frequency and time selective effects. The uplink symbol rate is sufficiently low, such that we may ignore the impact of any frequency selectivity, i.e., we assume no inter-symbol interference. Typically the physical environment changes slowly over time, so the symbols experience *slow fading* multipath conditions. Additionally, the antennas at the reader and tag are spaced sufficiently far apart so that the fading at each antenna is statistically independent; an assumption we'll justify shortly. A model that captures these features is the (M, L, N) *dyadic backscatter channel* (DBC) introduced by [19]. The downlink and uplink path gains are represented respectively by two matrices with independent and identically distributed (iid) complex Gaussian channel elements. The first represents the (pairwise) channel from each of M transmit (reader) antennas to the L tag antennas on the forward path, and the second captures the pairwise channel between any

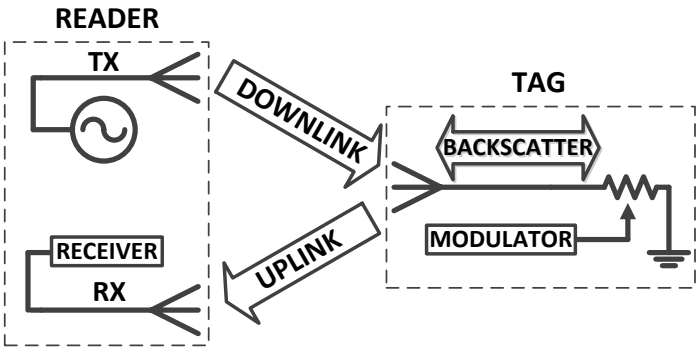


Figure 3.1: Single reader and tag architecture with separate uplink and downlink fading paths.

L tag antennas to the N receive (reader) antennas on the reverse. Clearly, the composite (dyadic) channel gains are no longer Gaussian and hence the fading is no longer Rayleigh distributed.

The physical form factor and cost are key attributes that determine the practicality of RFID systems. To have multiple antennas on both the tag and reader, the tag form factor must increase in order to separate antennas sufficiently (at least by one λ) to achieve independent fading. For RFID systems operating in the FCC 915 MHz ISM band, the required antenna dimensions make it difficult to design a compact system with independent fading at each antenna. To dramatically decrease the antenna size, designers have turned to operation in the unlicensed 5.1 GHz band instead of the 915 MHz band [19]. It is shown that a practical dual antenna tag with dimensions $3.8\,\text{cm} \times 7.7\,\text{cm}$ can be built [48, 49] for operation at 5.79 GHz. So advanced tags, such as the WISP, with two to four separated antennas operating at 5.8 GHz may be commercially feasible. Lastly, additional receiver chains on the reader is also economical, given today's prevalence of consumer grade MIMO systems (e.g. 802.11n, HSPA+, LTE).

3.1.1 Related Work

Previous works on MIMO applied to RFID based systems are sparse. The concept of using MIMO techniques for RFID was introduced by [50], but the work focuses on system level

descriptions. Of the existing literature, the majority has focused solely on multiple antennas at the reader, while ignoring the tags. [51] and [52] considers the hardware architecture needed for additional reader antennas, and [53] analyzes the channel corresponding to a single tag antenna with a MIMO reader. The multi-antenna tag channel and the DBC were first studied in detail by [19]. Later extensions of the work considers only the uncoded BER for the multiple input, single output (MISO) channel [54], correlated MISO channels [55], and simulation results [56, 57] for the Alamouti space time block code (STBC) [58]. It is also pointed out by [54] that the $(1, L, 1)$ DBC is equivalent to the $(L, 1)$ cascaded Rayleigh channel discussed by [59]. Accordingly, we are the first to systematically analyze the diversity order achievable for the MIMO DBC channel.

Since the completion of this work, the area of RFID MIMO systems has greatly increased in size and scope. The performance of space-time block codes is analyzed for the DBC Rayleigh fading case [60–62][1] and later for the Rician and Nakagami-m fading cases [63]. Instead of focusing on link design, other work has proposed reader antenna beamforming techniques in order to enhance power transfer the RFID tag [64, 65], which has the most impact for fully passive tags. Lastly, there has been progress on using multi-port theory to determine the individual channel gains of the DBC based on measurements only at the reader's receiver [66].

3.1.2 Contribution and Organization

This chapter is organized as follows. Section 3.2 discusses backscatter modulation in an MIMO environment, and discusses relationship between the existing pinhole MIMO channel model and the DBC model. Next, the DBC model [19] is extended to support our analysis in the MIMO DBC setting. Section 3.3 derives a bound on the symbol pairwise error probability for general space-time codes that is numerically computed; for orthogonal space-time codes (OSTBC), a closed form PEP bound is shown. Section 3.4 finds the asymptotic diversity order for OSTBCs based on results of Section 3.3. Section 3.5 evaluates the performance of OSTBCs and space-time trellis codes (STTC) using both analytical and

[1]Similar results by [60] was developed in parallel and independent of our work.

Monte Carlo simulation based results.

3.2 System Model - RFID Backscatter MIMO Channel

What distinguishes the dyadic backscatter channel (DBC) model from a standard (M, N) MIMO channel is that each tag antenna behaves as a *pinhole* [19] for uplink communication. Conceptually, a pinhole results when signals from the multiple paths superimpose at a single point, and the *combined signal re-scatters* [67–69]. The superposition at a pinhole occurs because the physical environment only supports one mode of propagation between the transmit and receive antenna arrays, despite having a scatter rich environment near each antenna. Pinholes in traditional (M, N) MIMO systems have been known to occur in many situations [70], e.g. a long hallway that supports only a single mode of propagation and others are discussed in [19]. To motivate the basis of the DBC model, the standard MIMO and 'pinhole' MIMO channels are first described.

The standard (M, N) MIMO system [18] is modeled by the received signal $\mathbf{y}_n \in \mathbb{C}^N$, at time instant n given by

$$\mathbf{y}_n = \mathbf{H}\mathbf{x}_n + \mathbf{w}_n, \tag{3.1}$$

where $\mathbf{H} \in \mathbb{C}^{N\times M}$ is the channel matrix, $\mathbf{x}_n \in \mathbb{C}^M$ is the transmitted signal, and $\mathbf{w}_n \in \mathbb{C}^N$ is AWGN. Typically, $\mathbf{H}$ is modeled as a full rank matrix with i.i.d. complex Gaussian entries. However in the pinhole MIMO channel, there exists two sets of paths traveling to and from each pinhole, so $\mathbf{H}$ is no longer Gaussian. The respective channel matrix for L independent pinholes is modeled as a sum of dyadic products [68][2]

$$\mathbf{y}_n = \sum_{i=1}^{L} s_i \mathbf{h}_i^{rx} \mathbf{h}_i^{tx\mathsf{T}} \mathbf{x}_n + \mathbf{w}_n = \mathbf{H}^{rx}\mathbf{Q}\mathbf{H}^{tx\mathsf{T}}\mathbf{x}_n + \mathbf{w}_n, \tag{3.2}$$

where $[\mathbf{h}_1^{tx}, \ldots, \mathbf{h}_L^{tx}] = \mathbf{H}^{tx} \in \mathbb{C}^{M\times L}$, $[\mathbf{h}_1^{rx}, \ldots, \mathbf{h}_L^{rx}] = \mathbf{H}^{rx} \in \mathbb{C}^{N\times L}$, $[s_1, \ldots, s_L]^{\mathsf{T}} = \mathbf{s} \in \mathbb{C}^L$, and $\mathrm{diag}\{\mathbf{s}\} = \mathbf{Q} \in \mathbb{C}^{L\times L}$. The matrices $\mathbf{H}^{tx}$ and $\mathbf{H}^{rx}$ represent the gains to and from each pinhole antenna pair, and $\mathbf{Q}$ represents the complex gain at each pinhole. Each pinhole gain s_i is determined by the physical environment and remains static. Furthermore, the channel

[2]Dyadic products are more commonly known as outer products.

fading is assumed flat and independent, whereby each element in $\mathbf{H}^{rx}$ and $\mathbf{H}^{tx}$ are drawn i.i.d from $\mathcal{CN}(0,1)$.[3]

Consider a RFID system consisting of a multi-antenna reader with M transmit antennas and N receive antennas, and a RFID tag with L antennas as in Figure 3.2. The corresponding DBC model has the form of (3.2) [19], but with the following key differentiating characteristic. For the pinhole MIMO channel, the sender attempts to communicate the signal $\mathbf{x}_n$ (unknown to the receiver) via a noisy observation at the receiver, $\mathbf{y}_n$. However for the RFID scenario, the signal $\mathbf{x}_n$ is the carrier wave generated by the reader, and the diagonal elements of $\mathbf{Q}$ (s_i) are determined by the tag antenna's impedance loading, and no longer by the environment as in the (M,N) pinhole channel. Hence, the tag attempts to communicate the signal $\mathbf{Q}$ (unknown at the receiver) via a noisy observation at the receiver, $\mathbf{y}_n$, with $\mathbf{x}_n$ known a-priori at the receiver.[4] In other words, the unknown information for the reader to decode resides in $\mathbf{Q}$.

While (3.2) accurately models the system for a single time instant, it needs to be extended over multiple symbol periods to allow for space-*time* coding, i.e. when $\mathbf{Q}$ changes over different slots. Both the traditional (M,N) MIMO or pinhole MIMO channels are easily extended over K time slots by collecting $\mathbf{x}$ and $\mathbf{y}$ as time-indexed matrix columns, e.g. $[\mathbf{x}_1,\ldots,\mathbf{x}_K]$, as follows. First, observe that $\mathbf{Q}\mathbf{H}^{tx\mathsf{T}}\mathbf{x}_n = \mathbf{s}\circ(\mathbf{H}^{tx\mathsf{T}}\mathbf{x}_n)$, where $\circ$ is the Hadamard product.[5] This results in

$$\mathbf{y}_n = \mathbf{H}^{rx}(\mathbf{s}\circ(\mathbf{H}^{tx\mathsf{T}}\mathbf{x}_n)) + \mathbf{w}_n. \tag{3.3}$$

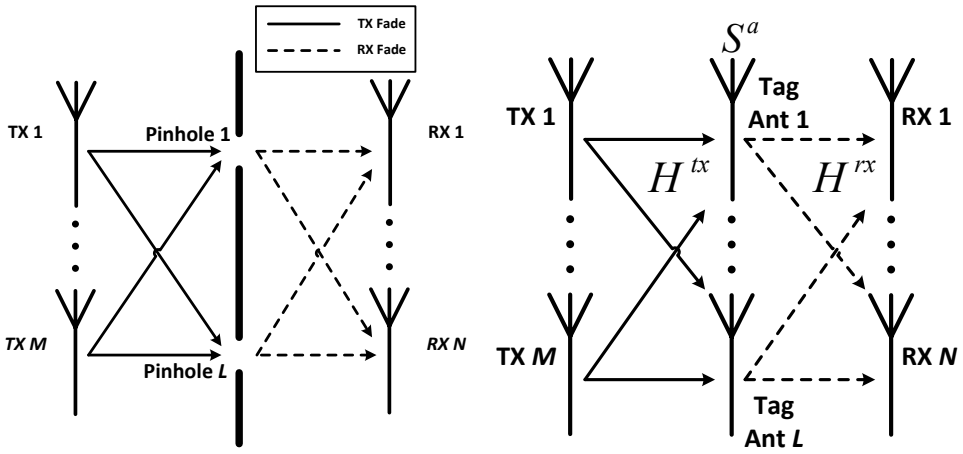
Note that if $L=1$ the model defaults to the $(M,1,N)$ DBC where no space time coding is possible, which is discussed in [19].

Now, consider a *space-time* block code over K symbol durations (slots). The reader's N receive antennas observes the signal $[\mathbf{y}_1,\ldots,\mathbf{y}_K] = \mathbf{Y} \in \mathbb{C}^{N\times K}$ in the presence of additive noise $[\mathbf{w}_1,\ldots,\mathbf{w}_K] = \mathbf{W} \in \mathbb{C}^{N\times K}$. The (known) sequence sent by the reader's transmit

[3] $\mathcal{CN}$ denotes the circular symmetric complex normal distribution.

[4] $\mathbf{x}_n$ also leaks from the transmitter to the receiver in most scenarios. For the remainder of the paper, we ignore such leakage - under the assumption that it is adequately canceled.

[5] The $\circ$ operator indicates the Hadamard product of matrices, i.e., point-wise multiplication $[A\circ B]_{ij} = A_{ij}B_{ij}$.



(a) (b)

Figure 3.2: (a) A two pinhole MIMO channel with multiple transmit and receive antennas and multiple pinholes. The signal emitted by each pinhole is the sum of the paths arriving at each pinhole. (b) The (M, L, N) DBC, where the signal observed at each receive antenna is the cascaded of the $\mathbf{H}^{tx}$ and $\mathbf{H}^{rx}$ channels.

antennas is $[\mathbf{x}_1, \ldots, \mathbf{x}_K] = \mathbf{X} \in \mathbb{C}^{M \times K}$. The tag antennas send the impedance-modulated code word $[\mathbf{s}_1^a, \ldots \mathbf{s}_K^a] = \mathbf{S}^a \in \mathbb{C}^{L \times K}$ via

$$\mathbf{Y} = \mathbf{H}^{rx} \left(\mathbf{S}^a \circ \left(\mathbf{H}^{tx\mathsf{T}} \mathbf{X} \right) \right) + \mathbf{W}, \tag{3.4}$$

where each code word $\mathbf{S}^a$ (indexed by a) is drawn from a code book $\mathcal{C}$.

Consider the simplest case, where the reader transmits a constant carrier wave with net energy E equally shared among all transmit antennas, i.e., $\mathbf{X} = \sqrt{E/M}\mathbf{1}_M \mathbf{1}_K{}^{\mathsf{T}}$, where $\mathbf{1}_M$ is a column vector of M ones. Rewriting (3.4) for a constant transmit signal yields:

$$\mathbf{Y} = \mathbf{H}^{rx} \left(\mathbf{S}^a \circ \left(\sqrt{E/M} \mathbf{H}^{tx\mathsf{T}} \mathbf{1}_M \mathbf{1}_K{}^{\mathsf{T}} \right) \right) + \mathbf{W}. \tag{3.5}$$

Lemma 3.2.1. *The statistics of* $\mathbf{Y}$ *are invariant to the number of TX antennas* M*; i.e., a* $(1, L, N)$ *configuration is statistically identical to a* (M, L, N) *configuration.*

Proof. Consider the matrix of random variables $\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M \mathbf{1}_K{}^{\mathsf{T}}$. From the properties of independent Gaussian random variables, each element of the vector $\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M$ is i.i.d. $\mathcal{CN}(0, 1)$ for any M, implying that $\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M \mathbf{1}_K{}^{\mathsf{T}}$ has identical statistics for any

M. Even with space-time coding it is not possible exploit any diversity from the paths described by $\mathbf{H}^{tx}$. □

This result agrees with the results for space only signaling [19]; increasing M has no benefit.

It is straightforward to show

$$\text{diag}\{\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M\}\mathbf{S}^a = \mathbf{S}^a \circ \left(\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M\mathbf{1}_K{}^{\mathsf{T}}\right). \tag{3.6}$$

For compactness, write $\mathbf{D}^{tx} = \text{diag}\{\sqrt{1/M}\mathbf{H}^{tx\mathsf{T}}\mathbf{1}_M\}$, resulting in a simplified version of (3.5) as

$$\mathbf{Y} = \sqrt{E}\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{S}^a + \mathbf{W}. \tag{3.7}$$

From Lemma 3.2.1, the diagonal elements of $\mathbf{D}^{tx}$ are i.i.d. Gaussian and the statistics are invariant to the choice of M. Notice that (3.7) follows the form of a standard MIMO channel with the difference that the elements of the product $\mathbf{H}^{rx}\mathbf{D}^{tx}$ are not i.i.d. Gaussian. This form of the DBC model is used to analyze the error performance of space-time codes.

3.3 PEP Analysis

The pairwise error probability (PEP) is the probability that a transmitted code word $\mathbf{S}^a$ is incorrectly detected as some other valid code word $\mathbf{S}^b$, which provides insight on code performance under ideal assumptions [18,31]. Its derivation undertakes the following steps:

1. Compute the PEP, conditioned on the channel state.

2. Find an upper bound on the conditional PEP, such as the Chernoff Bound.

3. Average over the channel state variables and obtain the unconditional PEP.

3.3.1 PEP for General Codebook $\mathcal{C}$

Let $p\left(\mathbf{S}^a \to \mathbf{S}^b\right)$ denote the PEP for the symbol pair a, b. By law of total probability, $p(\mathbf{S}^a \to \mathbf{S}^b) = \mathbf{E}_{\mathbf{H}^{rx}\mathbf{D}^{tx}}\left[p\left(\mathbf{S}^a \to \mathbf{S}^b \,\middle|\, \mathbf{H}^{rx}, \mathbf{D}^{tx}\right)\right]$, via averaging over the receive and transmit channel gains. For the measurement model in additive white Gaussian noise (AWGN)

described by (3.7), classical detection theory leads to a receiver that chooses the *nearest* codeword, i.e. one with the smallest Euclidean distance from the received signal. Hence,

$$p\left(\mathbf{S}^a \to \mathbf{S}^b \left| \mathbf{H}^{rx}, \mathbf{D}^{tx} \right.\right) = \mathcal{Q}\left(\sqrt{\frac{\mathrm{SNR}}{2}} \|\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{C}\|_{\mathbf{F}}\right) \tag{3.8}$$

where $\mathrm{SNR} = E/N_0$ and $\mathbf{C} = \mathbf{S}^a - \mathbf{S}^b$ is the codeword difference. The PEP is defined in terms of the well-known indefinite integral $\mathcal{Q}(.)$ [6] that renders averaging out the channel variables infeasible. However, using the Chernoff bound for the $\mathcal{Q}(.)$ function $\mathcal{Q}(x) \leq \frac{1}{2}\exp\left(\frac{-x^2}{2}\right)$, yields

$$p\left(\mathbf{S}^a \to \mathbf{S}^b \left| \mathbf{H}^{rx}, \mathbf{D}^{tx} \right.\right) \leq \frac{1}{2}\exp\left(-\frac{\mathrm{SNR}}{4}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{C}\|_{\mathbf{F}}^2\right). \tag{3.9}$$

Lemma 3.3.1. *$\|\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{C}\|_{\mathbf{F}}^2 = \sum_{i=0}^{L}\psi_i\gamma_i$, where ψ_i are the eigenvalues of the random matrix $\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger}$. Also, $p(\gamma_i|\mathbf{D}^{tx}) \sim \Gamma(N,1)$ [7] where γ_i is (conditionally) independent of γ_j, given $\mathbf{D}^{tx}$ $\forall i \neq j$.* [8]

Proof. Expanding the Frobenius norm

$$\begin{aligned}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{C}\|_{\mathbf{F}}^2 &= \mathrm{tr}\left[\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger}\mathbf{H}^{rx\dagger}\right] \\ &= \sum_{i=1}^{N}\mathbf{h}_i^{\dagger}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger}\mathbf{h}_i,\end{aligned} \tag{3.10}$$

where $[\mathbf{h}_1, \ldots, \mathbf{h}_N] = \mathbf{H}^{rx\dagger}$. For any $\mathbf{D}^{tx}$, $\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger}$ is Hermitian. Since all Hermitian matrices are diagonalized by a unitary matrix, i.e. $\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger} = \mathbf{U}\mathbf{\Psi}\mathbf{U}^{\dagger}$ where $[\mathbf{u}_1, \ldots, \mathbf{u}_L] = \mathbf{U}$ and $\mathbf{U}\mathbf{U}^{\dagger} = \mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}$ and the diagonal eigenvalue matrix $\mathbf{\Psi}$ composed of elements ψ_i. Each term in the sum of (3.10)

$$\mathbf{h}_i^{\dagger}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^{\dagger}\mathbf{D}^{tx\dagger}\mathbf{h}_i = \sum_{j=1}^{L}\psi_j|\mathbf{h}_i^{\dagger}\mathbf{u}_j|^2. \tag{3.11}$$

Since the vectors $\mathbf{u}_j$ form an orthonormal basis and the elements of all the vectors $\mathbf{h}_j$ are i.i.d $\mathcal{CN}(0,1)$, the random variable $\mathbf{h}_i^{\dagger}\mathbf{u}_j$ is i.i.d. $\mathcal{CN}(0,1)$ $\forall i, j$.

---

[6] $\mathcal{Q}(x) = 1/\sqrt{2\pi}\int_x^{\infty}\exp(-x^2/2)\mathrm{d}x$.

[7] The $\Gamma(m,\lambda)$ distribution has the pdf $f_X(x) = \lambda e^{-\lambda x}(\lambda x)^{m-1}/(m-1)!$ [71].

[8] $\dagger$ indicates conjugate transpose.

Combining (3.11) into (3.10) results in

$$\sum_{i=1}^{N}\sum_{j=1}^{L}\psi_j|\mathbf{h}_i^{\dagger}\mathbf{u}_j|^2=\sum_{j=1}^{L}\psi_j\gamma_j. \tag{3.12}$$

where γ_i is the sum of N magnitude squared i.i.d $\mathcal{CN}(0,1)$ random variables. Thus $p(\gamma_i|\mathbf{D}^{tx})\sim\Gamma(N,1)$ and form an i.i.d set [71]. □

Inserting into (3.9)

$$p\left(\mathbf{S}^a\rightarrow\mathbf{S}^b\mid\psi,\gamma\right)\leq\frac{1}{2}\exp\left(-\frac{\mathrm{SNR}}{4}\sum_{i=0}^{L}\psi_i\gamma_i\right). \tag{3.13}$$

Lemma 3.3.2. *For* $\gamma_i\sim\Gamma(N,1)$, $\mathbf{E}_{\gamma_i}\exp\left(-\frac{\mathrm{SNR}}{4}\psi_i\gamma_i\right)=\left(1+\frac{\mathrm{SNR}}{4}\psi_i\right)^{-N}$.

Proof. The expectation over γ_i of the Chernoff bound is equivalent to computing the moment generating function of the $\Gamma(N,1)$ distribution [72]. Hence, $M_\Gamma(s)=\mathbf{E}_{\gamma_i}e^{s\gamma_i}=(1-s)^{-N}$. Setting $s=-\frac{\mathrm{SNR}}{4}\psi_i$ yields $\left(1+\frac{\mathrm{SNR}}{4}\psi_i\right)^{-N}$. □

Unfortunately, the distribution of the eigenvalues ψ_i of the random matrix in (3.13) is unavailable in closed form, and an alternate route is needed, as developed next. [9] Assume $\mathrm{rank}\{\mathbf{CC}^{\dagger}\}=\bar{L}$ where $\bar{L}\leq L$. Since $\mathrm{rank}\{\mathbf{D}^{tx}\}=L$ with probability 1, then the $\mathrm{rank}\{\mathbf{D}^{tx}\mathbf{CC}^{\dagger}\mathbf{D}^{tx\dagger}\}=\bar{L}$. [10] As a consequence, the count of non-zero ψ_i equals the rank of $\mathbf{CC}^{\dagger}$. This suggests that the code matrix rank affects how quickly the PEP decreases with SNR. From Lemma 3.3.1 and 3.3.2, the conditional upperbound is

$$\begin{aligned}p\left(\mathbf{S}^a\rightarrow\mathbf{S}^b\mid\psi\right)&\leq\frac{1}{2}\prod_{i=1}^{\bar{L}}\left(1+\frac{\mathrm{SNR}}{4}\psi_i\right)^{-N}\\&=\frac{1}{2}\left|\mathbf{I}+\frac{\mathrm{SNR}}{4}\mathbf{D}^{tx}\mathbf{CC}^{\dagger}\mathbf{D}^{tx\dagger}\right|^{-N},\end{aligned} \tag{3.14}$$

where the equality reflects the fact that the product of eigenvalues equals the determinant, and $1+\frac{\mathrm{SNR}}{4}\psi_i$ for $i=1,\ldots\bar{L}$ are the $\bar{L}$ non-zero eigenvalues of $\mathbf{I}+\frac{\mathrm{SNR}}{4}\mathbf{D}^{tx}\mathbf{CC}^{\dagger}\mathbf{D}^{tx\dagger}$. In general the distribution of random determinants is not known, so the conditional PEP is

[9] Note that the original MIMO diversity analysis for standard Rayleigh channel in [31] did not encounter the random eigenvalue problem as above.

[10] A standard result in linear algebra states for any $\mathbf{B}$, the $\mathrm{rank}\{\mathbf{AB}\}=\mathrm{rank}\{\mathbf{B}\}$ if $\mathbf{A}^{\dagger}$ is full rank.

averaged out using Monte Carlo evaluation. Further details on the evaluation of this bound are discussed in Section 3.5.2.

To provide better intuition into code performance, we show that $|\mathbf{C}\mathbf{C}^\dagger|$ approximates the available coding gain. The Minkowski determinant inequality [73] states for any positive semi-definite and Hermitian matrices $\mathbf{A}_1, \mathbf{A}_2$, $|\mathbf{A}_1 + \mathbf{A}_2|^{1/n} \geq |\mathbf{A}_1|^{1/n} + |\mathbf{A}_2|^{1/n}$. Since $\mathbf{I}$ and $\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^\dagger\mathbf{D}^{tx\dagger}$ are Hermitian matrices, applying the above to (3.14) for $n = 1$ results in

$$\left|\mathbf{I} + \frac{\mathrm{SNR}}{4}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^\dagger\mathbf{D}^{tx\dagger}\right| \geq 1 + \left|\frac{\mathrm{SNR}}{4}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^\dagger\mathbf{D}^{tx\dagger}\right| , \tag{3.15}$$

thereby proving

$$p\left(\mathbf{S}^a \to \mathbf{S}^b \mid \psi\right) \leq \frac{1}{2}\left(1 + \left|\frac{\mathrm{SNR}}{4}\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^\dagger\mathbf{D}^{tx\dagger}\right|\right)^{-N} . \tag{3.16}$$

Additionally, the above is asymptotically tight, i.e., as $\mathrm{SNR} \to \infty$ the inequality in (3.16) becomes an equality.

Form the the properties of determinants, $|\mathbf{D}^{tx}\mathbf{C}\mathbf{C}^\dagger\mathbf{D}^{tx\dagger}| = |\mathbf{C}\mathbf{C}^\dagger||\mathbf{D}^{tx}\mathbf{D}^{tx\dagger}|$, and $|\mathbf{D}^{tx}\mathbf{D}^{tx\dagger}| = \prod |D^{tx}_{ii}|^2$. If $\mathbf{C}\mathbf{C}^\dagger$ is full rank, i.e $\bar{L} = L$, the above yields

$$p\left(\mathbf{S}^a \to \mathbf{S}^b \mid \psi\right) \leq \frac{1}{2}\left(1 + \left(\frac{\mathrm{SNR}}{4}\right)^L |\mathbf{C}\mathbf{C}^\dagger| \prod_{i=1}^{L} |D^{tx}_{ii}|^2\right)^{-N} . \tag{3.17}$$

Note that $|\mathbf{C}\mathbf{C}^\dagger|$ (which multiplies the SNR terms) represents a coding gain, similar to the result for Rayleigh channel STTCs [31]. However, the distribution of a product of Rayleigh random variables is not available in closed form [74], so Monte Carlo evaluation is necessary. Lastly, we report that (3.17) is quite loose at moderate SNR when compared to (3.14).

3.3.2 OSTBC Design

STBCs originated from the goal of achieving transmit diversity with low decoding complexity. Orthogonal STBCs (OSTBC) use transmitted signals from different antennas that are orthogonal and as a result, the received signal can be equivalently written in terms of parallel and independent single input single output (SISO) channels, which leads to very efficient (low-complexity) decoding [75, 76]. This comes at a price: OSTBCs in general are *not* full rate (i.e. $R = 1$), where the rate is defined as the number of information symbols sent per symbol period.

Using the channel model described by (3.7), define $\mathbf{S}^a$ as an OSTBC codeword of dimension $L \times K$ (antennas $\times$ symbol slots). Each OSTBC codeword draws from a set of vector information symbols $\mathbf{c}^a \in \mathbb{C}^{RK}$ that belong to some symbol alphabet. The Alamouti code is the best known example of an OSTBC for a two antenna ($L = 2$) system that achieves $R = 1$, whose codeword is given by (rows denote antennas, columns denote time) [58]:

$$\mathbf{S}^a = \begin{pmatrix} c_1^a & -c_2^{a*} \\ c_2^a & c_1^{a*} \end{pmatrix}. \tag{3.18}$$

For a general MIMO system defined by an arbitrary channel matrix $\mathbf{H}$, a OSTBC scheme decouples the MIMO channel into RK independent AWGN channels [76]. The equivalent SISO channel model is

$$\mathbf{y}' = \sqrt{E}\|\mathbf{H}\|_{\mathbf{F}}\, \mathbf{c}^a + \mathbf{w}' \tag{3.19}$$

with $w_i' \sim \mathcal{CN}(0, N_0)$, $\mathbf{w}' \in \mathbb{C}^{RK}$, and $\mathbf{y}' \in \mathbb{C}^{RK}$. The symbol PEP for $\mathbf{c}^a \to \mathbf{c}^b$ depends only on the statistics of $\|\mathbf{H}\|_{\mathbf{F}}$.

Comparing (3.7) to (3.19), the channel matrix for the DBC is $\mathbf{H} = \mathbf{H}^{rx}\mathbf{D}^{tx}$. Thereby the performance of OSTBC for the DBC depends solely on the statistics of $\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}}$. Using the same Chernoff bound approach as for the general codebook case, the conditional PEP bound is now written as

$$p\left(\mathbf{S}^a \to \mathbf{S}^b \,\middle|\, \mathbf{H}^{rx}, \mathbf{D}^{tx}\right) \leq \frac{1}{2}\exp\left(-\frac{\mathrm{SNR}}{4}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}}^2\|\mathbf{c}^a - \mathbf{c}^b\|_2^2\right). \tag{3.20}$$

where the signal to noise ratio is defined as $\mathrm{SNR} = \frac{E}{N_0}$ as before and we may normalize $\|\mathbf{c}^a - \mathbf{c}^b\|_2^2 = 1$ without any loss of generality in the following.

Lemma 3.3.3. *$\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}}^2 = \sum_{i=1}^{L} r_i^2\gamma_i$, where r_i are standard Rayleigh[11] i.i.d. and $p(\gamma_i) \sim \Gamma(N, 1)$ i.i.d.*

[11] $\sqrt{x^2 + y^2}$ where $x, y \sim \mathcal{N}(0, \frac{1}{\sqrt{2}})$ is standard Rayleigh distributed.

Proof. Following the same approach as in the general case; expand the Frobenius norm

$$\begin{aligned}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}}^2 &= \mathrm{tr}\left[\mathbf{H}^{rx}\mathbf{D}^{tx}\mathbf{D}^{tx\dagger}\mathbf{H}^{rx\dagger}\right] \\ &= \sum_{i=1}^{N}\mathbf{h}_i^{\dagger}\mathbf{D}^{tx}\mathbf{D}^{tx\dagger}\mathbf{h}_i \\ &= \sum_{i=1}^{N}\sum_{j=1}^{L}|D_{jj}^{tx}|^2|h_{ij}|^2, \end{aligned} \tag{3.21}$$

where $[\mathbf{h}_1,\ldots,\mathbf{h}_N] = \mathbf{H}^{rx\dagger}$ and the last equality follows because $\mathbf{D}^{tx}$ is diagonal. Setting $r_j = |D_{jj}^{tx}|$ and $\gamma_j = \sum_{i=1}^{N}|h_{ij}|^2$, then

$$\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}}^2 = \sum_{i=1}^{L} r_i^2\,\gamma_i. \tag{3.22}$$

From the definitions of $\mathbf{D}^{tx}$ and $\mathbf{H}^{rx}$, r_i is standard Rayleigh i.i.d. and $\gamma_i \sim \Gamma(N,1)$ i.i.d. □

Using (3.22), the norm in the exponent of (3.20) is replaced with a sum. Thus, the exponential of a sum can be broken into a product of exponentials. Since the vector variables $\mathbf{r}$ and γ terms are mutually independent, the expectation of a product is equal to the product of expectations:

$$\frac{1}{2}\mathbf{E}_{\mathbf{r},\gamma}\left[\exp\left(-\sum_{i=1}^{L}\frac{\mathrm{SNR}}{4}r_i^2\gamma_i\right)\right] = \frac{1}{2}\prod_{i=1}^{L}\mathbf{E}_{r_i,\gamma_i}\left[\exp\left(-\frac{\mathrm{SNR}}{4}r_i^2\gamma_i\right)\right]. \tag{3.23}$$

Applying Lemma 3.3.2, $\mathbf{E}_{\gamma_i}\exp\left(-\frac{\mathrm{SNR}}{4}r_i^2\gamma_i\right) = \left(1+\frac{\mathrm{SNR}}{4}r_i^2\right)^{-N}$.

Lemma 3.3.4. *For r_i standard Rayleigh variables* [12]

$$\mathbf{E}_r\left(1+\frac{\mathrm{SNR}}{4}r_i^2\right)^{-N} = \left[\frac{4}{\mathrm{SNR}}\right]^{N}\exp\left[\frac{4}{\mathrm{SNR}}\right]\Gamma\left(1-N,\frac{4}{\mathrm{SNR}}\right). \tag{3.24}$$

Proof. Similar to the approach in [59], rewrite $\left(1+\frac{\mathrm{SNR}}{4}r_i^2\right)^{-N}$ as $\left[\frac{4}{\mathrm{SNR}}\right]^{N}\left(\frac{4}{\mathrm{SNR}}+r_i^2\right)^{-N}$. Under change of variables, $r_i^2 = \xi_i$; $\mathrm{d}\xi_i = 2r_i\mathrm{d}r_i$,

$$\begin{aligned}\int_0^{\infty}\left(\frac{4}{\mathrm{SNR}}+r_i^2\right)^{-N}p_R(r_i)\mathrm{d}r_i &= \int_0^{\infty}\left(\frac{4}{\mathrm{SNR}}+r_i^2\right)^{-N}e^{-r_i^2}2r_i\mathrm{d}r_i \\ &= \int_0^{\infty}\left(\frac{4}{\mathrm{SNR}}+\xi_i\right)^{-N}e^{-\xi_i}\mathrm{d}\xi_i \\ &= \exp\left[\frac{4}{\mathrm{SNR}}\right]\Gamma\left(1-N,\frac{4}{\mathrm{SNR}}\right). \end{aligned} \tag{3.25}$$

[12] $\Gamma(s,x) = \int_x^{\infty}t^{s-1}e^{-t}\mathrm{d}t$ is the upper incomplete gamma function.

where the last line uses the definition of the incomplete Γ function [77]. The change of variables is also equivalent to the change of variables between the square of a Rayleigh random variable and a exponential random variable. $\square$

Theorem 3.3.5. *Assume a (M, L, N) backscatter channel with constant CW transmit signal across all M antennas and full channel state information (CSI) at the receiver. For an OSTBC, the PEP can be bounded above by*

$$p(\mathbf{S}^a \to \mathbf{S}^b) \leq \mathrm{SNR}^{-LN}\, g(\mathrm{SNR}, L, N). \tag{3.26}$$

Proof: Applying the independence $\mathbf{r}$ and γ, and the results of Lemma 3.3.2 and 3.3.4, the final upper bound for the PEP for a $(1, L, N)$ system is:

$$p(\mathbf{S}^a \to \mathbf{S}^b) \leq \frac{1}{2}\left(\frac{4}{\mathrm{SNR}}\right)^{LN} \exp\left(\frac{4L}{\mathrm{SNR}}\right) \Gamma\left(1-N, \frac{4}{\mathrm{SNR}}\right)^{L}. \tag{3.27}$$

The PEP bound shares the familiar form of the (M, N) MIMO Rayleigh channel; the SNR term raised to a negative exponent. However, the function g is monotonic increasing with SNR, in contrast with Rayleigh channel [31] where the multiplicative factor is a constant. It is anticipated that the presence of $g(.)$ for the DBC channel will offset some of the gain of diversity order, as compared to the Rayleigh case. Interestingly for the $(M, L, 1)$ DBC, (3.27) is equivalent to the bound derived in [59] for the $(L, 1)$ cascaded Rayleigh channel but differs for $N > 1$. Next, we formally derive the diversity order for OSTBC over DBC.

3.4 Diversity Analysis

The diversity order d of a system is the asymptotic rate at which the probability of error decays as a function of (average) SNR of the fading channel, and is defined as

$$d = \lim_{\mathrm{SNR}\to\infty} -\frac{\log p\left(\mathbf{S}^a \to \mathbf{S}^b\right)}{\log \mathrm{SNR}}. \tag{3.28}$$

Two useful facts about the incomplete gamma function that will be required, are listed.

$$1.\ \lim_{x\to 0} \frac{\Gamma(s,x)}{x^s} = -\frac{1}{s} \quad s < 0 \tag{3.29}$$

$$2.\ \frac{\partial \Gamma(s,x)}{\partial x} = -\frac{x^{s-1}}{e^x} \tag{3.30}$$

The first fact can be verified using L'Hopital's rule, and the second fact is a standard result of the incomplete gamma function [78].

Theorem 3.4.1. *A (M, L, N) system employing OSTBC achieves a diversity order of L.*

Proof. From the definition of diversity order and (3.27),

$$d = LN - \lim_{\mathrm{SNR}\to\infty} L\frac{\log \Gamma\left(1-N, 4/\mathrm{SNR}\right)}{\log \mathrm{SNR}}. \tag{3.31}$$

To find the remaining limit, apply L'Hopital's rule. The derivative of the numerator and denominator follow -

$$\frac{\mathrm{d}\log \Gamma\left(1-N, 4/\mathrm{SNR}\right)}{\mathrm{dSNR}} = \frac{1}{\Gamma\left(1-N, 4/\mathrm{SNR}\right)}\frac{(4/\mathrm{SNR})^{-N}}{e^{1/\mathrm{SNR}}}\frac{4}{\mathrm{SNR}^2} \tag{3.32}$$

$$\frac{\mathrm{d}\log \mathrm{SNR}}{\mathrm{dSNR}} = \frac{1}{\mathrm{SNR}}. \tag{3.33}$$

Hence

$$\begin{aligned}\lim_{\mathrm{SNR}\to\infty} L\frac{(4/\mathrm{SNR})^{-N+1}}{\Gamma\left(1-N, 4/\mathrm{SNR}\right)e^{1/\mathrm{SNR}}} &= \lim_{\mathrm{SNR}\to\infty} Le^{-1/\mathrm{SNR}}\left[\frac{\Gamma\left(1-N, 4/\mathrm{SNR}\right)}{(4/\mathrm{SNR})^{1-N}}\right]^{-1} \\ &= L(N-1),\end{aligned} \tag{3.34}$$

where property (3.29) was used. Substituting in Eq. (3.31) results in

$$d = LN - L(N-1) = L. \tag{3.35}$$

□

A diversity of order (L) equal to the number of RFID tag antennas suggests that *transmit/receive antenna diversity is not beneficial in RFID MIMO over DBC*, i.e., adding additional transmit/receive antennas has no impact on the asymptotic PEP. However, since the diversity order is only an *asymptotic* (in SNR) result, changing the number of transmit/receive antennas *does impact the PEP at any realistic SNR* - which is explored next. Re-consider the channel matrix norm in (3.21), $\sum_{j=1}^{L}|D_{jj}^{tx}|^2\sum_{i=1}^{N}|H_{ij}^{rx}|^2$. We explore this

quantity asymptotically with respect to N for a fixed SNR. To normalize against the array processing gain, divide the receive signal by $\sqrt{N}$, i.e.,

$$\frac{1}{\sqrt{N}}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}} = \sqrt{\frac{1}{N}\sum_{j=1}^{L}|D_{jj}^{tx}|^2\sum_{i=1}^{N}|H_{ij}^{rx}|^2} \tag{3.36}$$

Corollary 3.4.2. *For sufficiently large N, (M, L, N) DBC with OSTBC is equivalent (in distribution) to a $(L, 1)$ Rayleigh fading channel with OSTBC.*

Proof. Consider the equivalent OSTBC channel (3.19) normalized with respect to array processing gain -

$$\mathbf{y}' = \sqrt{E/N}\|\mathbf{H}^{rx}\mathbf{D}^{rx}\|_{\mathbf{F}}\ \mathbf{c}^a + \mathbf{w}'. \tag{3.37}$$

Rearranging (3.36), $\sum_{j=1}^{L}|D_{jj}^{tx}|^2\frac{1}{N}\sum_{i=1}^{N}|H_{ij}^{rx}|^2$, the inner sum is the sample mean of i.i.d. random variables. Thus, invoking the weak law of large numbers for large N yields[13] $\frac{1}{N}\sum_{i=1}^{N}|H_{ij}^{rx}|^2 \xrightarrow{p} 1$. Hence

$$\lim_{N\to\infty}\|\mathbf{H}^{rx}\mathbf{D}^{tx}\|_{\mathbf{F}} \xrightarrow{p} \|\mathbf{D}^{tx}\|_{\mathbf{F}}. \tag{3.38}$$

Plugging this back into (3.37)

$$\sqrt{\frac{E}{N}}\|\mathbf{H}^{rx}\mathbf{D}^{rx}\|_{\mathbf{F}}\mathbf{c}^a + \mathbf{w}' \xrightarrow{p} \sqrt{E}\|\mathbf{D}^{rx}\|_{\mathbf{F}}\mathbf{c}^a + \mathbf{w}'. \tag{3.39}$$

which is equivalent to a MISO $(L, 1)$ Rayleigh fading channel with OSTBC. □

A key conclusion from the above is that the dyadic backscatter channel performs no better than the standard $(L, 1)$ MISO Rayleigh fading channel in the limit $N \to \infty$, i.e., its diversity order approaches that of the Rayleigh MISO case. Additionally, to benefit from the diversity order at a reasonable SNR, the reader's receiver requires multiple antennas; i.e. $N > 1$. Thus for fixed N, the diversity order for the DBC is achieved at a much higher SNR, as compared to the Rayleigh MISO channel, although increasing N lowers the SNR at which the diversity order is achieved. These issues are explored in the following section.

[13]The notation $\xrightarrow{p}$ defines convergence in probability.

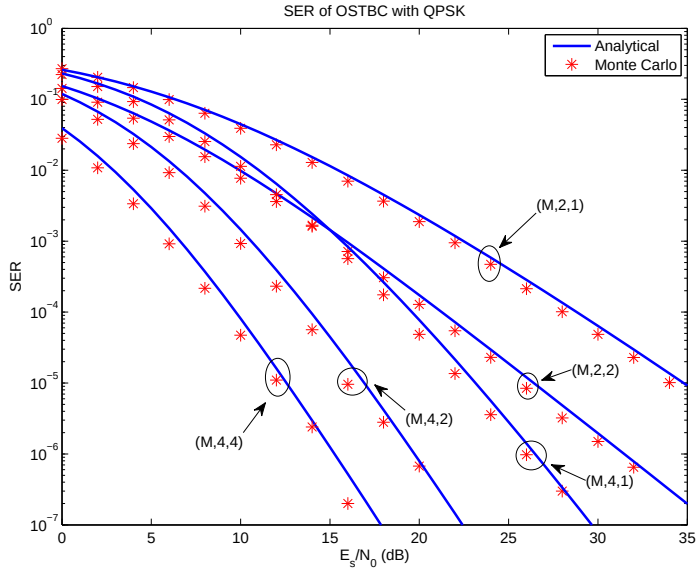


Figure 3.3: Analytical bound (3.27) and Monte Carlo simulations of E_s/N_0 vs SER for OSTBC with QPSK modulation. The channel model simulated is (3.19), where the receiver has perfect knowledge of the channel state.

3.5 Performance Evaluation

In this section, we evaluate the performance of OSTBC and STTC over the MIMO DBC, with quadrature phase shift keying (QPSK) modulation. The derived PEP bounds and diversity order in the previous sections are compared to symbol error rate (SER) obtained from Monte Carlo simulations.[14] The performance metrics used for OSTBC are the coding gain for a fixed SER and the *threshold* E_s/N_0 (average energy per symbol to noise ratio) at which the (asymptotic) diversity order is achieved, i.e. the slope of $-\log(\text{PEP})$ is within 10% of the asymptotic true value (Table 3.1); the performance metrics used for STTC are the coding gain for a fixed SER and the diversity order (Table 3.2). The coding gains are estimated with respect to the worse performing SER curve of Fig. 3.3 and 3.5, respectively. These metrics are measured with respect to the Monte Carlo simulation results; except for the diversity order threshold metric, which uses the analytical bound.

[14] All Monte Carlo simulations and evaluations were conducted in MATLAB environment.

3.5.1 OSTBC Evaluation

This section evaluates the OSTBC based designs discussed in Section 3.3.2 and 3.4. Figure 3.3 plots the analytical bound (3.27) for QPSK modulated OSTBCs, and the corresponding SER Monte Carlo simulations. The pairwise error for the analytical bound corresponds to a pair of QPSK symbols that are nearest neighbors.

The simulation does not conduct a full end-to-end encoding and decoding of OSTBC, but instead uses the OSTBC equivalent SISO model defined by (3.19). Input to the QPSK symbol mapper are i.i.d. equiprobable Bernoulli random variables and our results are averaged over at most 4×10^7 channel instances. Figure 3.4 plots the slope of the analytically derived $-\log(\text{PEP})$ vs E_s/N_0. The key metrics (coding gain and threshold SNR) are summarized in Table 3.1.

In Figure 3.3, the results show that a coding gain occurs for any increase of receive antennas (N) - as denoted by a horizontal shift of the SER curve, and the diversity order increases with increase of tag antennas (L) - as denoted by a change of the slope, confirming the results in Section 3.4. For example, at SER $= 10^{-5}$, the coding gain in increasing $N = 1 \rightarrow 2$ ($L = 2$) is nearly 9 dB whereas $N = 2 \rightarrow 4$ ($L = 4$) is 5 dB, suggesting diminishing returns while increasing N. However since a tag is typically constrained for space, there is a practical limit on the feasible values of L. So for practical systems, the choice of N also has a significant impact on uplink reliability, because the L is fixed to small values.

To illustrate this further, Figure 3.4 plots the numerically calculated slope of $-\log(\text{PEP})$ using the analytical results in Figure 3.3. The final column in Table 3.1 lists the *threshold* E_s/N_0 at which the diversity order - determined by the number of tag antennas ($L = 2$ or $L = 4$) is achieved as N varies. Clearly, increasing $N = 1 \rightarrow 2$ significantly impacts this threshold value. The results suggest that two to four receive antennas would suffice in practice to capture the full diversity order available to OSTBC.

Remark on diversity order and L, N parameters: The analytical and simulation results present a seemingly non-intuitive result. Theoretically, multiple receive antennas are not useful in the high SNR regime, but simulation demonstrate significant coding gains at low

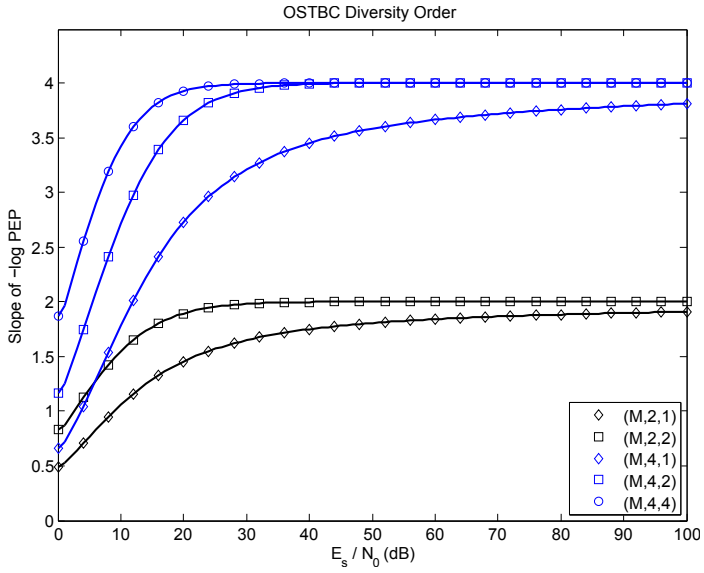


Figure 3.4: Slope of $-\log \mathrm{PEP}$ for OSTBC with QPSK modulation, i.e. diversity order. The analytical upper bound (3.27) is used.

to moderate SNR. To qualitatively explain this relationship, consider Corollary 3.4.2. The OSTBC coded channel is shown to be a sum of L products of two i.i.d. random variables, $\sum_{j=1}^{L} |D_{jj}^{tx}|^2 \sum_{i=1}^{N} |H_{ij}^{rx}|^2$. So there are effectively only L independent paths to transmit information over, which limits the the diversity order to L. At low SNR, the SER is severely affected by fades from either $|D_{jj}^{tx}|$ or $\sum_{i=1}^{N} |H_{ij}^{rx}|$ terms. Increasing N decreases the probability of $\sum_{i=1}^{N} |H_{ij}^{rx}|$ being in a deep fade, but if $|D_{jj}^{tx}|$ is in a deep fade then increasing N will not help much. Hence reliability at low SNR is improved by larger N, but remains ultimately limited by there being only L independent paths.

We note that the diversity results described in this section were independently presented in [60]. Additionally, they find the exact error curves for the DBC and show that the diversity order of uncoded MIMO over the DBC is $d = \min(L, N)$. The latter explains the observation in [19] regarding uncoded BER performance; increasing L and N results in diversity order gains, but increasing only L or N does not improve diversity. The next chapter - on the diversity-multiplexing tradeoff for the DBC - will explain this difference in diversity between the uncoded and OSTBC coded case and explore the benefit of increasing

Table 3.1: Performance Evaluation of OSTBC for the MIMO Dyadic Backscatter Channel

Channel	Coding Gain (SER $= 10^{-5}$)	Diversity Order	Diversity Order Threshold
$(M, 2, 1)$	0dB	2	49dB
$(M, 2, 2)$	9dB	2	16dB
$(M, 4, 1)$	12dB	4	52dB
$(M, 4, 2)$	17dB	4	19dB
$(M, 4, 4)$	22dB	4	12dB

N at the reader.

3.5.2 STTC Evaluation

This section evaluates STTC based designs discussed in Section 3.3.1. The specific STTC used is the full-rate antenna delay diversity (DD) code for two transmit antennas described in [31]. As an example, consider the DD code for two antennas with the information symbol sequence $c_1^a, c_2^a, \ldots, c_{n-1}^a, c_n^a$. The codeword corresponding to this sequence is

$$\mathbf{S}^a = \begin{pmatrix} c_1^a & c_2^a & c_3^a & \ldots & c_n^a & 0 \\ 0 & c_1^a & c_2^a & \ldots & c_{n-1}^a & c_n^a \end{pmatrix} \tag{3.40}$$

Effectively, each antenna transmits a time delayed version of a symbol stream. The corresponding trellis code for this scheme with QPSK modulation has four states and a constraint length of two. Input to the trellis code consists of two bits (representing a QPSK symbol) and the output consists of four bits (representing the pair of QPSK symbols transmitted by the two antennas). While the rate of this trellis code is $R = \frac{1}{2}$, two inputs $\rightarrow$ four output bits, the two antenna system transmits *two QPSK symbols* each time instant. Hence the overall system has $R = 1$. The key performance results are summarized in Table 3.2.

In Figure 3.5, the PEP upper bound and end to end simulation results of the DD coding scheme with QPSK modulation are evaluated for different DBC parameters. Simulation

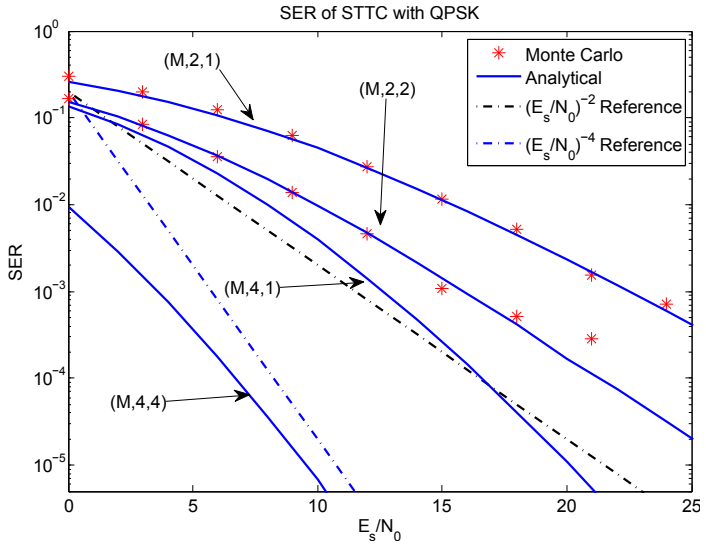


Figure 3.5: Monte Carlo evaluated PEP bound (3.14), denoted by the solid line, and end to end system Monte Carlo simulations of E_s/N_0 vs SER for STTC coded QPSK modulation, denoted by asterisks, where the delay diversity STTC is used. The simulated channel is defined by (3.7), where the receiver has perfect knowledge of the channel state.

results are only provided for the case of $L = 2$, because the decoding complexity increases exponentially in L. To calculate the upper bound, Monte Carlo evaluation over 10^6 iterations is used to average out the random variable in the conditional PEP (3.14). The codeword pair used to measure the codeword difference, i.e. $\mathbf{C} = (\mathbf{S}^a - \mathbf{S}^b)$, is chosen such that the codewords are nearest neighbors, which implies that the PEP and the SER are within a constant factor at high E_s/N_0. To validate the PEP bound of the DD code, the corresponding SER is obtained via Monte Carlo simulations of the DBC model described by (3.7). The input to the STTC encoder is a random bit sequence with each bit i.i.d. as equiprobable Bernoulli random variables. The decoder performs maximum likelihood (ML) detection, and it is assumed the receiver has perfect knowledge of the channel matrix, i.e. $\mathbf{H}^{rx}\mathbf{D}^{tx}$. The channel stays constant over a frame length of 200 symbol periods. The SER is averaged across 4000 frames. For the ML decoder, the Viterbi algorithm (VA) is used, [15] where the VA's decision metric is Euclidean distance and the traceback length used is 10

[15]The author's implemented their own Viterbi algorithm due to deficiencies of MATLAB's implementation.

Table 3.2: Performance Evaluation of STTC for the MIMO Dyadic Backscatter Channel

Channel	Coding Gain (SER = 10^{-3})	Diversity Order
$(M, 2, 1)$	0dB	2
$(M, 2, 2)$	8dB	2

times the code's constraint length.

Fig. 3.5 shows that the diversity order of the Monte Carlo simulation results match the diversity order of the PEP. As for OSTBC, the diversity order of STTC is determined solely by the number of tag antennas. The PEP result of the $L = 4$ DD code shows that its corresponding diversity order should be four. Lastly, the results in Fig. 3.5 demonstrates that increasing the number of receive antennas (N) results only in higher coding gains; suggesting that STTC based designs should follow the design recommendations similar to that of OSTBC.

3.6 Conclusion

This chapter presents OSTBCs and STTCs for the dyadic backscatter channel (DBC), with multiple antennas at RFID tag and reader. Our results show that the maximum diversity order is solely quantified by the number of tag antenna (L). However, the number of receive antennas (N) affects the coding gain and the SNR threshold at which the maximum diversity order is achieved. A RFID reader only needs two to four receive antennas and one transmit antenna; additional receive antennas provide diminishing gains. The most effective way to improve the DBC link reliability is to increase the number of tag antennas.

With regards to space-time code choice, OSTBCs achieve the maximum diversity order with *linear* decoding complexity, while STTCs achieve the same diversity order but have a much higher decoding complexity because STTCs require Viterbi decoding. Although both schemes do share linear encoding complexities, which can be implemented using digital state machines. So from the tag's perspective both code types are practical options, but decoder

cost at the reader remains a factor in the overall design. Both of these code designs can utilize standard QPSK or high order modulations schemes, which have been implemented on passive RFID tags [79].

Chapter 4

DIVERSITY MULTIPLEXING TRADEOFF FOR DYADIC BACKSCATTER CHANNEL

4.1 Introduction

Unlike the AWGN channel, the wireless fading channel cannot meet arbitrarily low probabilities of error by increasing a code's block length [18]. As such, the ability of any MIMO coding scheme to simultaneously achieve a desired code rate (multiplexing gain) and reliable communications (high diversity order) in a fading channel is bounded by the Diversity-Multiplexing Tradeoff (DMT) [32]. The characterization of the DMT curve is determined by the channel's fading statistics, so it is necessary to recalculate the DMT for new channel fading models. While the DMT for the (M, N) Rayleigh fading channel is fully characterized [32], the corresponding result for the (M, L, N) dyadic backscatter channel is presently unknown, to the best of our knowledge. Using the original problem formulation of the Rayleigh DMT, and results used to characterize the DMT of the double scattering MIMO channel [80], this chapter characterizes the DMT for the dyadic backscatter channel.

Recall the canonical (M, N) MIMO Gaussian channel is defined by

$$\mathbf{y} = \sqrt{\mathrm{SNR}}\mathbf{H}\mathbf{x} + \mathbf{w}, \tag{4.1}$$

where $\mathbf{w} \sim \mathcal{CN}(0, 1)$ and $\mathrm{Tr}\{E[\mathbf{x}\mathbf{x}^\dagger]\} = 1$ and the fading channel matrix $\mathbf{H}$ elements follow a suitable statistical characterization. With channel information only at the receiver, the capacity for a single channel instance of $\mathbf{H}$ is $\log\det(\mathbf{I} + \mathrm{SNR}\,\mathbf{H}\mathbf{H}^\dagger)$ [81]. For short range backscatter systems with a small number of antennas on the tag, it is reasonable for the receiver to acquire channel state through pilot symbols.

In the sequel, we assume the sender has no channel state knowledge and the receiver has full channel state knowledge. The sender transmits a codeword with rate $R(\mathrm{SNR})$ and incurs average probability of error $p_e(\mathrm{SNR})$. At high SNR, the performance is quantified

by two metrics, multiplexing gain (r) and diversity order (d) [32]

$$\lim_{\mathrm{SNR}\to\infty} \frac{R(\mathrm{SNR})}{\log \mathrm{SNR}} = r \quad \lim_{\mathrm{SNR}\to\infty} \frac{p_e(\mathrm{SNR})}{\log \mathrm{SNR}} = -d(r) \tag{4.2}$$

The optimal tradeoff is defined as $d^*(r) = \sup d(r)$, the supremum of the diversity order over all possible schemes achieving a multiplexing gain of r, [32]. In other words, for a system with $R(\mathrm{SNR}) \doteq r \log \mathrm{SNR}$, the corresponding probability of error is no better than $p_e(\mathrm{SNR}) \doteq \mathrm{SNR}^{-d^*(r)}$. The notation $\doteq$ denotes exponential equality, $f(x) \doteq g(x)$ if

$$\lim_{x\to\infty} \frac{\log f(x)}{\log x} = \lim_{x\to\infty} \frac{\log g(x)}{\log x}. \tag{4.3}$$

So for high SNR, $\mathrm{SNR}^{-d^*(r)}$ is an effective first order approximation for $p_e(\mathrm{SNR})$.[1]

Following the steps first taken by [32], the multiplexing gain is bounded by the ergodic channel capacity [81] of the fast fading MIMO channel

$$\mathcal{C}(\mathrm{SNR}) = E_{\mathbf{H}} \left[\log \det(\mathbf{I} + \mathrm{SNR}\ \mathbf{H}\mathbf{H}^\dagger) \right]. \tag{4.4}$$

At high SNR, the ergodic capacity can be approximated by $C(\mathrm{SNR}) \approx \min(M, N) \log \mathrm{SNR}$, and hence $r \leq \min(M, N)$. To achieve the ergodic capacity, a codeword must be averaged over many channel instances, which may not occur in practice with *coherent* narrowband systems. In a slow fading channel, the transmitted codeword experiences a single channel realization, thus the probability that a codeword with rate $r \log \mathrm{SNR}$ is in outage is given by [81]

$$p_{\mathrm{out}}(\mathrm{SNR}) = p\left(\log \det(\mathbf{I} + \mathrm{SNR}\ \mathbf{H}\mathbf{H}^\dagger) < r \log \mathrm{SNR} \right). \tag{4.5}$$

Corresponding to the diversity order, the outage probability exponent is defined by

$$\lim_{\mathrm{SNR}\to\infty} \frac{\log p_{\mathrm{out}}(\mathrm{SNR})}{\log \mathrm{SNR}} = -d_{\mathrm{out}}(r). \tag{4.6}$$

For a fixed multiplexing gain r, the relationship between the outage exponent and the diversity order is described next.

As discussed in [32, 80], suppose $\det(\mathbf{I} + \mathrm{SNR}\mathbf{H}\mathbf{H}^\dagger)$ has a pdf $p(\boldsymbol{\alpha})$ defined by a vector $[\alpha_1, \alpha_2, \ldots] = \boldsymbol{\alpha}$. If this pdf has a high SNR approximation

$$p(\boldsymbol{\alpha}) \doteq \mathrm{SNR}^{-\epsilon(\boldsymbol{\alpha})}, \tag{4.7}$$

[1] The operators $\dot{\leq}$ and $\dot{\geq}$ are defined similarly.

then $p_{\text{out}}(\text{SNR})$ can be approximated by the integral

$$\text{SNR}^{-d_{\text{out}}(r)} \doteq \int_{\mathcal{O}(\boldsymbol{\alpha},r)} \text{SNR}^{-\epsilon(\boldsymbol{\alpha})} \text{d}\boldsymbol{\alpha} \tag{4.8}$$

over the outage set $\mathcal{O}(\boldsymbol{\alpha}, r)$ such that $\log\det(\mathbf{I}+\text{SNR}\mathbf{H}\mathbf{H}^{\dagger}) < r\log\text{SNR}$ for a single instance of $\mathbf{H}$. For sufficiently large SNR, $d_{\text{out}}(r)$ can be approximated

$$d_{\text{out}}(r) = \inf_{\mathcal{O}(\boldsymbol{\alpha},r)} \epsilon(\boldsymbol{\alpha}). \tag{4.9}$$

via Laplace's method [82], as used by [32, 80]. This work assumes that the solution to (4.9) solves for the corresponding $d_{\text{out}}(r)$ for the DBC.

Note that the outage probability serves as a lower bound to the average probability of error, i.e., $p_{\text{out}}(\text{SNR}) \dot{\leq} p_e(\text{SNR})$, for a codeword with an arbitrary block length K [32]. Block length K is defined as the number of symbols periods required to transmit a space-time codeword, for example, the Alamouti code [58] has $K = 2$. For sufficiently large K, $p_{\text{out}}(\text{SNR}) \doteq p_e(\text{SNR})$, hence, $d_{\text{out}}(r) = d^*(r)$.

The remainder of this section are organized as follows. Sec. 4.2 calculates the DMT tradeoff function $d_{\text{out}}(r)$. Sec. 4.3 derives the block length requirements for $d_{\text{out}}(r) = d^*(r)$ and states the optimal tradeoff curve $d^*(r)$, which is the primary novel contribution of this chapter. Sec. 4.4 discusses how the DMT (an asymptotic metric) correlates with performance in the finite SNR regime.

4.2 Evaluation of $d_{out}(r)$

The signal model for the (M, L, N) DBC is

$$\mathbf{y} = \sqrt{\text{SNR}}\mathbf{H}^{\mathbf{rx}}\mathbf{D}^{\mathbf{tx}}\mathbf{x} + \mathbf{w},$$

where $\mathbf{H}^{\mathbf{rx}}\mathbf{D}^{\mathbf{tx}}$ is defined as in (3.7). Note that $\mathbf{x}$ is the signal backscattered by the RFID tags. Evaluating the spatial multiplexing gain r is straight forward. The matrix $\mathbf{H}^{\mathbf{rx}}\mathbf{D}^{\mathbf{tx}}$ has rank equal to $\min(L, N)$ with high probability, so in the high SNR limit the channel

capacity [32]

$$\begin{aligned}\mathcal{C}(\text{SNR}) &= E\left[\log\det(\mathbf{I}+\text{SNR}\ \mathbf{D^{tx\dagger}H^{rx\dagger}H^{rx}D^{tx}})\right]\\ &\approx \min(L,N)\log\text{SNR}+\sum_{i=1}^{\min(L,N)}E[\log\lambda_i]\\ &\doteq \min(L,N)\log\text{SNR},\end{aligned}\tag{4.10}$$

where the $\lambda_1 \geq \lambda_2 \geq \ldots$ are the non-negative ordered eigenvalues of $\mathbf{D^{tx\dagger}H^{rx\dagger}H^{rx}D^{tx}}$. Consequently, the multiplexing gain satisfies $0 \leq r \leq \min(L,N)$ Since the outage probability depends on the statistics of a random matrix, we first provide a necessary definition and lemma relating to random matrices.

Definition 4.2.1. *Define the $m \times m$ matrix $\mathbf{W} = \mathbf{GG}^\dagger$ where the columns of $\mathbf{G} \in \mathbb{C}^{m\times n}$ are i.i.d. $\mathcal{CN}(0,\mathbf{\Sigma})$ vectors. Then $\mathbf{W}$ follows the Wishart distribution $\mathcal{W}_m(n,\mathbf{\Sigma})$ [83].*

Lemma 4.2.2. *[80,83] Suppose $\mathbf{W} \sim \mathcal{W}_m(n,\mathbf{\Sigma})$. Define $a_1 \geq a_2 \geq \ldots \geq a_q \geq 0$ as the non-negative, ordered eigenvalues of $\mathbf{W}$ where $q = \min(m,n)$. Define $b_1 \geq b_2 \geq \ldots \geq b_m \geq 0$ as the ordered eigenvalues of $\mathbf{\Sigma}$. For $n \geq m$, the joint pdf $p(\mathbf{a})$ of the ordered eigenvalues of $\mathbf{W}$ is given by*

$$K_{m,n}Det\left[e^{-a_j/b_i}\right]_{i,j}\prod_{i=1}^{m}b_i^{m-n-1}a_i^{n-m}\prod_{i<j}^{m}\frac{a_i-a_j}{b_i-b_j}\tag{4.11}$$

and $p(\mathbf{a})$ for $m > n$

$$J_{m,n}Det(\mathbf{\Xi})\prod_{i<j}^{m}\frac{1}{b_i-b_j}\prod_{i<j}^{n}(a_i-a_j)\tag{4.12}$$

with $\mathbf{\Xi}$

$$\begin{bmatrix}1 & b_1 & \ldots & b_1^{m-n-1} & \frac{b_1^{m-n-1}}{exp(a_1/b_1)} & \cdots & \frac{b_1^{m-n-1}}{exp(a_n/b_1)}\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 1 & b_m & \ldots & b_m^{m-n-1} & \frac{b_m^{m-n-1}}{exp(a_1/b_m)} & \cdots & \frac{b_1^{m-n-1}}{exp(a_n/b_m)}\end{bmatrix}.$$

$K_{m,n}$ and $J_{m,n}$ are normalizing (positive) constants. Observe that the joint distribution of $\mathbf{a}$ depends only on the eigenvalues $\mathbf{b}$ of $\mathbf{\Sigma}$.

Following the analysis used by [32], the outage probability of the DBC is

$$p(\log\det(\mathbf{I}+\text{SNR}\ \mathbf{D^{tx\dagger}H^{rx\dagger}H^{rx}D^{tx}}) < r\log\text{SNR}).\tag{4.13}$$

Using well-known results for the matrix determinant -

$$\det(\mathbf{I} + \mathrm{SNR}\mathbf{D}^{\mathbf{tx}\dagger}\mathbf{H}^{\mathbf{rx}\dagger}\mathbf{H}^{\mathbf{rx}}\mathbf{D}^{\mathbf{tx}}) = \prod_{i=1}^{\min(L,N)} (1 + \mathrm{SNR}\lambda_i), \tag{4.14}$$

where λ_i is defined as before. Using the method described by [32], the eigenvalues of the channel matrix related to the outage probability for a given spatial multiplexing gain. Assume $\alpha_i = -\log \lambda_i / \log \mathrm{SNR}$, so that $\lambda_i = \mathrm{SNR}^{-\alpha_i}$ and $\alpha_1 \leq \alpha_2 \leq \ldots$. Each product term in the determinant can be written as $1 + \lambda_i \mathrm{SNR} \doteq \mathrm{SNR}^{(1-\alpha_i)^+}$ where $(x)^+ = \max(0, x)$. The outage probability in the high SNR regime is approximated by

$$p_{\mathrm{out}}(\mathrm{SNR}) \approx p\left(\prod_{i=1}^{\min(L,N)} \mathrm{SNR}^{(1-\alpha_i)^+} < \mathrm{SNR}^r \right). \tag{4.15}$$

Define $\mathcal{A} = \{\boldsymbol{\alpha} | \sum_{i=1}^{\min(L,N)} (1-\alpha_i)^+ < r, \alpha_1 \leq \alpha_2 \leq \ldots\}$ as the set of $\boldsymbol{\alpha}$ that cause an outage and $p(\boldsymbol{\alpha})$ as the joint pdf of the eigenvalues λ_i under the change of variables to α_i. The outage probability is given by the integral

$$p_{\mathrm{out}}(\mathrm{SNR}) = \int_{\mathcal{A}} p(\boldsymbol{\alpha}) \mathrm{d}\boldsymbol{\alpha}. \tag{4.16}$$

Evaluating the above integral in closed form is difficult because the distribution $p(\boldsymbol{\alpha})$ is unknown. Similar to the work on DMT of the double scattering channel [80], it is easier to find a joint distribution $p(\boldsymbol{\alpha}, \boldsymbol{\beta})$, whose pdf can be expressed in closed form, and solve for (4.16) over both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which avoids the explicit calculation of $p(\boldsymbol{\alpha})$.

For a fixed $\mathbf{D}^{\mathbf{tx}}$ and random $\mathbf{H}^{\mathbf{rx}\dagger}$, the columns of $\mathbf{G} = \mathbf{D}^{\mathbf{tx}\dagger}\mathbf{H}^{\mathbf{rx}\dagger}$ are iid Gaussian vectors with the correlation matrix $\mathbf{D}^{\mathbf{tx}}\mathbf{D}^{\mathbf{tx}\dagger}$. Therefore $\mathbf{G}\mathbf{G}^\dagger$, for fixed $\mathbf{D}^{\mathbf{tx}}$, generates a Wishart matrix (defn. 4.2.1) with distribution $\mathcal{W}_L(N, \mathbf{D}^{\mathbf{tx}}\mathbf{D}^{\mathbf{tx}\dagger})$, and $p(\boldsymbol{\lambda}|\mathbf{D}^{\mathbf{tx}})$ follows the distribution of the eigenvalues of a Wishart matrix given by lemma 4.2.2.

Define the ordered eigenvalues of $\mathbf{D}^{\mathbf{tx}}\mathbf{D}^{\mathbf{tx}\dagger}$ as $\mu_1 > \mu_2 > \ldots > \mu_L > 0$. Let $\beta_i = -\log \mu_i / \log \mathrm{SNR}$, then $p(\boldsymbol{\alpha}|\boldsymbol{\beta}) \sim \mathcal{W}_L(N, \mathbf{D}^{\mathbf{tx}}\mathbf{D}^{\mathbf{tx}\dagger})$ under the proper change of variables. Combining the known results on the distribution of $p(\boldsymbol{\alpha}|\boldsymbol{\beta})$ [80] and deriving the distribution of $p(\boldsymbol{\beta})$, the joint distribution $p(\boldsymbol{\alpha}, \boldsymbol{\beta})$ can be expressed in closed form. Using Bayes rule, the outage probability (4.16) is equivalently written as [80]

$$p_{\mathrm{out}}(\mathrm{SNR}) = \int_{\mathcal{A}\cap\mathcal{B}} p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta} \mathrm{d}\boldsymbol{\alpha}. \tag{4.17}$$

where $\mathcal{B} = \{\beta_1 \leq \ldots \leq \beta_L\}$. (4.9) will be used to approximate this integral in the high SNR regime. Next we find $\epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta})$ such that $p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \doteq \mathrm{SNR}^{-\epsilon(\boldsymbol{\alpha},\boldsymbol{\beta})}$.

Since $\mathbf{D^{tx}D^{tx\dagger}}$ is diagonal matrix, the eigenvalues equal the diagonal elements. Each element, $D_{ii}^{tx} D_{ii}^{tx*}$ follows the exponential distribution with parameter 1 and are iid, so the ordered eigenvalue distribution is

$$p(\boldsymbol{\mu}) = L! \exp\left[-\sum_{i=1}^{L} \mu_i\right], \tag{4.18}$$

and under a change of variables, the high SNR approximation to the pdf in terms of $\boldsymbol{\beta}$ is

$$p(\boldsymbol{\beta}) \doteq \mathrm{SNR}^{-\sum_{i=1}^{L} \beta_i} \tag{4.19}$$

for all non-negative, i.e. $\beta_i \geq 0$, otherwise $p(\boldsymbol{\beta}) \to 0$ exponentially.

$d_{out}(r)$ - $N \geq L$ Case:

Lemma 4.2.3. *For $N \geq L$, the joint distribution $p(\boldsymbol{\alpha}, \boldsymbol{\beta})$ has the high SNR approximation*

$$\begin{aligned} p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \doteq & SNR^{\sum_{i=1}^{L} -(L-N)\beta_i - (N-L+1)\alpha_i} \\ & \cdot SNR^{\sum_{i=1}^{L} -(L-i)(\alpha_i - \beta_i)} SNR^{\sum_{i<j}^{L} -(\alpha_i - \beta_j)^+} \end{aligned} \tag{4.20}$$

if $\alpha_i \geq \beta_i, \boldsymbol{\alpha} \geq 0, \boldsymbol{\beta} \geq 0$ - otherwise $p(\boldsymbol{\alpha}, \boldsymbol{\beta})$ decays to zero exponentially.

Proof. See Appendix C.1. □

Since (4.20) decays exponentially if the constraints in lemma 4.2.3 are not met, that range of values can be excluded from the outage integrand (4.17) without affecting its high SNR approximation [32, 80]; this reduces the sets $\mathcal{A}$ and to $\mathcal{B}$ to $\mathcal{A}' = \{\boldsymbol{\alpha} | \sum_{i=1}^{L} (1 - \alpha_i)^+ < r, 0 \leq \alpha_1 \leq \ldots \leq \alpha_L\}$ and $\mathcal{B}' = \{\boldsymbol{\beta} | \forall i,\ \beta_i \leq \alpha_i,\ \beta_1 \leq \ldots \leq \beta_L,\ \beta_i \geq 0\}$, respectively. Using (4.20), we may write $p(\boldsymbol{\alpha}, \boldsymbol{\beta}) \doteq \mathrm{SNR}^{-\epsilon(\boldsymbol{\alpha},\boldsymbol{\beta})}$ where

$$\epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \sum_{i=1}^{L} (i - N)\beta_i + (N + 1 - i)\alpha_i + \sum_{i<j}^{L} (\alpha_i - \beta_j)^+, \tag{4.21}$$

Using (4.9) the outage probability (4.17) is evaluated by $\mathrm{SNR}^{-d_{\mathrm{out}}(r)}$, where the exponent $d_{\mathrm{out}}(r)$ is obtained as the result of the following optimization

$$d_{\mathrm{out}}(r) = \inf_{\mathcal{O}(\boldsymbol{\alpha},\boldsymbol{\beta},r)} \epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta}) \tag{4.22}$$

and $\mathcal{O}(\boldsymbol{\alpha},\boldsymbol{\beta},r) = \mathcal{A}' \cap \mathcal{B}'$. Observe that $\epsilon(\boldsymbol{\alpha},\boldsymbol{\beta})$ is decreasing for increase in β_i, and recall that $\beta_i \leq \alpha_i$. So for any $\boldsymbol{\alpha}$, $\epsilon(\boldsymbol{\alpha},\boldsymbol{\beta})$ is minimized by choosing $\beta_i = \alpha_i$. Rewriting the optimization in terms of only $\boldsymbol{\alpha}$ yields

$$d_{\text{out}}(r) = \inf \sum_{i=1}^{L} \alpha_i$$
$$\text{subject to: } \sum_{i=1}^{L} (1-\alpha_i)^+ < r, \;\; 0 \leq \alpha_1 \leq \ldots \leq \alpha_L \tag{4.23}$$

The minimum occurs when $\sum_{i=1}^{L}(1-\alpha_i)^+ = r$, or equivalently $L - r = \sum_{i=1}^{L} \alpha_i$ given $0 \leq \boldsymbol{\alpha} \leq 1$. The corresponding outage exponent is

$$d_{\text{out}}(r) = L - r. \tag{4.24}$$

$d_{out}(r)$ - $L > N$ Case:

We employ the same strategy as the $N \geq L$ case to find $d_{\text{out}}(r)$.

Lemma 4.2.4. *For $L > N$, the joint distribution $p(\boldsymbol{\alpha},\boldsymbol{\beta})$ has the high SNR approximation*

$$\begin{aligned} p(\boldsymbol{\alpha},\boldsymbol{\beta}) \doteq & SNR^{-\sum_{i=1}^{N+1}(L-N-1)\beta_i} SNR^{-\sum_{i=1}^{L}\beta_i} SNR^{-\sum_{i=1}^{N}\alpha_i} \\ & \cdot SNR^{-\sum_{i=N+2}^{L}(L-i)\beta_i} SNR^{\sum_{i<j}^{L}\beta_i} SNR^{-\sum_{i<j}^{N}\alpha_i} \\ & \cdot SNR^{-\sum_{i<j}^{N}(\alpha_j-\beta_i)^+} SNR^{-\sum_{i=1}^{N}\sum_{N+1}^{L}(\alpha_i-\beta_j)^+} \end{aligned} \tag{4.25}$$

if $\forall 1 \leq i \leq N,\ \alpha_i \geq \beta_i,\ \boldsymbol{\alpha} > 0,\ \boldsymbol{\beta} > 0$ - otherwise $p(\boldsymbol{\alpha},\boldsymbol{\beta})$ decays to zero exponentially.

Proof. See Appendix C.2. □

Since (4.25) decays exponentially if the constraints in lemma 4.2.4 are not met, that range of values can be excluded from the outage integrand (4.17) without affecting its high SNR approximation - reducing the sets $\mathcal{A}$ and to $\mathcal{B}$ to $\mathcal{A}' = \{\boldsymbol{\alpha} | \sum_{i=1}^{L}(1-\alpha_i)^+ < r, 0 \leq \alpha_1 \leq \ldots \leq \alpha_N\}$ and $\mathcal{B}' = \{\boldsymbol{\beta} | \forall 1 \leq i \leq N,\ \beta_i \leq \alpha_i,\ \beta_1 \leq \ldots \leq \beta_L,\ \beta_i \geq 0\}$.

As before, the terms in the exponent of (4.25) in lemma 4.2.4 can be expressed as the function

$$\begin{aligned} \epsilon(\boldsymbol{\alpha},\boldsymbol{\beta}) = & \sum_{i=1}^{N}(N-i+1)\alpha_i + \sum_{i=1}^{N+1}(i-N)\beta_i + \sum_{i=N+2}^{L}\beta_i \\ & + \sum_{i<j}^{N}(\alpha_i-\beta_j)^+ + \sum_{i=1}^{N}\sum_{j=N+1}^{L}(\alpha_i-\beta_j)^+ \end{aligned} \tag{4.26}$$

$d_{\text{out}}(r)$ is given by the solution to the optimization problem (4.22) over the outage set $\mathcal{O}(\boldsymbol{\alpha}, \boldsymbol{\beta}, r) = \mathcal{A}' \cap \mathcal{B}'$. To solve this optimization problem, first, find the $\boldsymbol{\beta}$ that minimizes $\epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta})$ for any fixed $\boldsymbol{\alpha}$, and finally, solve for the optimal $\boldsymbol{\alpha}$.

The components of $\epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta})$ that depend on $\boldsymbol{\beta}' = [\beta_{N+1}, \ldots, \beta_L]^{\mathrm{T}}$ are

$$\sum_{i=N+1}^{L} \beta_i + \sum_{i=1}^{N} \sum_{j=N+1}^{L} (\alpha_i - \beta_j)^+. \tag{4.27}$$

For any $\boldsymbol{\alpha}$, it will be shown that $\boldsymbol{\beta}' = \alpha_N \mathbf{1}_{L-N}$ minimizes the above.[2] Suppose $\beta_L \leq \alpha_N$, then the contribution of the β_L terms to (4.27) is

$$\beta_L + \sum_{i=1}^{N} (\alpha_i - \beta_L)^+, \tag{4.28}$$

otherwise for $\beta_L > \alpha_N$, the above reduces to β_L. To minimize the above it is necessary that $\sum_{i=1}^{N-1} (\alpha_i - \beta_L)^+ = 0$, and due to the ordering of α_i, $\beta_L = \alpha_N$ is the minimization solution. The same logic can be recursively applied to $\beta_{L-1}, \ldots, \beta_{N+1}$ to show that $\boldsymbol{\beta}' = \alpha_N \mathbf{1}_{L-N}$ is sufficient to minimize (4.27).

Substituting $\boldsymbol{\beta}' = \alpha_N \mathbf{1}_{L-N}$ back into the original equation results in

$$\begin{aligned} \epsilon(\boldsymbol{\alpha}, \beta_1, \ldots, \beta_N) = & \sum_{i=1}^{N} (N - i + 1)\alpha_i + (L - N)\alpha_N + \\ & \sum_{i=1}^{N} (i - N)\beta_i + \sum_{i<j}^{N} (\alpha_i - \beta_j)^+ \end{aligned} \tag{4.29}$$

As in $N \geq L$ case, ϵ is decreasing in increasing β_i. Therefore, choose $\alpha_i = \beta_i$, $\forall 0 < i \leq N$. The reduced minimization problem in terms of only $\boldsymbol{\alpha}$ is now

$$\begin{aligned} d_{\text{out}}(r) &= \inf \sum_{i=1}^{N-1} \alpha_i + (L - N + 1)\alpha_N \\ \text{Subject to: } & \sum_{i=1}^{N} (1 - \alpha_i)^+ < r, \;\; 0 \leq \alpha_1 \leq \ldots \leq \alpha_N \end{aligned} \tag{4.30}$$

In the objective, the term with the largest coefficient (α_N) should be minimized first before reducing the remaining terms. To satisfy the ordering constraints of α_i while minimizing

[2] $\mathbf{1}_N$ represents an all 1 column vector with dimension N.

α_N, it is necessary to set $\alpha_i = \alpha_N \forall i \in [1, N-1]$. Setting $\forall i\ \alpha_i = a$, a dummy variable where $0 < a$, the objective is minimized when the constraints are met with equality i.e. $N(1-a)^+ = r$. The corresponding solution is

$$d_{\text{out}}(r) = L(1 - r/N). \tag{4.31}$$

4.3 DBC DMT

Codeword Length Considerations:

The $d_{\text{out}}(r)$ serves as an upper bound to the optimal tradeoff $d^*(r)$. To show that $d_{\text{out}}(r) = d^*(r)$ for certain cases, we apply the union bound method used by [32] in the Rayleigh DMT calculation. The error event can be expressed as the union of two exclusive events

$$p_e = p(\text{error}|\text{outage})p_{\text{out}} + p(\text{error and no outage}), \tag{4.32}$$

Since the goal is to prove $p_e \dot{\leq} p_{\text{out}}$, consider the following upper bound

$$p_e \leq p_{\text{out}} + p(\text{error and no outage}). \tag{4.33}$$

An error occurs from either the channel being in outage or the additive noise causes an error despite the channel not being in outage. The latter can be driven to an arbitrarily small value by extending a code's block length towards infinity, however, for a finite block length, p(error, no outage) may dominate the outage probability p_{out}.

Since an outage event does not occur, i.e. $\log\det(\mathbf{I} + \text{SNR}\mathbf{H}\mathbf{H}^\dagger) > r\log\text{SNR}$, the multiplexing gain is constrained by $r > \sum_i (1-\alpha_i)^+$. Define the no outage set as $\mathcal{A}'' = \{\boldsymbol{\alpha} | \sum_{i=1}^{L}(1-\alpha_i)^+ > r, 0 < \alpha_1 \leq \ldots \leq \alpha_{\min(L,N)}\}$, similar to the definition of $\mathcal{A}'$. Using the joint distribution of $p(\boldsymbol{\alpha}, \boldsymbol{\beta})$, the p(error and no outage) is bounded by [32]

$$p(\text{error and no outage}) \dot{\leq} \int_{\mathcal{A}''\cap\mathcal{B}'} \text{SNR}^{-\epsilon(\boldsymbol{\alpha},\boldsymbol{\beta})}\text{SNR}^{-K[\sum(1-\alpha_i)^+ - r]}\mathrm{d}\boldsymbol{\beta}\mathrm{d}\boldsymbol{\alpha} \tag{4.34}$$

As in [32], this integral is evaluated using (4.9), by finding the $\boldsymbol{\alpha}, \boldsymbol{\beta}$ that minimize the

exponent given that the channel state is such that the system is not in outage, i.e.

$$d_{\mathrm{U}}(r) = \inf \ \epsilon(\boldsymbol{\alpha}, \boldsymbol{\beta}) + K[\sum_{i}(1-\alpha_i)^+ - r]$$

$$\text{Subject to: } \alpha_i \geq \beta_i; \ \ 0 \leq \alpha_1 \leq \alpha_2 \leq \ldots$$

$$0 \leq \beta_1 \leq \beta_2 \leq \ldots; \quad \sum_{i=1}^{\min(L,N)} (1-\alpha_i)^+ > r \tag{4.35}$$

As before, this optimization can be written in terms of only $\boldsymbol{\alpha}$ by finding the optimal $\boldsymbol{\beta}$ terms first.

Case $\underline{N \geq L}$, the minimization

$$\inf \sum_{i=1}^{L} \alpha_i + K[\sum_{i=1}^{L}(1-\alpha_i)^+ - r] \tag{4.36}$$

subject to the constraints in (4.35). If $K \geq 1$, then the minimum occurs when $\sum_{i=1}^{L}(1-\alpha_i)^+ - r = 0$. Therefore $d_{\mathrm{U}}(r) = L - r$, i.e. the same solution as the outage exponent (4.24).

Case $\underline{L > N}$, we have

$$\inf \sum_{i=1}^{N-1} \alpha_i + (L-N+1)\alpha_N + K[\sum_{i=1}^{N}(1-\alpha_i)^+ - r] \tag{4.37}$$

subject to the constraints in (4.35). If $K \geq \lceil L/N \rceil$, then the minimum occurs when $\sum_{i=1}^{N}(1-\alpha_i)^+ - r = 0$, which gives the same solution as the outage exponent (4.31).

Returning to the upper bound on error probability (4.33) it can be seen that if the code block length K meets the above constraints, then the upper and lower bounds are tight, per below [32]:

$$\begin{aligned} p_e(r) &\dot{\leq} \mathrm{SNR}^{-d_{\mathrm{out}}(r)} + \mathrm{SNR}^{-d_{\mathrm{U}}(r)} \\ &\doteq \mathrm{SNR}^{-d_{\mathrm{out}}(r)} \end{aligned} \tag{4.38}$$

since $d_{\mathrm{out}}(r) = d_{\mathrm{U}}(r)$. Hence, the probability of error can be upper and lower bounded by $d_{\mathrm{out}}(r)$, which proves $d^*(r) = d_{\mathrm{out}}(r)$. Our desired result is summarized in the following theorem 4.3.1. Note that the code length constraint must be met for $d^*(r) = d_{\mathrm{out}}(r)$.

Main Result and Discussion:

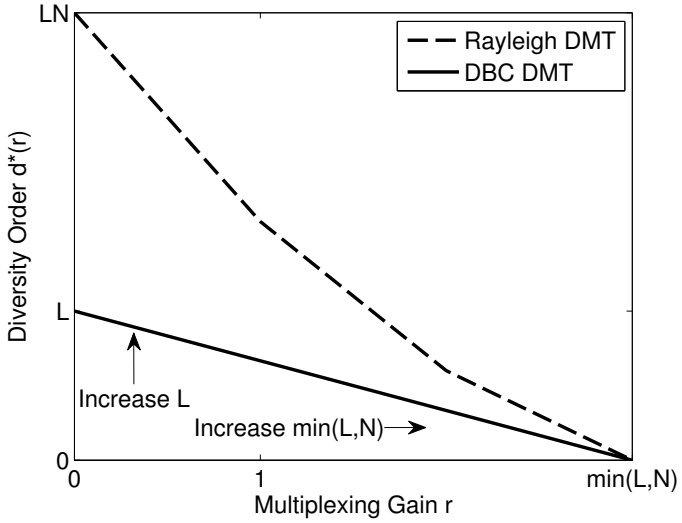


Figure 4.1: (M, L, N) DBC and (L, N) Rayleigh Channel Diversity Multiplexing Tradeoff - the curves are plotted for $L = 5, N = 3$.

Theorem 4.3.1. *For a (M, L, N) dyadic backscatter channel with code length K, the optimal DMT curve $d^*(r)$ is given by the linear function*

- *$N \geq L$ and $K \geq 1$*

$$d^*(r) = L - r, \quad 0 < r \leq L \tag{4.39}$$

- *$N < L$ and $K \geq \lceil L/N \rceil$*

$$d^*(r) = L(1 - r/N), \quad 0 < r \leq N \tag{4.40}$$

Proof. The diversity order exponent $d_{\text{out}}(r)$ is derived, for both cases, in Sec 4.2 at (4.24) and (4.31). The argument posed by (4.38) establishes the block length restrictions such that $d_{\text{out}}(r) = d^*(r)$. □

Fig. 4.1 illustrates the tradeoff for the two cases, and as comparison, it plots the Rayleigh DMT for a system with the same antenna configuration. Effectively, the DBC DMT is a linear function connecting the points L and $\min(N, L)$ on the $d(r)$ and r axises. For comparison, the (L, N) Rayleigh channel DMT is characterized by the piece-wise linear

function with the corner points as $d^*(r) = (L - r)(N - r)$ for integer values of r; see theorem 2 of [32]. The results reveals some interesting new characteristics - while the PEP bounds and diversity order results in Chapter 3 suggest that an increase in receive antennas N has no asymptotic benefits, the DMT proves this is not the case. First for $N < L$, increasing N has multiple effects. It improves the maximum multiplexing gain while increasing the reliability $d^*(r)$ for $0 < r < N$. So while the diversity for $r = 0$ remains unchanged, it improves for all other r. Additionally, increasing N reduces the length of the space-time code required to achieve the optimal DMT. This explains why diversity gains are feasible in the DBC without any space-time codes, because for $N \geq L$, coding only over space (antennas) is sufficient to achieve the optimal $d^*(r)$. These results agree with prior work [60] that reports a maximum diversity of $\min(N, L)$ for space only codes - for $L > N$ a (M, L, N) system cannot have worse performance than a (M, N, N) system.

Increasing N beyond L appears to only have a minor impact in system performance; diversity or spatial multiplexing gains remain bounded by L. This is in clear contrast to the Rayleigh fading channel, where increasing N provides substantial diversity order gains (up to LN) for small values of r. While the DBC can offer the same range of multiplexing gains as the Rayleigh channel, the DBC lacks capability to offer the same level of reliability. Essentially, the DBC is more likely than its Rayleigh counterpart to encounter fading, but DBC does not suffer from the loss of rank as found in keyhole or pinhole channels. These DMT results suggest that a system with $L = N$ will strike a balance in performance (diversity and multiplexing) and complexity/cost, which also agrees with the recommendations by [60].

Lastly, the case for which $d^*(r)$ is not solved for is $L > N$ with $K < \lceil L/N \rceil$ - we highlight this as a case that will not occur commonly in practice. Due to space constraints on a RFID tag, it is unlikely for $L > N$ to occur in most situations, and the code length constraint remains small for any practical occurrence of $L > N$. This characteristic is different than (M, N) Rayleigh case, where the code length increases as $M + N - 1$ [32]. So deriving the remaining case for $d^*(r)$ may not offer much practical guidance in choosing L and N.

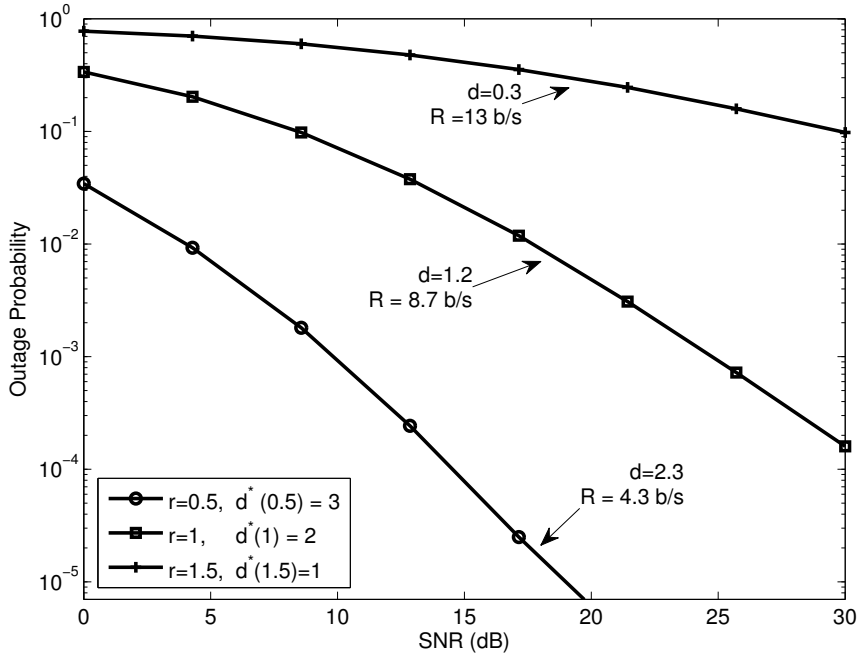


Figure 4.2: Monte carlo evaluation of (4.42) for the $(M, 4, 2)$ DBC. The slope of the diversity curve and $r\log(1+g\text{SNR})$ is evaluated at 17.1dB for different multiplexing gains r, and the corresponding optimal DMT is denoted by $d^*(r)$.

4.4 DBC DMT - Finite SNR Regime

While the DMT is originally defined as an asymptotic metric, it is still useful to quantify the behavior of the multiplexing and diversity order gains at *realistic* values of SNR. Existing work [84] on the Rayleigh and Rician MIMO fading channels derive a practical lower bound to the outage probability with respect to the multiplexing gain. Coupled with simulations to verify the bound, these results indicate a significant gap between the asymptotic DMT curve and what can be achieved in practice. This section builds on the results of [84] and formulates the finite SNR DMT for the DBC and provides results derived from simulations, however, developing a closed form to the lower bound is outside of the scope of this manuscript.

The multiplexing gain and diversity order have the following corresponding metrics in

finite SNR [84]:

$$r = \frac{R(\mathrm{SNR})}{\log(1+g\mathrm{SNR})} \quad d(r,\mathrm{SNR}) = -\frac{\partial \log p_{\mathrm{out}}(\mathrm{SNR})}{\partial \log \mathrm{SNR}}. \tag{4.41}$$

For small SNR, $\det\log(\mathbf{I}+\mathrm{SNR}\mathbf{H}\mathbf{H}^{\dagger}) \approx \log(1+\mathrm{SNR}\|\mathbf{H}\|_F^2)$, so a fair value to choose for the array gain is $g = E[\|\mathbf{H}\|_F^2] = LN$ [84]. The outage probability is redefined slightly,

$$p(\det\log(\mathbf{I}+\mathrm{SNR}\mathbf{H}\mathbf{H}^{\dagger}) < r\log(1+g\mathrm{SNR})), \tag{4.42}$$

in order to account for the new definition of multiplexing gain.

Monte Carlo evaluations of (4.42) over a range of SNR and for various r represent the main new contributions of this section. Fig. 4.2 plots the outage probability for the $(M,4,2)$ DBC for various multiplexing gains r. Similar to the Rayleigh case [84], only a fraction of the diversity order is achievable at reasonable SNR, and the diversity order worsens as the multiplexing gain is increased. So to achieve high spectral efficiency with high reliability, backscatter communication will require significant time or frequency diversity; both of which may be a challenge for narrow band backscatter systems. Lastly, these results confirm the conclusion of the asymptotic DMT analysis - that DBC is primarily limited by reliability (vulnerability to fading) rather than code rate (multiplexing gain).

4.5 Conclusion

The diversity-multiplexing tradeoff offers new insight on the performance of MIMO RFID systems. Compared to the Rayleigh MIMO channel, the DBC supports the same range of spatial multiplexing gains but with substantially less diversity (reliability). Increasing the number of tag antennas always improves the diversity, and it also improves the multiplexing gain if $L < N$. Previous analysis of the DBC focuses mainly on the maximum diversity order, which suggests the the number of receive antennas has little impact on performance. To the contrary, the DMT indicates improved diversity for any increase in N if $N < L$ and the multiplexing gain is non-zero. The minimum space-time code length required by the DMT also decreases as N increases. For $N \geq L$, space only codes are sufficient to achieve the entirety of the DMT curve.

There still remains open questions with respect to RFID MIMO systems, namely, how to code the reader's continuous wave (CW) transmission to improve diversity. The assumption of constant CW transmission results in the statistical equivalence of the $(1, L, N)$ and (M, L, N) dyadic backscatter channels. It would be interesting to see if coding the CW transmission (without channel state knowledge) can improve diversity.

Chapter 5

HYBRID-ARQ IMPROVEMENTS FOR EPC GEN2

5.1 Introduction

Compared to most modern wireless communication standards (e.g. cellular and WiFi), the Gen2 protocol is relatively simplistic in its design. During nominal operation, RFID tags are first identified with a frame slotted ALOHA (FSA) [28] random access scheme. A RFID reader transmits a query command, and each RFID tag will randomly choose whether to respond with a 16 bit random number (RN16) or to stay silent. After an error free response by a tag, the reader ACKs the RN16 by retransmitting the sequence. The corresponding tag responds with its electronic product code (EPC, up to 528 bits in length) or payload. The reliability of the tag's EPC is insured by a cyclic redundancy check (CRC) and a stop and wait Automatic Repeat reQuest (ARQ) protocol. If the CRC fails, the reader ACKs the RN16 again, and the tag retransmits its payload. If the RN16 is decoded incorrectly or a collision occurs, no attempt to recover is made; the reader progresses to the next slot. Fig.5.1 illustrates this process as a state transition diagram.

Improving the efficiency of either the FSA random access scheme or the EPC transmission will increase the average tag read rate. Much of the existing work has focused on incremental improvements to the FSA algorithm [85], and previously there has only been little attention paid to the physical and media access control (MAC) layer design of the EPC transmission. A simple retransmission scheme using BCH codes is proposed by [86], but fails to address signaling overhead and uses only hard decision decoding. Varying the tag's symbol rate to control the link quality is proposed by [87], but no error correction is considered. Other work considers sophisticated multiuser detection and error correction based on belief propagation to improve link throughput [88], however, their system differs greatly from Gen2 and focuses on experimental results.

In parallel to this work, the application of coding to RFID systems has received increased

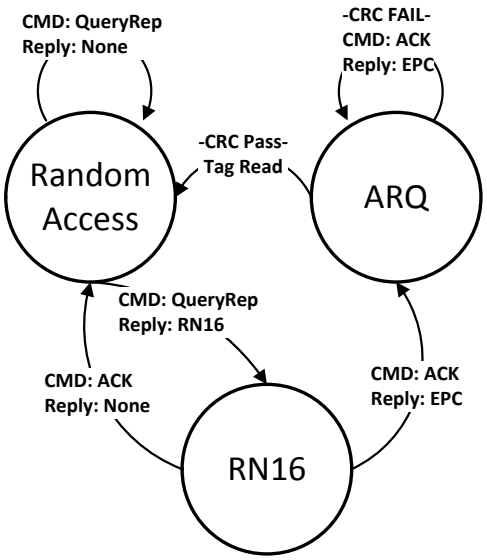


Figure 5.1: State transition diagram of EPC Gen2 RFID tag singulation.

attention. The application of the Alamouti space-time codes [58] concatenated with BCH codes offer both improved diversity and coding gains [89] as compared to only coding gains for a single antenna tag [90]. The work most similar to this chapter is [91],[1] which proposes and analyzes ARQ and HARQ schemes based on forward error correction (FEC) codes similar to those discussed in the remainder of the chapter. Nearly all the work on applying FEC to RFID treats the FM0/Miller modulation and the FEC code as independent, however, iterative joint decoding using the Turbo principle offers modest coding gains [92]. Lastly even for the standard EPC Gen2 system, the CRC-16 appended to each EPC message has the capability to correct single bit errors and some double bit errors [93], which has immediate practical use because it requires no changes the Gen2 protocol.

The purpose of this chapter is to show low complexity and small improvements, in regards to FEC and ARQ, to the EPC message of the Gen2 protocol, which result in increased read rates and robustness to varying SNR. The proposed changes to Gen2 have a limited scope, and the FEC encoder/decoder algorithms used in this work are mature and widely used in industry, so modifications are ready for implementation in future tags and readers.

Hybrid ARQ (HARQ) [94, 95] is a subfield in communication theory that studies the

[1]This work was developed parallel and independently of our own.

codesign of FEC and ARQ algorithms. In general there are two types of HARQ, *chase combining* (CC) [96] and *incremental redundancy* (IR) [97]. CC is the maximum likelihood (ML) combining of identical ARQ retransmissions. Effectively, CC behaves as a simple repetition code. IR sends additional parity bits of a FEC on each ARQ retransmission, so an incrementally lower rate code is used with each retransmission. These IR FEC codes are constructed using a coding technique called rate compatible puncturing (RCP) [98], which allows for efficient encoder and decoder constructions. We investigate the benefit of HARQ CC and IR methods with respect to the baseline Gen2 protocol, and describe what protocol changes are needed for RFID to use HARQ. Our innovation is demonstrating how to use existing digital communication techniques that are practical and beneficial for RFID systems.

5.2 System Model and Baseline Performance Analysis

5.2.1 Gen2 System Model

This section develops a mathematical model that captures EPC Gen2's main modes of operation. The EPC Gen2 protocol is reader driven, where a reader command is transmitted for a tag to respond. A delimiter, a frame sync, and a variable length sequence of symbols comprise a reader command message. Each of these elements are generated by Pulse Interval Encoding (PIE) of the reader's continuous wave (CW).[2] Within a fixed amount of time following a reader's command, a tag chooses whether to respond to the reader's command based on the state transition diagram in Fig 5.1. A tag's response consists of a known preamble sequence with good autocorrelation and a binary encoded (Miller sub-band or FM0) information sequence, which depending on the reader's command is the RN16 or EPC message. We assume the tag's preamble is sufficient to detect the presence of a tag response over a wide range of SNR. Based on the tag's response, the reader decides on which command to issue next. We assume that the reader's commands are always decoded error free by all tags.

[2]We defer the specifics of these elements to the Gen2 standard document, because they remain unchanged in this work.

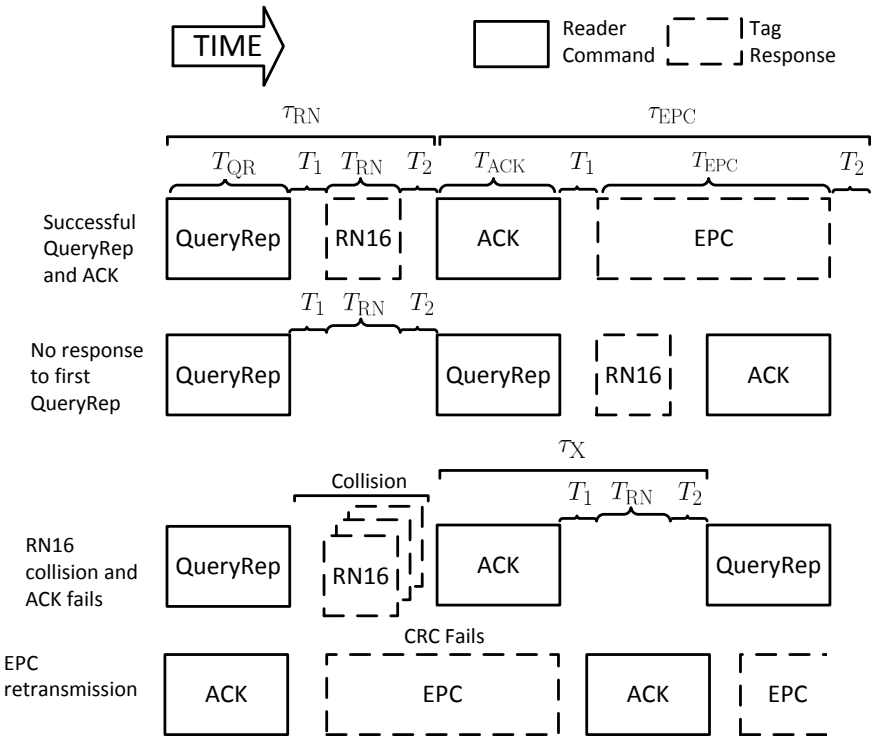


Figure 5.2: Timing diagram for 4 key scenarios: tag read success, empty RN16 slot, RN16 collision, and a EPC retransmission. The time periods listed above are derived from the Gen2 standard document.

Since this work is interested in the average tag read rate, we shall only consider two reader commands (QueryRep and ACK), and the two corresponding tag responses (RN16 and EPC response). Before requesting a tag's EPC with the ACK command, the reader must first identify the tag through a FSA process. Multiple tags might experience a collision during their QueryRep response, because FSA is a random contention based algorithm. To simplify analysis, we assume collisions always result in an incorrectly decoded RN16. After decoding a RN16, the reader echoes the decoded random 16 bit sequence with a ACK command. If the RN16 message was decoded error free by the reader, then the corresponding tag responds with its EPC message.[3] The payload integrity is checked by a 16 bit CRC, and if the CRC fails the reader can request for a retransmission from the tag by repeating its ACK

[3]The EPC message consists of a 16 bit protocol control (PC) bits, 496 bit stored product code, and a CRC-16.

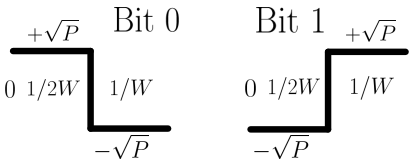


Figure 5.3: Backscattered waveforms from the tag for coherent BPSK signaling with power P. The signaling frequency is W and The SNR of either symbol is $\frac{P}{WN_0}$.

command. Fig. 5.2 illustrates multiple scenarios of the tag and reader query processes.

The time between a reader command and a tag response (T_1) and the next reader command (T_2) are taken from the response times stated in the Gen2 standard; the lengths are 10 and 20 tag symbol periods respectively. For the scenarios where a tag fails to respond, it is assumed that it takes the reader T_{RN} seconds (length of an RN16) to correctly detect the absence of the tag. To simplify notation, the time for certain reader and tag actions are denoted as follows. The time for the reader to transmit a QueryRep and the time to allow for a tag to respond is

$$\tau_{\text{RN}} = T_{\text{QR}} + T_1 + T_{\text{RN}} + T_2. \tag{5.1}$$

The value is the same for a collision free slot, a slot with collisions, and an empty slot. The time for a reader to transmit the ACK[4] and the tag to respond is

$$\tau_{\text{EPC}} = T_{\text{ACK}} + T_1 + T_{\text{EPC}} + T_2. \tag{5.2}$$

Lastly, the time for the reader to transmit an incorrect ACK and to detect the lack of a response by a tag is

$$\tau_{\text{X}} = T_{\text{ACK}} + T_1 + T_{\text{RN}} + T_2. \tag{5.3}$$

The specific values of these reader and tag response times are derived from the Gen2 standard document.[5]

To offer different bit error of performance across a range of SNR, the tag must vary its symbol rate. Lower symbol rates allow the reader's receiver to reduce the received

[4]The ACK response by the reader is variable in time because of the different lengths of the '0' and '1' PIE symbols. So we use the averaged length over all possible 2^{16} ACK sequences.

[5]This analysis considers the link timings for uplink symbol rates from 58.125 KHz to 465 KHz.

noise power but at the expense of raw throughput. Conversely, increasing the symbol rate improves the raw throughput but the received noise power increases. The two types of binary encoding used by Gen2, Miller and FM0, have similar performance when normalized against their energy per symbol, and their performance is comparable to differential binary phase shift keying (DBPSK) and orthogonal signaling [33]. To further simplify the model, we shall instead assume coherent antipodal modulation binary phase shift keying (BPSK); Fig. 5.3 defines the bit '1' and '0' waveforms. We assume that the channel state is perfectly estimated through the tag's preamble sequence, so differential modulation is unnecessary.

The system SNR is defined with respect to an operating frequency of 465KHz (TRCal = 17.2μs), where SNR $= \frac{P}{W'N_0}, W' = 465$KHz and P is the fixed signal power at the reader's receiver.[6] Thus, increasing or decreasing the symbol frequency W manifests as a linear SNR gain defined by $\gamma = W'/W$. The received signal at the reader can be readily modeled as a BPSK with additive white Gaussian noise (AWGN),

$$y = \sqrt{\gamma \text{SNR}} x + w \tag{5.4}$$

where $y \in \mathbb{C}$, $x \in \{-1, +1\}$, $w \sim \mathcal{CN}(0, 1)$. The BER for an uncoded signal is[7]

$$p_e = \mathcal{Q}(\sqrt{2\gamma \text{SNR}}). \tag{5.5}$$

The remainder of this work shall use (5.4) as the received signal model.

5.2.2 Baseline Throughput Analysis

The figure of merit is the average number of tags read per second at a fixed SNR value. Since the Gen2 standard dictates a fixed symbol rate across all tags, the reader cannot dynamically adjust the link's reliability to varying channel conditions. Consequently, only the choice of symbol frequency and the tags' link SNR determine the average throughput. We shall consider the case with identical link SNR for all tags in order to illustrate the deficiencies of EPC Gen2 and the potential of HARQ algorithms.

---

[6] N_0 is the power spectral density of thermal noise.

[7] Define $\mathcal{Q}(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-a^2/2} \mathrm{d}a$.

As discussed previously, EPC Gen2 uses FSA to first identify tags by their RN16 and a stop and wait ARQ mechanism to reliably recover its EPC payload. FSA consists of a frame divided into a fixed number of time slots, and after each slot, the reader either acks a collision free RN16 to trigger a EPC transfer or the reader signals the start of the next FSA slot. At the start of each frame, the reader broadcasts the number of slots in a frame N, and the K present tags choose at random a slot index from 1 to N. When the reader announces the slot index chosen by a tag, the tag responds with its RN16. At the completion of the current frame and the start of the following frame, the same group of K tags continue participate in new frame irregardless of previous success of failure.

Following basic results of random processes, the throughput can be written in terms of transmission length of the RN16, the EPC, and their corresponding probabilities of error. Define $R(t) = \sum_{i=1}^{U(t)} R_i$, where $U(t)$ is the number of successfully transmitted frames in the interval $(0, t]$, R_i is the number of successfully read tags in the i-th frame, and L_i is the time duration of the i-th frame. Lastly, the statistics of each frame is independent and identically distributed (iid), so R_i and L_i are iid. The average throughput ($\mathcal{T}$) can be written using a renewal reward process (Thm 3.6.1 [99])

$$\mathcal{T} = \lim_{t\to\infty} \frac{R(t)}{t} = \frac{E[R_i]}{E[L_i]}. \tag{5.6}$$

Each frame has N slots and every tag chooses to transmits its RN16 in one randomly chosen slot. The expected number of tags read in a single ALOHA frame is [100]

$$E[R_i] = (1 - \mathcal{E}_{\mathrm{RN}})K(1 - 1/N)^{K-1}, \tag{5.7}$$

where $1 - \mathcal{E}_{\mathrm{RN}} = (1 - p_e)^{16}$ is the probability of decoding the RN16 correctly.

The duration of each frame is the summation of the time spent:

$$L_i = N\tau_{\mathrm{RN16}} + \sum_{j=1}^{R_i} V_j \tau_{\mathrm{EPC}} + (B_i + C_i)\tau_{\mathrm{X}}, \tag{5.8}$$

where V_j is the number of transmissions to decode the j-th tag correctly, B_i is the number of incorrectly decoded RN16s, and C_i is the number of RN16s that experience a collision. From linearity of expectations $E[L_i]$ is calculated by finding the expectation of each term of the sum.

By Wald's equation [99], the expected number of EPC transmissions that occur in the entire frame is

$$E\left[\sum_{j=1}^{R_i} V_j\right] = E[R_i]E[V_j] \tag{5.9}$$

The distribution of V_j is geometric with parameter $1-\mathcal{E}_{\text{EPC}}$, where $1-\mathcal{E}_{\text{EPC}} = (1-p_e)^{528}$ is the probability of decoding the EPC message correctly. The expected number of incorrect RN16s is

$$E[B_i] = \mathcal{E}_{\text{RN}} K(1-1/N)^{K-1}. \tag{5.10}$$

Since the RN16 has no error detection protection, a collision is likely to generate a RN16 detection but with an incorrectly value. The probability that a slot is empty or has only a single message is $(1-1/N)^K + \frac{K}{N}(1-1/N)^{K-1}$. The marginals for each slot is identical, so the number of collision slots can be written as a sum of Bernoulli random variables with probability $1-(1-1/N)^K - \frac{K}{N}(1-1/N)^{K-1}$, i.e.,

$$E[C_i] = N\left(1-(1-1/N)^K - \frac{K}{N}(1-1/N)^{K-1}\right) \tag{5.11}$$

Finally, the expected time for a frame to complete is

$$E[L_i] = N\tau_{\text{RN}} + \frac{E[R_i]\tau_{\text{EPC}}}{1-\mathcal{E}_{\text{EPC}}} + (E[B_i]+E[C_i])\tau_{\text{X}} \tag{5.12}$$

To further simplify $E[R_i]$ and $E[L_i]$, consider the approximation $(1-x)^y \approx e^{-xy}$ for small x [28]. Denote $G = \frac{K}{N}$,

$$E[R_i] \approx (1-\mathcal{E}_{\text{RN}})NGe^{-G} \tag{5.13}$$

and

$$E[C_i] \approx N(1-e^{-G}-Ge^{-G}). \tag{5.14}$$

The N and K terms can be canceled and replaced with generalized parameter G, which describes the offered load per slot. Bringing these results together, the average throughput is

$$\mathcal{T} = \frac{(1-\mathcal{E}_{\text{RN}})NGe^{-G}}{N\tau_{\text{RN}} + \frac{E[R_i]\tau_{\text{EPC}}}{1-\mathcal{E}_{\text{EPC}}} + (E[B_i]+E[C_i])\tau_{\text{X}}} \tag{5.15}$$

In most ALOHA style random access systems maximal throughput occurs at $G = 1$, however, the reader may mistake collisions as valid RN16s and transmit spurious ACK

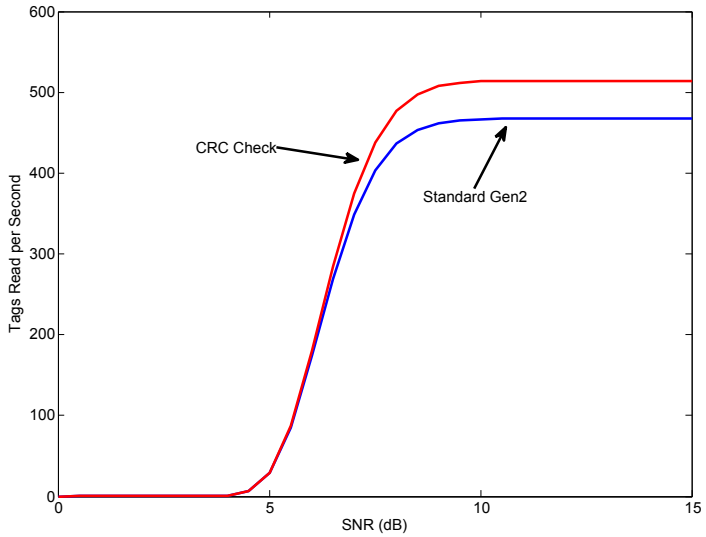


Figure 5.4: Comparison of the baseline Gen2 protocol and the CRC enhancement that can detect RN16 collisions. The tag's symbol frequency is operating at 465KHz.

messages. So G should be slightly smaller than 1 in order to reduce the number of collisions. Numerically, the impact of optimizing over G is rather small since the reader can recover from an incorrect ACK rather quickly, so for simplicity, we set $G = 1$ throughput the remainder of this work.

One simple enhancement adds a cyclic redundancy check (CRC) to the RN16, which allows the receiver to easily detect collisions and incorrectly decoded RN16s. A practical choice is the 5 bit CRC already utilized by the reader's downlink command [27], which provides a probability of undetected error of at least 2^{-5} [98], and thereby nearly eliminating incorrect RN16 ACKs by the reader. The timing and error probability values are adjusted to account for the additional symbols, and the contributions of B_i and C_i are removed. So the expected time for a frame to complete is now

$$N\tau_{\mathrm{RN}} + NGe^{-G}(1-\mathcal{E}_{\mathrm{RN}})(1-\mathcal{E}_{\mathrm{EPC}})^{-1}\tau_{\mathrm{EPC}}, \tag{5.16}$$

which is shorter as compared to the standard frame duration (5.12). Applying the same renewal process argument as before, the average throughput is rewritten as

$$\mathcal{T}_{\mathrm{CRC}} \approx \frac{(1-\mathcal{E}_{\mathrm{RN}})Ge^{-G}}{\tau_{\mathrm{RN}} + Ge^{-G}(1-\mathcal{E}_{\mathrm{RN}})(1-\mathcal{E}_{\mathrm{EPC}})^{-1}\tau_{\mathrm{EPC}}}. \tag{5.17}$$

Fig. 5.4 illustrates the tag read rate throughput for baseline Gen2 case (5.15) and the CRC enhanced case where the collisions are perfectly detected (5.17). Despite the added overhead of the CRC, throughput is slightly improved. The remainder of the this work shall use the CRC enhanced throughput as the baseline to measure against.

5.3 A RFID Based HARQ Protocol

HARQ protocols can be described as a subclass of *rateless codes*, which are defined as codes with no fixed rate. The code construction allows the rate of the code to automatically adapt to the channel capacity. Defined formally, a rateless code has the property of that the prefix of any codeword remains as a valid codeword but at a higher rate [101]. Under this architecture, the sender can continuously generate codeword symbols until the receiver is able to decode the message error free, and the receiver can then instruct the sender to stop transmitting. This section describes how a HARQ protocol can operate with respect to RFID and develops a theoretical limit on the performance of *any* HARQ or rateless coding scheme within the context of EPC Gen2. Specific HARQ examples are discussed in Sec. 5.4 and 5.5. Note that we assume the length of the EPC is fixed and known to reader and as a running example we use 528 bits. This issue is addressed in Sec. 5.6.

5.3.1 Protocol Description

As described in the previous section, the tag performs retransmissions based on feedback (in the form of ACKs) from the reader. For the feedback to occur, the reader and tag do not utilize the channel while decoding the received signal at each retransmission (see T_1 and T_2 in Fig. 5.2). These wasted timing gaps can be removed by leveraging properties unique to non-active RFID systems.

RFID is unique with regards to how the tag uses the reader's CW to backscatter its signal. Since the tag does not actively generate its own RF signal, it is feasible to sense the the presence of the reader's CW during backscatter operation. So instead of transmitting a fixed length response and then waiting for the reader response, it has been previously

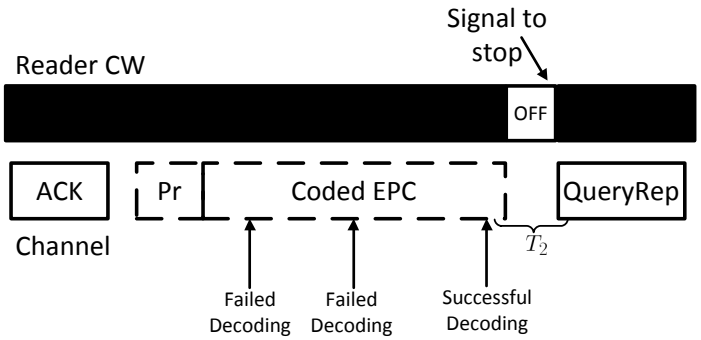


Figure 5.5: HARQ protocol showing 3 attempts by the reader to decode the message. The reader CW label indicates that the CW is on when the bar black and off when white. The label Pr represents the tag's preamble.

proposed that the tag continuously backscatters symbols until the reader blanks[8] the CW for a brief interval [88]. Effectively, feedback from the reader to tag can operate in full duplex mode rather than half duplex. While the blanking commands can be misinterpreted at the tag due to changes in channel fading, the PIE coding of the reader's commands suffers from the same issue. This type of signaling was also demonstrated experimentally by [88]. In contrast to [88], our contributions focus on analytical results and evaluating the suitability of using HARQ algorithms within Gen2.

The protocol algorithm is defined as the following. After the tag undergoes successful contention resolution via FSA, the tag receives a ACK message from the reader that matches its corresponding RN16. The tag shall

1. start backscattering its preamble and then its coded message. The symbols are continuously generated.

2. On its own discretion, the reader attempts to decode message as more symbols are collected at the receiver.

3. Once the CRC succeeds, the reader sends the 'blanking' command to the tag.

4. The tag stops backscattering and the reader initiates another QueryRep round.

[8]The blanking period can be on the order of $12.5 - 25\mu$s. We assume it can be incorporated within the existing gap between the tag's message and the reader's response.

Fig. 5.5 illustrates this protocol starting after the initial ACK. Note that it is up to the reader to chose when to start decoding the message and how often to retry the decoding process.

5.3.2 Information Theoretic Limit on the Protocol

The performance limit of any HARQ or rateless coding scheme can be established through the use of a perfect capacity achieving code for the EPC component of the tag's transmission. It is well understood that the random coding strategy used by the capacity limit derivations is also a rateless code [101]. So the channel coding capacity limit serves as an optimistic bound on the proposed protocol's performance.

The number of symbols it takes to decode the EPC is the message size (our running example is 528 bits) divided by the channel capacity $\mathcal{C}$ (bits/symbol). The amount of time required to exchange the ACK and EPC message is now

$$\tau_{\mathrm{HA}} = T_{\mathrm{ACK}} + T_1 + \frac{1}{W}\left(18 + \frac{528}{\mathcal{C}}\right) + T_2. \tag{5.18}$$

Note that the preamble consists of 18 symbols and remains unchanged from standard Gen2. With regards to (5.17), the term τ_{EPC} is replaced by τ_{HA} and the EPC probability of error term is removed. Thereby resulting in

$$\mathcal{T}_{\mathcal{C}} \approx \frac{(1-\mathcal{E}_{\mathrm{RN}})Ge^{-G}}{\tau_{\mathrm{RN}} + Ge^{-G}(1-\mathcal{E}_{\mathrm{RN}})\tau_{\mathrm{HA}}}. \tag{5.19}$$

The preamble sequences and the RN16 frames remains unchanged, because the message length is far too short to assume the existence of a code that can achieve capacity.[9]

We consider two practical channel capacity results: BPSK and AWGN. A closed form of BPSK capacity is not available, but it can be readily calculated through numerical integration techniques [46]. Following the system model described by (5.4) with $\gamma = 1$, the

[9]Theoretically, reliable channel codes are constructed over an infinitely large block size; this work defers to the classic Shannon capacity results.

capacity is[10] [46]

$$\mathcal{C}_{\text{BPSK}} = \max_{p(X)} I(X;Y) \tag{5.20}$$

$$= 1 - \int_{-\infty}^{\infty} H(X|Y=y)p(y)\mathrm{d}y \text{ bits/symbol.} \tag{5.21}$$

Where the input distribution is chosen to be $p(X=+1) = p(X=-1) = 0.5$. At high enough SNR, the constellation size limits the capacity to 1 bit per symbol.

The capacity of the AWGN channel follows that of the classic Shannon result, where the channel model has no constraint on the constellation size

$$\mathcal{C}_{\text{AWGN}} = \log_2(1 + \text{SNR}) \text{ bits/symbol.} \tag{5.22}$$

Where the input distribution is chosen to be Gaussian distributed, $\mathcal{CN}(0,1)$. The Gaussian codebook scheme has no limit in constellation size, so the capacity grows unbounded with SNR. These two bounds are discussed with respect to the different HARQ strategies in the following sections.

5.4 Hybrid ARQ: Chase Combining

5.4.1 System Description

CC is defined as the ML combining of repeated packet transmissions (FEC coded or uncoded) [96], which effectively make it rateless repetition code. For simplicity we only evaluate CC for an uncoded EPC. While CC can operate on either bits or symbols, we shall only consider CC across symbols since it has the better performance of the two.

After the tag receivers the reader's matching ACK command, the tag continuously backscatters copies of its EPC until the reader tells it to stop. An example of the tag sending three repetitions of its EPC is demonstrated in Fig. 5.6. Following the reader receiver model (5.4), each received packet can be written in vector form

$$\mathbf{y}_i = \mathbf{x}_i + \mathbf{w}_i, \tag{5.23}$$

[10] Capacity is defined in terms of information theoretic quantities. $H(X) = E_X[-\log_2 p(X)]$ is the entropy of a random variable and $I(X;Y) = H(X) - H(X|Y)$ is the mutual information [73]

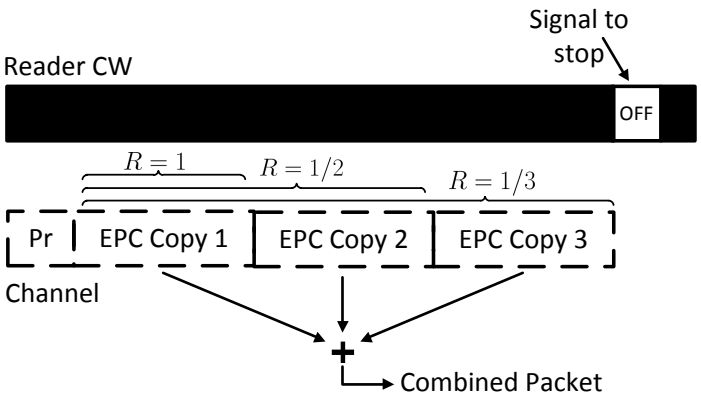


Figure 5.6: A HARQ CC example that requires three EPC frame transmissions for successful decoding. At each decoding attempt, the reader soft combines the previous transmissions.

where $i \in 1, 2, \ldots$ refers to the i-th transmission of the EPC and $\mathbf{y}_i, \mathbf{x}_i, \mathbf{w}_i \in \mathbb{C}^n$ and n is the number of symbols sent in each EPC copy. ML decoding follows standard BPSK detection of the signal

$$\mathbf{y} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{y}_i, \tag{5.24}$$

where $M - 1$ is the number of repetitions by the tag. The effective coding rate is $1/M$ and the probability of bit error at the M-th transmission is $\mathcal{Q}(\sqrt{2M\gamma \mathrm{SNR}})$.

A key practical observation is that CC can be implemented without changing the Gen2 standard. Instead of discarding each of the retransmissions of stop and wait ARQ, they can be combined in the same manner described above. Lastly, CC is implementable on the tag with no major change of the on tag memory requirements, because the message is simply repeated from the tag's memory.

5.4.2 Simulation Results

Our simulation methodology is to find the expected number of symbols required for correct decoding of the EPC at varying levels of SNR. Define $\mathcal{S}_{\mathrm{CC}}$ as expected number of symbols needed to correctly decode the EPC message at a fixed SNR. The expected amount of time to complete one EPC round is

$$\tau_{\mathrm{CC}} = T_{\mathrm{ACK}} + T_1 + \frac{1}{W}\left(18 + \mathcal{S}_{\mathrm{CC}}\right) + T_2. \tag{5.25}$$

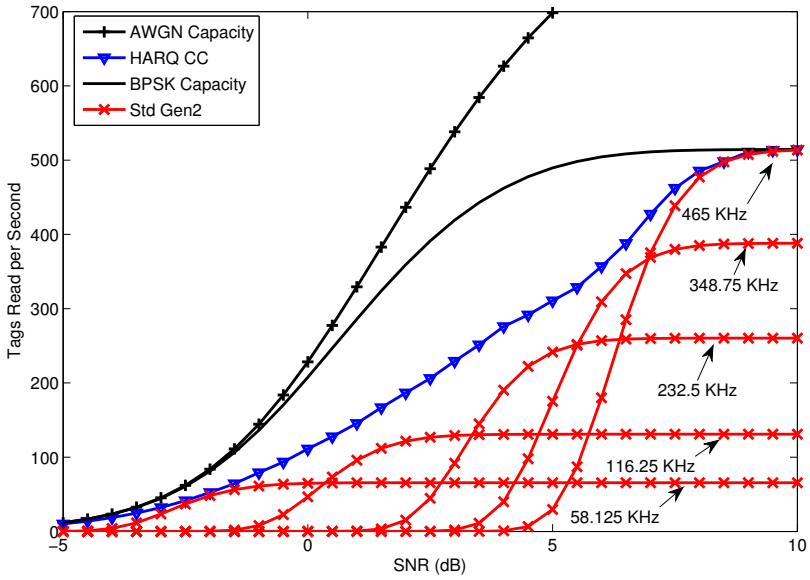


Figure 5.7: Tags read per second vs SNR for standard Gen2 at varying W values, HARQ CC operating at 465KHz, and the BPSK/AWGN capacity bounds. The SNR is taken with respect to a symbol with power P and period $1/W$ where $W = 465$KHz.

The expectation $\mathcal{S}_{\text{CC}}$ is obtained by Monte Carlo simulations over 1000 iterations. Used in conjunction with the analytical throughput model derived at (5.19), where τ_{CC} replaces τ_{HA}. The expected tag throughput is plotted in Fig. 5.7. Our derived tag read rate (approximately 500 tags per second) is comparable to claims by industry [37].[11]

The results demonstrate the main problem with the current Gen2 protocol. There exists an almost binary effect in terms of read rate throughput. If the SNR is above a certain threshold, the throughput quickly saturates. Increasing the SNR has no additional benefit. Similarly, if the SNR is below a threshold the throughput quickly drops to zero.

The CC scheme enables a graceful decline in throughput over a range of 15dB in SNR, despite no changes to the RN16 reliability. For moderate SNR, CC outperforms standard Gen2 by a large margin but a sizable gap between CC and BPSK capacity remains. At low SNR, CC begins to approach the BPSK capacity bound. The AWGN capacity bound

[11]Industry reports up to 1000 tags per second at a higher link frequency of 640 KHz and possibly with smaller EPC message sizes than our 528 bits.

suggests that higher order constellations might only be beneficial in medium or high SNR systems, but at low SNR, BPSK modulation should be sufficient.

5.5 Hybrid ARQ: Incremental Redundancy

5.5.1 System Description

In IR HARQ the sender first transmits a high rate code, and if the transmission is unsuccessful additional parity bits are sent until decoder success [95]. For implementation and practical reasons, the high rate code is constructed by *puncturing*[12] a low rate *mother code*. Puncturing removes a subset of the parity bits, and the receiver, knowing which bits are absent, has the decoder treat the punctured bits as erasures [98]. The mother code is the code that all the punctured versions are derived from, which allows the sender and receiver to use a single encoder and decoder.

The lowest rate of a punctured code is the base rate of the mother code. So to achieve lower code rates, a common approach is to use repetition coding after the lowest base rate has been reached. Fig. 5.8 demonstrates an example of using additional IR and repetition coding. After each incremental parity stage, the receiver combines any previous repetitions and attempts to decode.

The operation in Fig. 5.8 is written mathematically by the following. Suppose $\mathbf{a}_i, \mathbf{b}_i, \mathbf{c}_i, \mathbf{d}_i, \mathbf{e}_i$ are the received signals for the first to fifth IR stages at the i-th repetition. The receiver decodes by

$$\left.\begin{array}{c}(\mathbf{a}_1+\mathbf{a}_2)/2\\(\mathbf{b}_1+\mathbf{b}_2)/2\\\mathbf{c}_1\\\mathbf{d}_1\\\mathbf{e}_1\end{array}\right\} \rightarrow \text{Viterbi Decoder} \tag{5.26}$$

The IR scheme used in this work is derived from a set of well understood RCP convolution codes (RCPC codes) [1]. For convolutional codes, the additional complexity required by the encoder at the tag is manageable because it operates sequentially, has little state memory,

[12]Please see [1] to better understand how code puncturing is done. For brevity, we skip this description.

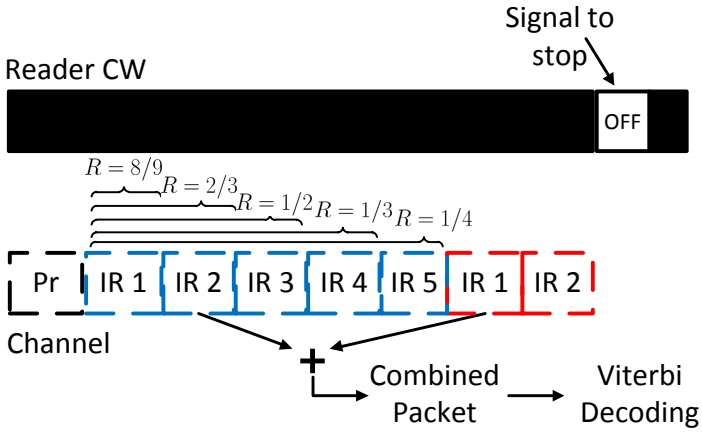


Figure 5.8: A HARQ IR example where multiple IR stages and CC stages are necessary. The blue and red outlines mark the copy of the packet that each IR stage belongs to. The reader attempts to decode at the end of each IR stage.

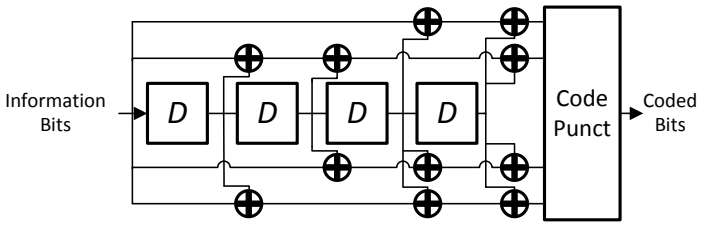


Figure 5.9: The encoder circuit used by the RCPC code simulated in this work. A stream of binary bits are fed into the encoder and then punctured. One input bit generates 4 output bits to the puncturing block. Each element labeled with D represents a delay memory element and $\oplus$ is the XOR operation.

and only requires bitwise operations. To minimize on tag memory usage, the coded bits can be calculated on the fly during each IR stage transmission. Fig. 5.9 depicts of the encoder circuit used in this work, which illustrates the simplicity of the decoder and its suitability for implementation in hardware. The encoder has a memory of four, so the corresponding decoder uses a 16 state trellis. The decoder is based on the Viterbi algorithm, so there exists well optimized implementations ready for use by the reader.

Punctured codes can have a wide range of offered rates, but for simplicity we choose from a set of five rates: $\frac{8}{9}, \frac{2}{3}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$. Table 5.1 lists the code generator polynomials and the different rates, d_{free}, and puncturing patterns.

While there exists better performing RCP variants (near capacity achieving) for Turbo

Code Polynomial [23, 35, 27, 33]			
Rate & d_{free}	Punct. Pattern	Rate & d_{free}	Punct. Pattern
Rate $= \frac{8}{9}$ $d_{\text{free}} = 2$	1111 0111 1000 1000 0000 0000 0000 0000	Rate $= \frac{2}{3}$ $d_{\text{free}} = 4$	1111 1111 1010 1010 0000 0000 0000 0000
Rate $= \frac{1}{2}$ $d_{\text{free}} = 7$	1111 1111 1111 1111 0000 0000 0000 0000	Rate $= \frac{1}{3}$ $d_{\text{free}} = 11$	1111 1111 1111 1111 1111 1111 0000 0000
Rate $= \frac{1}{4}$ $d_{\text{free}} = 15$	1111 1111 1111 1111 1111 1111 1111 1111		

Table 5.1: RCPC Code polynomial with memory 4, puncturing patterns with period 8, and corresponding d_{free} [1].

codes [102], the block lengths required to achieve high performance is quite large. Additionally, the encoders and decoders are significantly more complicated as compared to RCPC codes. We only consider RCPC codes, because RCPC performance is already quite close to the capacity bound.

5.5.2 Simulation Results

We perform the same methodology used by the HARQ CC section to calculate the average throughput for HARQ IR. The RCPC code encoder and decoder are simulated within MATLAB using the trellis encoder and Viterbi decoder provided by the Communication Toolbox. The Viterbi algorithm uses the log likelihood ratio (LLR) of each received symbol as the branch metric. Punctured symbols are initialized with a LLR of 0, and as additional redundancy is received, the LLR values are updated.

Our simulation results, Fig. 5.10, show that RCPC codes are sufficient to achieve a large fraction throughput capacity in terms of tags read per second. These results prove that a simple HARQ IR scheme can greatly improve the reliability of EPC Gen2 with only 1-bit

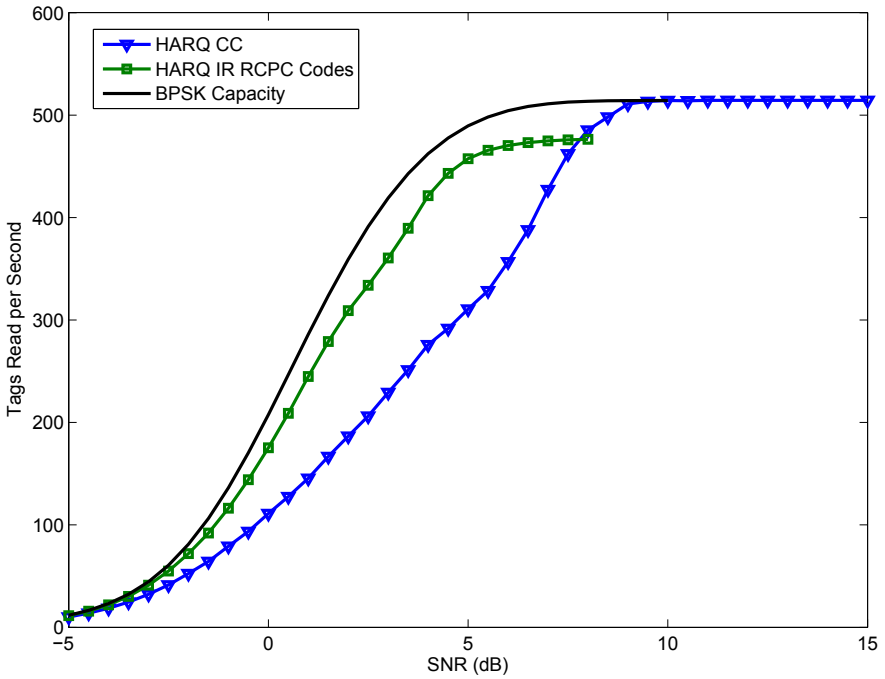


Figure 5.10: Tags read per second vs SNR for HARQ IR with RCPC codes. The SNR is taken with respect to a symbol with power P and period $1/W$ where $W = 465$KHz.

feedback from the reader (blanking signal). The encoders and decoders both have low cost in terms of energy, hardware, and engineering effort to implement in future tags and readers.

Although, note that the throughput is slightly lower than CC case and the BPSK bound, because the highest rate code used is a rate 8/9. To overcome this loss, it is possible to either slightly increase the tag's symbol frequency (W) or use a systematic RCPC code. The code used in this work is non-systematic. Using less puncturing stages, such as four instead of 5, results in less granularity in performance. Since five stages appear to achieve capacity, we do not believe any additional puncturing stages will be of any use.

5.6 Variable Length EPC

5.6.1 System Description

The previous protocol designs assume that message length is already known to the receiver. Without explicitly knowing the message length, it is not possible for the receiver to decode

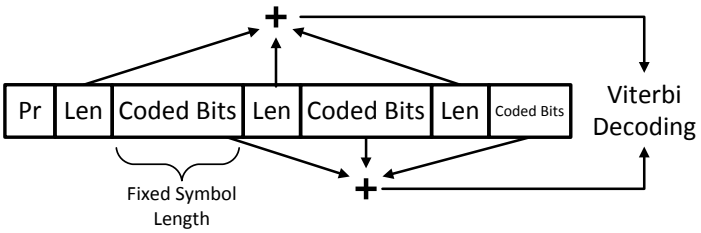


Figure 5.11: The variable length HARQ protocol inserts the length field periodically into the stream of the coded EPC symbols. To decode the EPC, the receiver uses the estimate of the length field. Each retransmission of length field provides the reader with a better estimate.

the message. For packet based wireless systems, a possible application for RFID, it is not always practical to have fixed length packets. In this section, we describe a method that insures reliable transmission of the length field while maintaining the rateless property of HARQ.

The length field and the EPC are coded using separate encoders and are both transmitted continuously. Every B transmitted symbols, the length field is repeated. The B symbols between each repetition are composed of the symbols generated by the EPC HARQ; either CC or IR HARQ can be used. The receiver first performs CC of the repeated length fields to remove any bit errors. Fixing the repetition to occur periodically on the number of symbols allows the receiver to know what symbols to correctly combine. After estimating the packet's length field, the decoder attempts HARQ decoding of the EPC. Until the CRC succeeds, the receiver continuously updates its estimate of the length field and attempts to decode the EPC. Fig. 5.11 illustrates the protocol operation.

5.6.2 Variable Length Simulation Results

Simulations of the RCPC codes follow the same methodology described in the previous section. The key difference is the length field is no longer protected by the FEC. Reliability is achieved through repetition coding/chase combining. While there is some inefficiency with repeating the length field, the additional cost is low because field only has 5 bits; the entire EPC message can be up to 528 bits.

To simplify our simulations, we take advantage of our assumption that the message

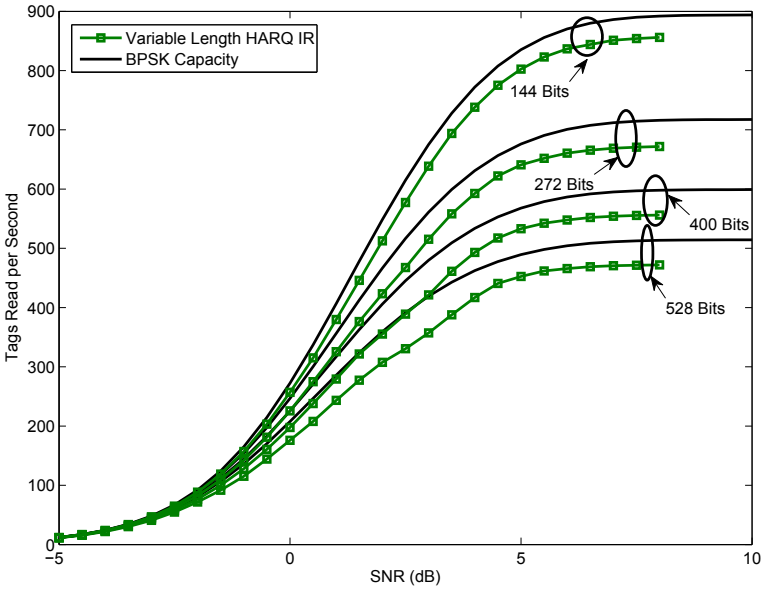


Figure 5.12: The throughput is plotted for multiple EPC sizes ranging from 528 bits to 144 bits. The SNR is taken with respect to a symbol with power P and period $1/W$ where $W = 465$KHz. The repetition period for the length field is $B = 250$ symbols.

CRC check is perfect. The simulation first repeats transmission of the length field until it is decoded correctly. It is assumed if the length field is not decoded correctly, the entire EPC message will always fail the CRC. Once the length field is correctly decoded, all of the symbols up to that point are collected and the decoding algorithm for HARQ IR begins. Most practical receivers will probably not try decoding on every length field repetition, since the receiver could wait until it has a reasonable probability of correct detection.

The results, Fig. 5.12, show that this scheme has little to no performance lost when compared to the case where the length is already known to the receiver. The system still operates within less than 3dB from the capacity bound. So the absence of length information does not inhibit HARQ in any practical fashion, as it can be recovered before the main message is decodable.

5.7 Conclusion

The results presented in this chapter suggest that HARQ or rateless protocols can greatly improve the reliability of RFID tag reading for a wide range of SNR. Since RFID can experience channel fading effects, this is a critical improvement. For WSN systems, this improves spectral efficiency, and for logistic systems this enables better object identification.

To implement these HARQ improvements, changes on both the reader and tag are necessary. The additional requirements to the reader are relatively minor. For CC, only some additional signal processing for combining the repeated packets will be required. For IR, the reader will need to use the Viterbi algorithm to decode the RCP codes. The tag will have to support the encoders for CC and IR, and have the circuitry to detect the reader's 'blanking' messages that tell the tag to stop backscattering.

Future work should investigate how to improve the throughput in the high SNR regime via high order constellations. Possible methods for additional throughput enhancements may include BICM [103], strider codes [104], or spinal codes [105].

Chapter 6

INCREMENTAL REDUNDANCY FOR FRAME SLOTTED ALOHA

6.1 Introduction

6.1.1 Background & Motivation

Since RFID tags cannot coordinate with other tags, a random access protocol is used to arbitrate channel access between tags. The random access scheme used by Gen2 is reservation frame slotted ALOHA (FSA), a variation of slotted ALOHA (SA) [28]. In FSA each round of random access is divided up into a frame containing N slots, where the frame length is set by the reader at the start of each round. In one frame instance a tag chooses a single slot to transmit its reservation request, which consists of a random 16 bit number (RN16). If the RN16 is decoded error free by the reader, the reader ACKs the RN16 on the downlink and the corresponding tag transmits its EPC payload contention free. If the RN16 cannot be decoded error free (e.g. a collision occurs), then that particular tag must wait till the next frame to attempt a reservation again. The efficiency of the random access protocol is important for RFID and backscatter systems for the following reasons.

- *High user density*: A typical scenario in RFID tagging applications involves a reader attempting to read tagged objects as quickly as possible. Looking at future applications of RFID, backscatter communication is being considered as a physical/link layer for the Internet of things, which may involve many sensors attempting to communicate with a central data fusion node [20–22].

- *Narrowband limited*: Due to their simplistic modulator design [7] and limited energy availability, backscatter based systems are primarily limited to narrowband operation around the reader's carrier frequency.

- *Short message lengths*: In the Gen2 standard, the tag's electronic product code length

varies between 96 and 512 bits [27]. Also for wireless sensor networks, each sensor may only have handful of bits of sensory data to transmit, and the length of connected sessions may be short. So the gains of using reservations instead of standard ALOHA are limited, as compared to wide area cellular networks with persistent data sessions.

- *No channel sensing*: For most short range wireless systems, carrier sense multiple access (CSMA) is a simple and effective method to avoid collisions [28], however, RFID tags only have a crude receiver [15] that can sense only tags in close proximity.

Next, we discuss the current performance limit of FSA and the Gen2 protocol.

Consider a scenario with K tags and a frame size of N slots. With the average system load defined as $G = K/N$, the optimal average throughput (in terms of RN16s) for FSA converges to $\mathcal{T}_{\text{FSA}} = Ge^{-G}$ as $K \to \infty$ as shown in Chapter 5. If K is known then the throughput optimal choice in frame size is $K = N$, which gives $\mathcal{T}_{\text{FSA}} = e^{-1}$ [106]. In terms of performance, FSA and SA are nearly identical in advantages and drawbacks, although terminating a frame slightly improves the throughput of FSA [107,108]. The Q algorithm, as mentioned in the Gen2 standard, is a variation on FSA that enables the reader to terminate and restart a frame with different lengths. This algorithm allows the reader to quickly adapt the frame size if the tag population size K is not known to the reader [109]. Similarly, work on Bayesian approaches to FSA attempts to find a good estimate of the tag population size based on previous collision slots, empty slots, and success slots [110, 111]. Effectively, the Q algorithm and Bayesian methods are capable of achieving efficiencies close to e^{-1}, the maximum efficiency for fixed frame sizes.

To go beyond the e^{-1} limit imposed by the FSA protocol, it is necessary to investigate more sophisticated schemes. Intuition on how to approach the problem can be drawn from the two main draw backs of FSA:

- slot collisions are destructive and irrecoverable and empty slots are wasted time;

- transmissions and retransmissions are independent of each other.

A successful slot only occurs if one user accesses the medium during that slot, any more or any less results wasted resources. Furthermore, the knowledge of past failures do not change the outcome of future transmission, because the next transmission event occurs independent of past events. FSA has poor performance, in terms of throughput and variation in load, because the successful transmission event is an unlikely corner case in the random access process. In general the probability of success and failure for a FSA is

$$p(\text{Success}) = \frac{K}{N}(1 - 1/N)^{K-1} \tag{6.1}$$

$$p(\text{Empty or Collision}) = 1 - \frac{K}{N}(1 - 1/N)^{K-1}. \tag{6.2}$$

Clearly it is not possible to drive p(Empty or Collision) to zero for any choice of N.

Two recent advances make progress in solving the ALOHA collision problem: coded ALOHA and sparse-recovery based multi-user detection. Coded ALOHA is a modification of FSA that allows interference cancellation between slots and retransmissions drawn from a pseudo-random Bernoulli process [112]. The scheme has also been extended for frameless case [113]. While the efficiency is improved over FSA, the protocol's efficiency is sensitive to the offered load G.

The area of multi-user detection (MUD) has received much attention in the context of spread spectrum or code division multiple access (CDMA) networks, although, the primary goal of the research was to reduce the detection complexity in terms of the number of users [114]. In classic CDMA scenario the receiver knows the users' identities and their corresponding spreading sequences, so the receiver's main challenge is mitigating the inter-user interference. While MUD does improve the throughput in random access systems [115, 116], it requires knowledge of the active users' code sequences. So called blind CDMA receivers still require knowledge of the desired user's spreading sequence [117]. In contrast to CDMA, the RFID reader's receiver has zero knowledge of what tags participate in a collision.

Recent work on CDMA examines the detection problem from a sparse signal point of view [118, 119], which draws on the research activity from the compressive sensing community. The receiver tries to detect the presence of a few spreading sequences out of a set contain many possible spreading sequences, and it has been recognized as a possible approach to the

random access problem [120]. In practical scenarios this approach has been experimentally studied at the system level by [121] for asynchronous communication links and by [88] in the context of RFID. Although [88] focuses on the overall system performance and it does not specifically isolate the performance of the random access portion of the protocol.

6.1.2 Contribution & Organization

The system of focus in this chapter is a reservation random access protocol called *incremental redundancy for frame slotted ALOHA* (IR-FSA), where error correction codes handle collision resolutions as opposed to a retransmission in a future time slot. Similar protocols have been discussed by [115, 122–124], but depend on either perfect channel knowledge at the receiver or the ability to estimate the channel state prior to decoding the active users. Another similar work [125] uses random codes as an error correction code and orthogonal matching pursuit to decode, which have inefficient encoders and the decoder has exponential complexity. Our focus is a to analyze the performance of IR-FSA under the condition of zero channel state knowledge at the receiver except for time synchrony between transmitters, and to investigate the performance of a Reed-Muller error correction code as a practical code construction for IR-FSA.

Sec. 6.2 first describes the IR-FSA protocol and its corresponding signal model. To quantify the theoretical performance of IR-FSA, Sec. 6.3 uses results on sparsity recovery [126] to develop a throughput and capacity model for IR-FSA. Sec. 6.4.1 presents the random access channel from a compressive sensing point of view and argues that compressive sensing algorithms offer the signal processing toolbox to implement the MUD decoder required by IR-FSA. In order to address the high computational complexity of most compressive algorithms, Sec. 6.4.2 proposes the use of rate compatible Reed-Muller codes as an error correction code for IR-FSA. Lastly Sec. 6.5 presents simulation results and compares the performance of the Reed-Muller code against recent work on RFID tag collision resolution [127–129].

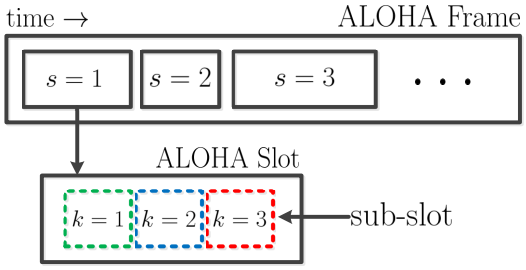


Figure 6.1: IR-FSA frame, slot, and sub-slot structure. Each sub-slot contains new parity symbols to assist the receiver in decoding colliding transmissions.

6.2 System Model & Protocol Description

Consider a framed transmission protocol with each frame is decomposed into N logical time slots (indexed from 1 to N) of variable length kT symbol periods (k is a variable positive integer). RFID tags access the channel following a random access procedure similar to FSA, as described within Gen2. At the start of each frame, each of K present tags choose a single slot index, at uniform random, to transmit their RN16. Tags cannot coordinate with each other prior to transmission, so like ALOHA, it is possible for multiple tags to collide within a single time slot. Additionally, the tags are time synchronous with the reader, such that the start of each slot is dictated by the reader. For example, the tags transmitting in slot s must wait for the announcement of the start of slot s by the reader. This synchronous behavior is necessary, because the length of each slot will vary based on a slot by slot basis.

A slot's length is determined by the capability of the reader to correctly resolve the colliding tags within the current slot, and the length is variable, because the number of colliding tags varies randomly on a slot by slot basis. Suppose each slot is divided by time into multiple sub-slots of length T symbols, and the tags within each sub-slot transmit a block of symbols as shown in Fig 6.1. The reader attempts to decode the tag's transmission after each sub-slot, and if decoding is successful, the current slot ends and the next slot starts. If decoding is not successful, the reader signals for an additional sub-slot symbol block, or it gives up to start the next slot. Any unsuccessfully decoded tag will wait until the next frame to reattempt its transmission.

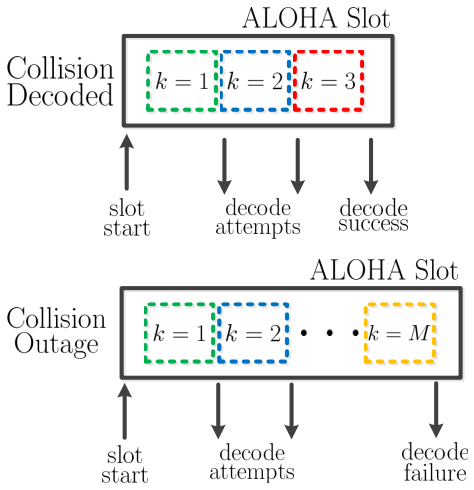


Figure 6.2: Example of IR-FSA decoding a collision with 3 sub-slot transmissions, and IR-FSA failing to decode the collision, thereby resulting in M sub-slot transmissions.

In order for the subsequent retransmissions in each additional sub-slot useful to the decoder, the symbols need to contain new information about the message. Simply repeating the same sequence in each sub-slot will reduce the thermal noise power, but will not reduce the interference between the colliding users. The framework of hybrid-ARQ incremental redundancy (HARQ-IR) protocols maps to the concept of sub-slot backoff in a natural fashion. Suppose each tag assigns their reservation message (RN16) to a codeword with rate R and length MT symbols, where M is the maximum number of sub-slot transmissions allowed. Assume R and MT are chosen such that there are a total of D possible messages or $\log D$ bits in a message. The codeword is broken into M separate chunks with T symbols each.

The transmission block of message i in sub-slot k is $\mathbf{x}_{i,k} \in \mathbb{C}^T$, where the collection of sub-slot transmissions $\mathbf{x}_{i,1}, \ldots, \mathbf{x}_{i,M}$ makes up the entire codeword. Additionally, the code has a rate compatible structure, such that $\mathbf{x}_{i,1}$ is a valid codeword of rate MR, $\mathbf{x}_{i,1}$ and $\mathbf{x}_{i,2}$ make up a valid codeword of rate $(M-1)R$ and so forth. The signal received at the reader in slot s and sub-slot k is $\mathbf{y}_k \in \mathbb{C}^T$, which is the superposition of the colliding users and

AWGN $\mathbf{w}_k \in \mathcal{CN}(0, \sigma^2\mathbf{I})$

$$\mathbf{y}_k = \sum_{i \in \mathcal{A}_s} h_i \mathbf{x}_{i,k} + \mathbf{w}_k. \tag{6.3}$$

The set $\mathcal{A}_s$ denotes the index set of active tags in slot s. An illustrative example of both a successful and failed collision resolution is shown in Fig. 6.2.

Since a wireless channel is subject multi-path effects, a $h_i \in \mathbb{C}$ models a random phase rotation and amplitude scaling. For simplicity, assume the channel follows a flat fading profile over the entire length of a slot. The SNR for tag transmitting message i is defined at the symbol level $\rho_i = |h_i|^2/\sigma^2$. It is assumed the receiver has no knowledge of $\mathcal{A}_s$ and h_i, but the receiver has perfect symbol timing knowledge.

In general for most wireless systems, the assumptions of flat fading and symbol level synchrony in a multi-user system are not correct. Either mobility or local oscillator offset will generate a time-varying channel state, and furthermore, varying distances from the receiver will void the assumption of time synchrony between users. This type of system from a RFID point of view does not share the same difficulties as a general wireless system. With a single active reader, tags all share the same carrier frequency, because each tag modulates the reader's carrier wave via backscatter modulation [7]. The message sent over the random access channel is short, so the channel state is unlikely to change over a single slot. Time synchronous transmission is a valid assumption, because RFID can only operate in short range scenarios.

Lastly, the RFID receiver encounters a large constant offset due to the simultaneous transmission of its carrier frequency and reception of the tags' backscatter signal. A simple method to remove the offset without complex channel state estimation is to restrict a tag's backscatter waveform to have zero mean, so a simple match filter will remove the constant offset without any impact on detection performance. The Miller and FM0 encodings currently used by Gen2 [27] already utilize a simple zero-mean waveform, i.e. $+1$ and -1 for the first and second halves of the symbol duration.

6.3 Information Theoretic Performance Analysis

If the set of active tags $\mathcal{A}_s$ and channel states h_i are known to the receiver, then the channel model described previously is essentially the classic multiple access channel (MAC). For which there exists a host of MUD techniques that can achieve good performance [114, 130, 131] and established mathematical tools to analyze the achievable channel capacity [73]. Protocols similar to IR-FSA have been analytically studied by [115] and [124] under the assumption of channel state knowledge, although, a true random access system these states are unknown to the receiver, thereby greatly increasing the difficulty of quantifying the actual capacity of the random access channel [120]. Recent results relate the MAC to a sparsity recovery problem [126], and under similar constraints of the random access channel. Their results provide sufficient conditions for correct detection as the message size goes to infinity. A similar approach is used by [120] to analyze the performance of slotted ALOHA (without retransmissions), but here we focus on modeling the performance of IR-FSA.

Since the results of [126] are proven only for a real valued channel model, consider the equivalent of the IR-FSA channel model (6.3) in the real domain. Assume the symbols for codeword in each sub-slot $\mathbf{x}_{i,k}$ are generated by an iid $\mathcal{N}(0,1)$ distribution, and the codewords are known to receiver. At the first sub-slot, the received signal is

$$\mathbf{y}_1 = \sum_{i \in \mathcal{A}_s} h_i \mathbf{x}_{i,1} + \mathbf{w}_1, \tag{6.4}$$

where the receiver does not h_i and $\mathcal{A}_s$, but for now assume the receiver knows the number of active users $A = |\mathcal{A}_s|$. If there are D possible messages and A users transmitting in the first sub-slot, then the sufficient condition to reliably recover all users within the first sub-slot is [126]

$$\lim_{D \to \infty} \sup \frac{\log D}{T_D} < q(\mathbf{h}) \tag{6.5}$$

with the sub-slot length T_D scaling with D and $q(\mathbf{h})$ as a constant defined by

$$q(\mathbf{h}) = \min_{\mathcal{F} \subset \mathcal{A}_s} \frac{1}{2|\mathcal{F}|} \log \left(1 + \sum_{i \in \mathcal{F}} \rho_i \right). \tag{6.6}$$

The typical worst case scenario for MUD is equal receive power for all users, which effectively maximizes the interference caused between users. For $\rho_i = \rho_j \ \forall i, j \in \mathcal{A}_s$, it can be shown

that $q(\mathbf{h}) = \log(1 + A\rho)/(2A)$ with ρ as the shared SNR. Correspondingly, the sufficient condition for successfully detected all A users in one sub-slot is $T > 2A \log D / \log(1 + A\rho)$. Since $A \log D$ is the effective sum-rate of all the messages, the sufficient condition says the MAC without channel state knowledge has a capacity of $\log(1 + A\rho)/2$. Interestingly, this is the capacity of the symmetric Gaussian MAC with channel state knowledge [73]. The lack of channel state does not impact the capacity of the MAC in symmetric case, and for the general case [126] suggests it does not affect the capacity in a meaningful way.

If the sufficient condition is not met, then it is assumed the receiver detects that it cannot resolve the collision and requests the next sub-slot. Since the codewords at each IR stage are generated by a iid random process, they form a rate compatible code or rateless code [132]. Therefore the sufficient condition for correctly decoding by sub-slot k is

$$kT > 2A \log D / \log(1 + A\rho), \tag{6.7}$$

and if this cannot be met by the last sub-slot $k = M$ then no messages are decoded. Effectively, the receiver requires an order of $A \log D$ symbols in order to decode all of the users. For the remainder of this section, assume the result of (6.7) holds for the original complex valued signal model, i.e. the 2 factor in (6.7) is removed.

Choosing the sub-slot length to be on the same order of $\log D / \log(1 + \rho_{\min})$, such that the receiver can decode a single message at some minimum SNR threshold $\rho_{\min}$. Clearly if two sub-slots are used, then at least two messages can be reliably decoded, and so forth for general A as shown by

$$A \frac{\log D}{\log(1 + A\rho)} < A \frac{\log D}{\log(1 + \rho_{\min})}. \tag{6.8}$$

Using this result, the throughput of IR-FSA is reasonably approximated by the following collision-MUD model.

- A users accessing the channel requires A sub-slots to decode the collision.
- $A > M$ cannot be decoded and M sub-slots are wasted in the decoding attempt.
- $K = 0$ uses one sub-slot.

While in practice multiple messages may be decoded in a single slot, it is difficult to capture these scenarios when analyzing the throughput at the protocol level. Instead, this simplified model enables straight forward throughput analysis of the IR-FSA protocol and it has some theoretical justification.

To analyze the aggregate throughput of the IR-FSA protocol, consider the saturated case - every tag has a new message to send at the start of each frame regardless of previous attempts and successes. The collision probability between the users is dictated by the ratio of users to frame size $G = K/N$, which the reader dictates at the start of each frame. Under the saturated condition, the length of each frame and the number of tags decoded per frame are jointly independent between frames. Therefore the average throughput ($\mathcal{T}$) follows the renewal-reward equation [99]

$$E[\mathcal{T}] = \frac{NE[\mathcal{S}]}{NE[\mathcal{L}]} \tag{6.9}$$

where $\mathcal{S}$ is the number of successful messages per slot and $\mathcal{L}$ is the length of a slot. Therefore, $NE[\mathcal{S}]$ is the average number of successful message detections per frame and $NE[\mathcal{L}]$ is the average length of a frame. Unlike the original throughput analysis for MUD capable FSA, it is necessary to capture average frame length frame length due to time varying back-off of the sub-slots.

As long as the number of users colliding in a slot is no greater than M, then the decoding all the messages is successful. The number of users colliding in a single slot follows the binomial distribution, hence, the expected number of successful users in a single frame is

$$E[\mathcal{S}] = \sum_{i=1}^{M} i \binom{K}{i} \left(\frac{1}{N}\right)^{i} \left(1 - \frac{1}{N}\right)^{K-i}. \tag{6.10}$$

The length of each frame is defined by the summation of the number of sub-slots used by every slot in a frame. As discussed previously, only a single sub-slot is used if zero or one users access than channel, A sub-slots are used if $A \leq M$, and M sub-slots are used if $A > M$. The expectation is evaluated as

$$E[\mathcal{L}] = \left(1 - \frac{1}{N}\right)^{K} + \sum_{i=1}^{M} i \binom{K}{i} \left(\frac{1}{N}\right)^{i} \left(1 - \frac{1}{N}\right)^{K-i} + M \sum_{i=M+1}^{K} \binom{K}{i} \left(\frac{1}{N}\right)^{i} \left(1 - \frac{1}{N}\right)^{K-i}. \tag{6.11}$$

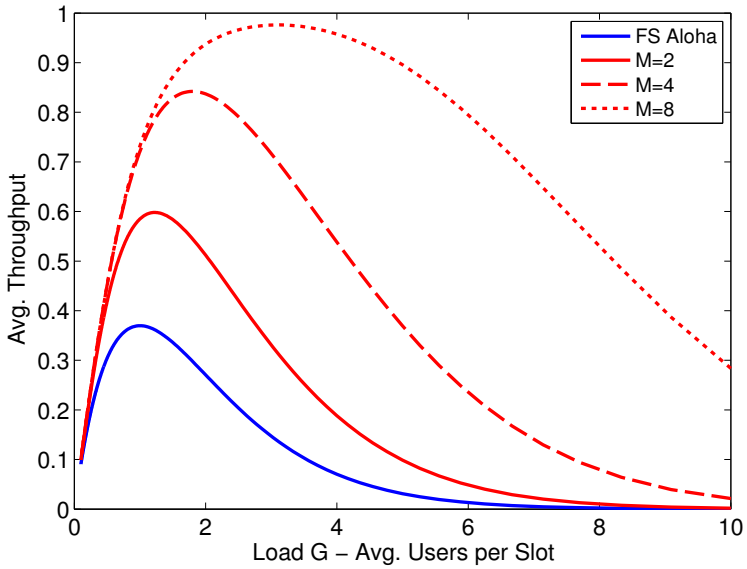


Figure 6.3: Baseline throughput of FSA and IR-FSA for $M = 2, 4, 8$. The curves are evaluated with a total tag population of $K = 100$.

Lastly, it is important to note that throughput never exceeds one message per channel use due to the channel capacity assumption of allowing at most one message detection per channel uses in a slot.

The average throughput (6.9) is plotted in Fig. 6.3 for different values of G and M. As illustrated by the figure, IR-FSA outperforms FSA in two key ways - increased throughput efficiency and improved tolerance to load variation. While increasing $M = 1 \rightarrow 2$ doubles the efficiency over FSA, the gains for larger M substantially drop off. For example, $M = 8$ achieves more than 95% useful utilization of the channel. FSA remains efficient only when $G = 1$ or near 1, however, IR-FSA offers high efficiency of a much wider range of G. In the case of $M = 8$, efficiencies of near 90% are met for loads approximately ranging over $1 \leq G \leq 6$. Just increasing the receiver's MUD capability to $M = 2$ offers substantially improved performance for moderate and high loads.

6.4 Multi-User Detection Model

This section first will describe the compressive sensing model for the random access channel, and how it supports the IR-FSA protocol described in Sec 6.2. Secondly, theoretical performance guarantees on correct recovery [119] are discussed and related to the information theoretic modeling in Sec.6.3. Lastly, the application of punctured second order Reed-Muller codes for computational efficient MUD is discussed.

6.4.1 Compressive Sensing Viewpoint

Compressive sensing (CS), in general, seeks to recover a high dimensional sparse vector by using as few as measurements as possible [133]. The random access channel was first recognized as an equivalent CS problem by [120], where the high dimensional sparse vector is the message space and the measurements are the received codeword symbols. As discussed by [120], a CS framework is able to address the short-comings of the information theoretic MAC analysis of the previous section. Primarily these short-comings are the assumption of arbitrarily long message lengths and little insight on how to design practical encoding/decoding algorithms. On the other hand the CS field provides a set of rich set of tools to tackle MUD in a random access setting and has been active area of research [118–120].

The received signal model for the random access channel is [120]

$$\mathbf{Y}_k = \mathbf{X}_k \mathbf{H} \mathbf{v} + \mathbf{w} \tag{6.12}$$

where $\mathbf{Y}_k \in \mathbb{C}^{kT}$ is a set of received symbols, $\mathbf{X}_k \in \mathbb{C}^{kT \times D}$ is the codebook matrix, $\mathbf{H} \in \mathbb{C}^{D \times D}$ is a diagonal matrix representing the channel gains, $\mathbf{v} \in \mathbb{C}^{D}$ is a vector composed of 1 and 0's, and lastly $\mathbf{w}$ is AWGN vector. The vector $\mathbf{v}$ indicates which messages are transmitted through the channel, e.g. $v_i = 1$ indicates that user with message i is transmitting. The columns of $\mathbf{X}_k$ are the codeword sequences assigned to each index in $\mathbf{v}$. The goal of the decoder is to estimate $\mathbf{v}$ given the observations $\mathbf{Y}_k$. Here the connection to CS is quite clear, since the dimension of $\mathbf{v}$ is much larger than $\mathbf{Y}_k$ and $\mathbf{v}$ has a small number of non-zero entries.

It is straight forward to extend (6.12) to support the IR-FSA protocol in a fashion similar

to [125]. After sub-slot k, the measurement vector can be written as the concatenation of the received symbols vectors from each sub-slot

$$\mathbf{Y}_k = [\mathbf{y}_1^\mathsf{T}, \mathbf{y}_2^\mathsf{T}, \ldots, \mathbf{y}_k^\mathsf{T}]^\mathsf{T} \tag{6.13}$$

and similarly for the codebook matrix

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{x}_{2,1} & \cdots & \mathbf{x}_{D,1} \\ \mathbf{x}_{1,2} & \mathbf{x}_{2,2} & \cdots & \mathbf{x}_{D,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{1,k} & \mathbf{x}_{2,k} & \cdots & \mathbf{x}_{D,k} \end{bmatrix} \tag{6.14}$$

Essentially after sub-slot, the receiver has additional samples to attempt to reconstruct $\mathbf{v}$.

Initial work on using CS in the random access channel [120] context proposes the use of an iid Gaussian distributed codebook $\mathbf{X}_k$ and a decoding algorithm called *sequential orthogonal matching pursuit* (SeqOMP) with power shaping active users. As addressed by [120], this approach has serious drawbacks for practical applications. A random codebook design cannot easily be stored on memory limited devices like RFID tags, and Gaussian distributed symbols cannot easily be generated with backscatter modulators. Although, the random coding strategy lends itself well to the IR-FSA protocol as discussed in Sec.6.3. Furthermore, power shaping is not a possible in a practical scenario, because users are not aware of their channel gains or other user's channel gains [120].

An alternative to the random codebook design is a deterministic construction of $\mathbf{X}_k$. The benefit of deterministic constructions are two fold: storage on devices like RFID tags is feasible, and it enables theoretical results that corroborate with the analysis of Sec.6.3. Seen as a way to reduce computation and to improve detection in CDMA multi-user detection, [119] proposes the use of Kerdock codes[1] as spreading sequences or codewords and orthogonal matching pursuit (OMP) as the decoder. The primary result of [119] proves that a randomly sub-sampled Kerdock code enables detection of all the messages with high probability if the number of samples is on the order of $A \log D$. The sub-sampling method

[1]The reader is directed to [119] for more information on Kerdock codes, as they are not the focus of this work.

used by [119] is equivalent to puncturing the codeword, so it is possible to construct $\mathbf{X}_k$ using Kerdock codes and maintain the rate compatible structure. This result agrees with the scaling discussed in Sec. 6.3, and does not require any coordination or channel state information, so viewing random access as a CS problem is a promising approach to tackle MUD in the random access channel.

The OMP detection algorithm is similar to successive interference cancellation (SIC) receivers used in CDMA systems [120]. As used by [119], OMP correlates the received signal with every possible message sequence, chooses the sequence with the strongest correlation, and lastly cancels that sequence out of the received signal. The primary drawback of OMP is the requirement to test every column in $\mathbf{X}_k$ against $\mathbf{Y_k}$, which results in a computational complexity of D. Each message has a length of $\log D$ bits, so the complexity of OMP grows exponentially in the message length. Scenarios with a small number of messages are not hindered by OMP's complexity, but the message used by RFID in EPC Gen2 behaves as a temporary tag ID number, which requires a large address space to insure no collisions between tags. To address OMP's computational complexity, it is necessary to use CS reconstruction methods with sub-linear complexity. This is the topic of the next discussion.

6.4.2 Second Order Reed-Muller for Efficient MUD

A (r, m) Reed-Muller (RM) code is an algebraic error correction code parametrized by its order (r) and its length (2^m symbols) [134]. Recently, $(2, m)$ RM codes were discovered to have excellent performance when used as deterministic sensing matrices in CS applications [135]. In comparison to random and some deterministic sensing matrix constructions, the RM code structure has a reconstruction complexity based on the code's length 2^m [135], unlike the complexity of D required by OMP. In the context of communication applications, MUD using the $(2, m)$ RM was first applied to neighbor discovery problem as a rapid on-off full duplex communication technique [136,137]. This section will first address the complexity of the $(2, m)$ RM encoder, then describe how to construct a rate compatible $(2, m)$ RM code, and lastly discuss how the $(2, m)$ RM code is used by the IR-FSA protocol.

Since RFID devices are computationally limited due to their ultra-lower power con-

straints, it is necessary to address the computational complexity of the $(2, m)$ RM code. A $(2, m)$ RM codeword $\phi_{\mathbf{b},\mathbf{c}}$ consists of quadrature phase shift keying (QPSK) symbols (e.g. $\{+1, -1, +j, -j\}$), which are generated by the function [136]

$$\phi_{\mathbf{b},\mathbf{c}}(\mathbf{a}) = \exp\left[j\pi\left(\frac{1}{2}\mathbf{a}^T\mathbf{P}(\mathbf{c})\mathbf{a} + \mathbf{b}^T\mathbf{a}\right)\right]. \tag{6.15}$$

The codeword symbols are indexed by the binary vector $\mathbf{a} \in \mathbb{Z}_2^m$, and the vectors $\mathbf{b} \in \mathbb{Z}_2^{n_1}, \mathbf{c} \in \mathbb{Z}_2^{n_2}$ comprise the message bits. The bit representation of the reservation message is split into two separate vectors $\mathbf{b}$ and $\mathbf{c}$ and zero-padded if needed, as described by [136].

The vector $\mathbf{c}$ is used to construct a $m \times m$ binary symmetric matrix $\mathbf{P}(\mathbf{c})$ from a basis set $\mathbf{B}(i)$ of n_2 $m \times m$ binary symmetric matrices [136]

$$\mathbf{P}(\mathbf{c}) = \sum_{i=1}^{n_2} c_i \mathbf{B}(i) \bmod 2. \tag{6.16}$$

Since the size of a reservation message is quite small, $\mathbf{B}(i)$ can be pre-calculated and stored in memory. In the examples discussed in the following section, values of $m = 6$ ($m = 7$) and $n_2 = 11$ ($n_2 = 13$) are used. So in the worst case calculating $\mathbf{P}(\mathbf{c})$ requires $n_2 m^2$ operations, although the calculation is only required once per message. The evaluation of $\mathbf{P}(\mathbf{c})$ lends itself well to embedded processors or standard digital hardware, because it is computed over modulo 2, hence, only simple XOR operations are required. The details of constructing $\mathbf{B}(i)$ are given by [136, 137] and are omitted for brevity.

The complexity of evaluating one symbol instance of a codeword is determined by the complexity calculating the exponent of (6.15). Since $\mathbf{P}(\mathbf{c})$ and $\mathbf{b}$ are fixed, only m^2 integer operations are required. Evaluating the exponential function in (6.15) can easily be accomplished by a look-up table mapping to the QPSK symbol set, because its exponent is always a multiple of $j\pi/2$. Lastly, generating the entire codeword requires 2^m iterations of (6.15), hence, the overall complexity of the encoder is $m^2 2^m$. Recall that the length of the code is $l = 2^m$, so the complexity of the encoder in terms of the code's length is $l \log^2 l$. While in terms of its length the encoder is super-linear in complexity, it is still less than polynomial and only requires binary and integer operations.

Rate compatible codes are constructed by successively puncturing a low rate code [98], however, the puncturing pattern should follow a structure that does not degrade the de-

coder's performance. The application of $(2, m)$ RM codes in network discovery [136] applies a structured random puncture pattern to achieve full duplex communication. In their scheme a user transmits its punctured codeword, and listens to other user's transmissions during its punctured symbol slots. The puncturing algorithm used by [136] was designed to create a fixed set of gaps in a user's transmission and not as an incremental redundancy scheme. So the existing puncturing algorithm requires slight modifications to support the rate compatible code structured required by IR-FSA.

Define a puncture pattern vector

$$\mathbf{r}_k = (r_{k,1}, r_{k,2}, \ldots, r_{k,2^m})^{\mathsf{T}} \tag{6.17}$$

indexed by the sub-slot transmission number k. The value of each element $r_{k,i}$ is 1 or 0, which indicate symbol i of $\phi_{\mathbf{b},\mathbf{c}}$ is to be received by sub-slot k. So the elements that change from 0 to 1 between the puncture patterns $\mathbf{r}_{k-1}$ and $\mathbf{r}_k$ indicate which symbols are sent in sub-slot k. By sub-slot M the entire codeword has been transmitted, so clearly $\mathbf{r}_M$ is a vector of all ones.

Similar to the method described by Algorithm 3 in [136], construct $\mathbf{r}_k$ by dividing the vector into 2^{m_0} consecutive segments of length 2^{m-m_0}. The first segment follows a certain pattern, to be described shortly, and the remaining segments are duplicates of the first segment. For $\mathbf{r}_k$, the first segment is generated as follows. Create an ordered set consisting of the integers $(1, 2^{m-m_0})$

$$\mathcal{I} = \{1, 2, \ldots, 2^{m-m_0}\}. \tag{6.18}$$

Generate a random permutation π and denote the permuted set as $\pi(\mathcal{I})$; this permutation remains fixed for the entire slot. Define the set $\mathcal{P}_n$ as the set of the first n elements of $\pi(\mathcal{I})$ and $\mathcal{P}_n^{\mathrm{C}}$ as the remaining $2^{m-m_0} - n$ elements of $\pi(\mathcal{I})$. Then the first segment in $\mathbf{r}_k$

$$r_{k,i} = 1 \;\; \forall i \in \mathcal{P}_n \quad \text{and} \quad r_{k,i} = 0 \;\; \forall i \in \mathcal{P}_n^{\mathrm{C}}. \tag{6.19}$$

The value of n is set by how many symbols need to be received by sub-slot k. As described previously, this pattern is duplicated for the remainder of $\mathbf{r}_k$. Since the permutation π is random it needs to shared with all of the users participating in a slot prior to transmission. One possible method is to link the choice of π to the current slot index.

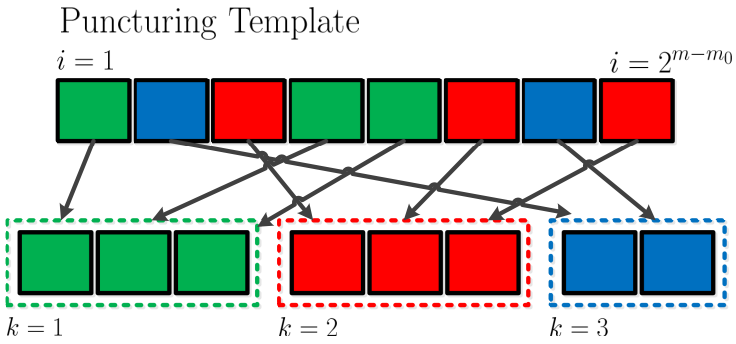


Figure 6.4: Example construction of a puncturing template for $2^{m-m_0} = 8$ with three stages of redundancy.

An example of generating a puncturing template is shown in Fig for $m - m_0 = 3$ and $M = 3$. In this example there are three incremental redundancy stages ($M = 3$) as indicated by the different coloring schemes. The corresponding permuted index set is

$$\pi(\mathcal{I}) = \{1, 4, 5, 3, 6, 8, 2, 7\} \tag{6.20}$$

and $n = 3, 6, 8$ for $k = 1, 2, 3$. So there are $3 \cdot 2^{m_0}$ symbols in the first two sub-slots and $2 \cdot 2^{m_0}$ symbols in the final sub-slot.

The decoding algorithm for a $(2, m)$ RM code with a puncturing pattern r_k remains unchanged from the description of Algorithm 4 in [136], so its details are omitted. At a high level, IR-FSA protocol uses the RM decoder in to perform SIC. The strongest signal is decoded, canceled out, and the process repeated until all the users are decoded. How the RM decoder is used by the IR-FSA protocol is detailed in Fig. 6.5, and it is in part based on the work by [136]. The IR-FSA decoding process consists of 5 key steps as highlighted by the following.

- At sub-slot k, input $\mathbf{Y}_k$ into the RM decoder to receive an estimate of an active user a and add it to the set of current active users $\hat{\mathcal{A}}$. Reconstruct the punctured codewords of the active users as the columns of the matrix $\mathbf{X}_{\hat{\mathcal{A}}}$.

- Estimate the channel gains of the active users by finding $\mathbf{h}$ such that $\|\mathbf{Y}_k - \mathbf{X}_{\hat{\mathcal{A}}}\mathbf{h}\|_2$ is minimized. Update $\mathbf{s} = \mathbf{Y}_k - \mathbf{X}_{\hat{\mathcal{A}}}\mathbf{h}$

procedure IR-FSA RM DECODER
  Input: $\mathbf{y}_1, \ldots, \mathbf{y}_M$
  Initialize $I_{\max}$, M, thresholds ϵ_1 and ϵ_2, and counters $i = 1$ and $k = 1$.
  while $k \leq M$ **do**
    Initialize $\mathbf{Y}_k$, $\mathbf{s} = \mathbf{Y}_k$ as the SIC update vector, and active user set $\hat{\mathcal{A}} = \emptyset$.
    while $i \leq I_{\max}$ **do**
      The RM decoder operates on $\mathbf{s}$, output message index a, and adds a to $\hat{\mathcal{A}}$.
      Reconstruct punctured codewords of active users as columns of $\mathbf{X}_{\hat{\mathcal{A}}}$.
      Est. the CSI of the active users $\mathbf{h} = \arg\min_{\mathbf{h}} \|\mathbf{Y}_k - \mathbf{X}_{\hat{\mathcal{A}}}\mathbf{h}\|_2$.
      Update the SIC update vector $\mathbf{s} = \mathbf{Y}_k - \mathbf{X}_{\hat{\mathcal{A}}}\mathbf{h}$.
      if $\|\mathbf{s}\|_2 < \epsilon_1$ **then**
        Prune from $a \in \hat{\mathcal{A}}$ if an active user's channel gain $h_a < \epsilon_2$.
        return $\hat{\mathcal{A}}$
      end if
    end while
    Could not successful decode, wait for next IR stage.
  end while
  return $\hat{\mathcal{A}} = \emptyset$.
end procedure

Figure 6.5: Pseudo-code for the IR-FSA RM Decoder.

- If $\|\mathbf{s}\|_2 < \epsilon_1$, then the received sequence $\mathbf{s}$ is a good estimate of true set of active users. To reduce the probability of incorrectly decoding a users prune any user whose channel coefficient $h_a < \epsilon_2$. Return the pruned set $\hat{\mathcal{A}}$ and exit the decoding process. Otherwise use the updated receive vector $\mathbf{s}$ as the input to the RM decoder, and repeat the decoding process until the energy threshold is met or the maximum number of iterations $I_{\max}$ is reached.

- If the number of iterations exceeds $I_{\max}$, then the current set of users cannot be

decoded by sub-slot k. Request the symbols for sub-slot $k+1$ and attempt the decoding process anew. If $k+1 > M$, then it is not possible to resolve the collision; return $\hat{\mathcal{A}} = \emptyset$.

6.5 System Simulation Study

To better evaluate the performance of the IR-FSA protocol, it is necessary to resort to numerical simulations as the performance modeling of IR-FSA discussed in Sec 6.3 only applies asymptotically, and the performance bounds derived using compressed techniques are quite loose. Numerical simulations of the punctured $(2, m)$ RM code are conducted and its performance is compared against the standard FSA model and recent advancements in the application of multi-packet reception in the context of RFID.

Two short length RM codes are considered, $m = 6$ and $m = 7$, which results in length 64 and 128 sequences. Short length codes are chosen in order to have a punctured code length that is similar to the length of an RN16 (16 symbols). For both cases the value of $m_0 = 2$ is chosen, and the 16 bit message is allocated to $\mathbf{b}$ and $\mathbf{c}$ by the following. For $m = 6$, $n_2 = 11$ so 11 bits is allocated to $\mathbf{c}$, and $n_1 = 6$ so the remaining five bits are allocated to $\mathbf{b}$. The sixth bit of $\mathbf{b}$ is zero padded. For $m = 7$, $n_2 = 13$ so 13 bits are allocated to $\mathbf{c}$, $n_1 = 7$ so the remaining three bits are allocated to $\mathbf{b}$. The four remaining bits of $\mathbf{b}$ are zero padded.

Two different puncturing schedules are chosen for the $m = 6$ and $m = 7$ codes. The first two sub-slot transmissions for $m = 6$ consist of 24 symbols each, and the third and last sub-slot transmission has 16 symbols. The $m = 7$ case consists has four sub-slot transmissions with 32 symbols each. These puncturing schedules were chosen because they exhibit good performance in simulations, whereas more aggressive puncturing schedules causes high error rates from the decoder. The puncturing patterns used in this work a effectively random and any unused message bits of a RM codeword are set to zero, so the poor performance under puncturing may be attributed to bad puncture patterns. This points to the need for research on developing puncturing patterns and code expurgation techniques [98] for the $(2, m)$ RM code.

The RM decoder's performance is quantified by the average number of messages decoded successfully and the average number of symbols transmitted when there is a fixed number

of active users A. Relating to the throughput model defined by (6.10) and (6.11), these two numerical results are the conditional expectations $E[\mathcal{S}|A]$ and $E[\mathcal{L}|A]$, respectively. The expectations are evaluated via Monte Carlo simulations for different values of A up to a maximum value of Z, where the decoder is unlikely to detect a large fraction of the collided messages. For $A > Z$, it is assumed that decoder will attempt all M sub-slots and fail to decode any message, hence, $E[\mathcal{S}|A] = 0$ and $E[\mathcal{L}|A] = 2^m$. Lastly for the case of no tag transmitting in a slot, assume $E[\mathcal{L}|A = 0]$ is equal to length of the first IR stage of the RM code.

Monte Carlo simulations are conducted over 1000 trials per value of A, where A ranges from 1 to 6 for the $m = 6$ case and 1 to 8 for the $m = 7$ case. For each trial, users have a channel gain distribution of $\mathcal{CN}(1, 0.5)$, i.e. the channel envelope follows a Rician distribution with the SNR of the line-of-sight component set to 15dB.

The throughput is calculated by slightly modifying the throughput equation developed in Sec. 6.3. Using iterative expectation and the renewal-reward theorem [99], the average throughput is calculated as

$$E[\mathcal{T}] = \frac{NE_A[E[\mathcal{S}|A]]}{NE_A[E[\mathcal{L}|A]]}. \tag{6.21}$$

Similarly, the numerator and denominator are written as

$$E[\mathcal{S}] = \sum_{i=1}^{Z} E[\mathcal{S}|A = i]\binom{K}{i}\left(\frac{1}{N}\right)^i\left(1 - \frac{1}{N}\right)^{K-i} \tag{6.22}$$

and

$$E[\mathcal{L}] = \sum_{i=0}^{Z} E[\mathcal{L}|A = i]\binom{K}{i}\left(\frac{1}{N}\right)^i\left(1 - \frac{1}{N}\right)^{L-i} + 2^m \sum_{i=Z+1}^{K}\binom{K}{i}\left(\frac{1}{N}\right)^i\left(1 - \frac{1}{N}\right)^{K-i}. \tag{6.23}$$

The simulation results are shown in Fig. 6.6 and the throughput is normalized against the length of the RN16 used by Gen2. In addition to comparing the simulation results with the theoretical throughput, they are compared against existing work that applies MUD techniques to the RN16 collision problem - the case of decoding one RN16 in the presence of a two tag collision [127], and the case of decoding two RN16s in the presence of a two tag collision [128, 129]. In both scenarios either one or two messages were decoded with high reliability, although [128] is able to decode more than two messages with much lower

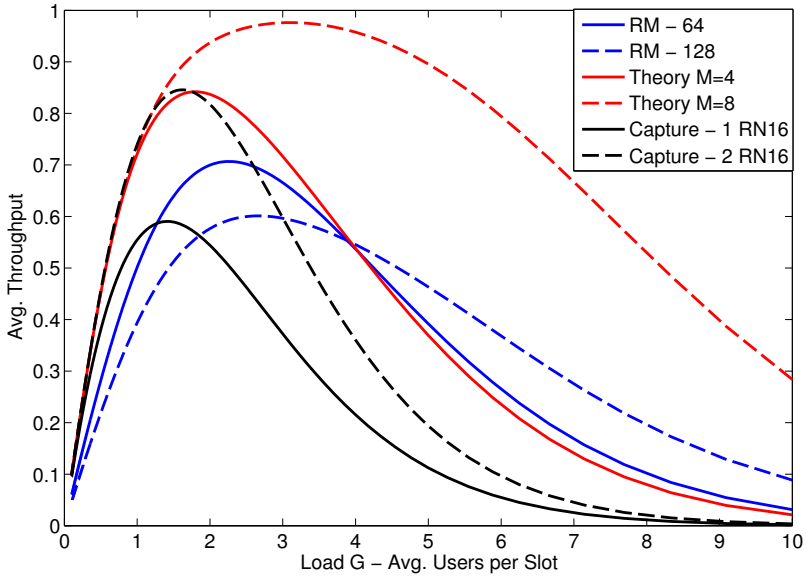


Figure 6.6: Average throughput of IR-FSA using punctured RM codes with length $2^6 = 64$ and $2^7 = 128$, theoretical model throughput, and throughput of prior work on RN16 MUD.

reliability. The $m = 6$ and $m = 7$ RM code can equally decode two messages with high reliability, and performance does not substantially degrade until there are more than three tags for $m = 6$ and four tags for $m = 7$. It is also important to note that these three works must perform channel state estimation of the tag's channel coefficients prior to decoding, whereas the RM codes only require state estimation for the SIC process.

The gap between the RM performance and the existing RN16 MUD work can be explained by the sub-optimal puncturing of the RM code. The minimum punctured code length for the $m = 6$ case is 24 symbols, as opposed to the RN16's 16 symbol length. The scenario for $m = 7$ is far more grim with a minimum length of 32 symbols. So immediately the RM code is put at a disadvantage due to the puncturing overhead. As mentioned earlier, under the current puncturing scheme it is not possible to further reduce the number of symbols sent in the first sub-slot without seriously degrading the code's performance. These issues emphasizes the need for both structured puncturing schemes and code expurgation. The latter is necessary, because the $m = 6$ RM code supports 27 message bits and the $m = 7$ supports 35 bits but a RN16 requires only 16 bits. The current decoder algorithm

does not take advantage of setting the remaining bits of the RM code to 0.

Differences between the theoretical throughput and the simulation generated throughput may also be explained in part by the sub-optimal performance of puncturing the RM code. The $m = 6$ RM code has performance most closely matching its theoretical equivalent of a $M = 4$ sub-slot backoff, and the $m = 7$ code falls far short of matching the $M = 8$ sub-slot backoff case. For both cases, the RM code exhibits two bottle necks: the smallest backoff of 24 (32) symbols as opposed to 16 symbols, and failure to decode 4 (8) users reliably. The issue with the former has already been addressed, and the latter may be solved by implementing heuristic improvements to the RM decoder as discussed in [138] and [139].

6.6 Conclusion

In RFID systems random access protocols plays a central role in identifying and reading RFID tags, however, such protocols are susceptible to collisions between tags. This work studies a random access protocol, IR-FSA, which uses error correction codes and incremental redundancy as a mechanism for collision resolution. The validity of such an approach is justified by existing results on compressive sensing. As an example of a code construction, punctured second order Reed-Muller codes are analyzed through simulations. While Reed-Muller codes do not out perform prior work on multi-user detection of tag collisions, its decoding algorithm and puncturing scheme still have room for improvement.

Chapter 7

CONCLUSION

7.1 Summary of Contributions

The proliferation of passive RFID into multitudes of commercial applications, and the emergence of WSNs based on re-purposed RFID systems clearly indicate the success and future potential of backscatter based communications. Backscatter modulation offers a new radio architecture that no longer requires active RF components at the tag. Instead of actively generating its own modulated RF signal, the tag modulates an externally generated RF source by switching its antenna impedance load between states, thereby modulating the RF signal scattering off of its antenna. As a consequence, a backscatter based communication link has new and different tradeoffs than those of an active transceiver architecture. Additionally, as commercial RFID systems and backscatter based WSNs expand their scope, the corresponding physical and link layer will require greater reliability and efficiency. This dissertation has three broad goals with respect to RFID: 1) analyze the co-design of backscatter modulation and error correction coding (Chapter 2), 2) characterize space-time coding and MIMO performance limits of the dyadic backscatter channel (Chapters 3 and 4), and 3) improving EPC Gen2 performance via hybrid-ARQ and random access enhancements (Chapters 5 and 6).

In Chapter 2, we model a RFID tag as an Thevenin equivalent circuit and derive the average power harvesting efficiency for an arbitrary backscatter modulator state. To quantify power and energy efficiency of backscatter, we define a new metric called backscatter power efficiency loss per bit. For a given 4-QAM constellation geometry, we find the modulator impedances that maximizes the average power harvesting efficiency. Using the corresponding optimal modulator design, we analyze two different coding schemes: equal error protection (EEP) and unequal error protection (UEP). EEP generates a symmetric constellation design similar to QPSK, however, UEP results in an asymmetric or skewed

constellation design. We find that UEP designs outperform EEP designs.

In Chapter 3, we analyze the performance of a MIMO RFID system using space-time codes. The (M, L, N) dyadic backscatter channel (DBC) captures the fading statistics of RFID system with M reader transmit antennas, L tag antennas, and N reader receive antennas. We derive a conditional pairwise error probability (PEP), which is amenable to Monte Carlo simulations, for a general codebook design. Additionally, the PEP for orthogonal space-time block codes (OSTBCs) and its corresponding diversity order is derived. We find that for OSTBCs the maximum diversity order is determined by L alone, and it does not depend on N. Although, increasing N provides a coding gain that diminishes as N increases. The analytical results are verified through simulations.

In Chapter 4, we derive the diversity multiplexing tradeoff (DMT) for the DBC. The tradeoff is characterized by a linear function between the maximum diversity order $d = L$ and the maximum multiplexing gain $r = \min(L, N)$. This reveals that a RFID MIMO system is primarily limited by reliability but not rate. The block length restrictions on the DMT curve reduce as N increases, and space-only codes are sufficient to achieve the channel capacity when $N = L$. The DMT suggests that a system with $N = L$ strikes a good balance between cost and performance.

In Chapter 5, we study the benefit of hybrid-ARQ (HARQ) techniques for the EPC Gen2 (or ISO 18000-6C) standard. We first characterize the baseline throughput of Gen2 with respect to receive SNR and symbol duration. As upper bounds to performance, we find the throughput of Gen2 if a tag uses either a Shannon capacity or BPSK capacity achieving code. Chase combining and incremental redundancy (IR) are the two types of HARQ under consideration in this chapter. The former provides considerable gains over the baseline result, however, a large gap exists between it and the BPSK capacity curve. The latter exhibits near capacity achieving performance for most ranges of SNR.

In Chapter 6, we propose a multiuser detection (MUD) based random access algorithm to improve the tag singulation phase of the EPC Gen2 protocol. The protocol, incremental redundancy for frame slotted ALOHA (IR-FSA), uses punctured error correction codes as a collision resolution method. The throughput of IR-FSA is characterized using existing results on sparse recovery. Increasing the MUD capability of the receiver results in im-

proved channel utilization and greater resilience to varying load conditions. We find that compressive sensing methods are a practical way to implement encoders and MUD based decoders for IR-FSA. As an example construction, we evaluate the average throughput of a punctured second order Reed-Muller code based IR-FSA. The performance is compared against existing results on multi-tag singulation.

7.2 Future Work

The contributions to the three core areas (coding/MIMO/MAC) studied within this dissertation reveal backscatter modulation as a communication channel with novel performance tradeoffs. As such, there exists the multiple new directions of research within the aforementioned areas.

7.2.1 Coded Modulation for Backscatter

Chapter 2 assumes both a linear power transfer model and a constant input impedance for the tag's power harvesting circuit, however, actual circuit constructions exhibit both a non-linear harvesting efficiency [30] and variable input impedance with respect to input power and frequency [38, 39]. In practice, RFID tags will encounter wide ranges of input power, so these two effects require further study to understand their role in system design. First, the proposed constellation design procedure no longer has its guarantee of optimality if the harvester's efficiency is non-linear. Second, the backscatter constellation geometry will suffer distortion if the harvester's equivalent impedance differs from its assumed value, which may have a serious impact on the link's symbol error performance. The first step for future work should quantify the sensitivity of the coded modulation design to varying input power through experimental measurements.

Irregardless of the tag's equivalent impedance, there exists a set of modulator impedances that generate any arbitrary constellation geometry if the tag's IC impedance stays constant. So if the tag is aware of its input power and if it can set its modulator impedance to any arbitrary value, the tag can prevent constellation distortion by adapting its modulator impedance in response to the measured input power. A tag's backscatter modulator typically must switch between a fixed set of values, hence, a tag cannot set any arbitrary

impedance state. Alternatively, by 'over designing' the modulator circuit with respect to the constellation size and the number of modulator states, e.g. a 4-QAM constellation requires a 16 state modulator (such modulators are demonstrated by [43])[1], the tag has a much wider choice of backscatter states. Based on the measured input power, the tag chooses the set of modulator states that map to the desired constellation the best.

This direction of research has numerous challenges from circuit/device design and system optimization. Through measuring the power envelope at the tag's antenna, the tag needs to estimate the harvesting circuit's impedance from a stored lookup table. Some challenges are the design of the power estimation circuit and generating this table for custom made devices and for mass produced tags. Another challenge is how to choose the modulator states to meet a desired $\mathcal{B}_{\text{loss}}$ metric, since the harvester circuit is no longer conjugate matched to the antenna in general and the impedance states are at discrete values.

7.2.2 RFID MIMO System Design

The analysis of Chapters 3 and 4 assumes a constant transmission of the reader's carrier signal, i.e. CW transmission at each of the reader's antennas. In conjunction with the assumption of complex normal distributed path gains, the (M, L, N) DBC is statistically equivalent to the $(1, L, N)$ DBC. Obviously, the choice of M plays no role in system performance, and increasing diversity or multiplexing gains require additional tag antennas (L) and reader receive antennas (N). The tag's form factor does not allow for the placement of more than two to four antennas at the tag, so increasing the diversity order through additional tag antennas remains limited.

Evaluating the potential of increasing diversity through reader transmit antennas (M) presents one direction for future work for both to experimental verification and theoretical code design. Instead of CW transmission, the reader should transmit a precoding sequence, similar to space-time code designs, over both space and time. Currently it is not known if precoding scheme will offer any diversity gains, so there exists ample opportunity to quantify the feasible gains and to construct precoding sequences. There also remains the practicality

[1]I thank Prof. Matt Reynolds and Vamsi Talla for suggesting the over-design/adaptive-impedance approach.

of decoding the backscatter in the presence the reader's changing carrier signal. The receiver will encounter modulated self-interference, and the observed channel state information of the backscatter link will change each time the reader's transmit signal changes. Clearly this may increase the difficulty of decoding the tag's backscatter.

7.2.3 MAC Enhancements: HARQ

Within the contributions of this dissertation, the EPC Gen2 HARQ enhancements proposed in Chapter 5 have the most potential for immediate realizable gains. Although there remains the experimental verification of the gains observed by theory and simulation, since it is not possible to model all the dynamical effects occurring within real RFID systems. Namely, the channel state may change during a tag's EPC transmission and the distribution SNR gains amongst a tag population may vary based on the application. The simulation model and decoder used within this work assumes a constant channel state, and the corresponding throughput analysis assume equal SNR across all RFID tags.

HARQ experiment work should focus on practical receiver design and implementation, for which there exists examples [140,141] of EPC Gen2 receivers that may serve as a starting point for experimental work. Additionally, the implementation of new anti-collision protocols on DSP/FPGA hardware [142] is one example of validating system level enhancements with realistic experiments. Automatically adapting the HARQ decoder algorithm (either Chase combining or incremental redundancy) to a time varying channel state represents the primary challenge for experimental work. One possible approach may use additional pilot symbols to assist with channel state estimation [143], and another approach might use the RCPC code structure to assist in channel state estimation [144].

With respect to building HARQ encoders on the tag, the WISP [15] is one such platform that allows for the implementation of repetition coding and RCPC codes in software. Lastly, there remains the issue of further validating the reliability of blanking the reader's CW to indicating successful EPC decoding, which was first proposed and demonstrated by [88]. It is not known if time varying fading affects the accuracy of CW detection circuit.

7.2.4 MAC Enhancements: IR-FSA

The contributions of Chapter 6 focus on establishing a performance baseline for the IR-FSA protocol. Basic theoretical throughput modeling of IR-FSA suggests large performance gains over conventional random access schemes (e.g. slotted ALOHA), however, there remains a large performance gap between simulation results and the theoretical model. Since IR-FSA's performance depends heavily on the receiver's MUD algorithm, future research should focus on further developing compressive sensing methods for MUD. First, the discussed Reed Muller decoder still has further room for improvement at the encoder and decoder. Second, other compressive sensing methods based on expander graphs [145,146] should be considered as potential candidates for MUD encoders/decoders.

The Reed Muller based MUD algorithm studied in Chapter 6 does not have any theoretical guarantees with respect to maximum likelihood detection for the MUD case or even the single user decoding case. The literature discusses heuristic improvements to the SIC Reed Muller decoder [138, 139], which are related to prior work on single user Reed Muller codes [147, 148]. Additionally, there remains the open question of applying joint iterative decoding of Reed Muller codes in the multi-user scenario [139]. Regarding the performance of the encoder, the improvement of the Reed Muller puncturing pattern design should also be a focus of future work, because simulation results exhibit poor performance at high code rates. Currently, randomly generated erasures comprise the puncturing pattern used within this work, and while deterministic puncturing patterns exist for Reed Muller codes [149], the suitability of this approach for MUD still needs to be evaluated.

Lastly, future work should also investigate alternative encoder/decoder architectures such as those based on expander graphs [145, 146]. Driven by the same motivation for using Reed Muller codes in compressive sensing, expander graphs have sub-linear decoding complexity and reported have good sparse recovery performance. As discussed by [146], expander graphs have structure similar to LDPC codes [134] and the recovery algorithm follows the well known belief propagation decoding algorithm. It remains unknown if MUD based on expander graphs offers better performance than Reed Muller codes with respect to error performance and ease of constructing rate compatible codes.

BIBLIOGRAPHY

[1] J. Hagenauer. Rate-compatible punctured convolutional codes (rcpc codes) and their applications. *Communications, IEEE Transactions on*, 36(4):389 –400, apr 1988.

[2] D. Dobkin. *The RF in RFID*. Elsevier Inc., 2008.

[3] Klaus Finkenzeller. *RFID Handbook*. John Wiley & Sons, 2003.

[4] J. Landt. The history of rfid. *Potentials, IEEE*, 24(4):8–11, 2005.

[5] Harry Stockman. Communication by means of reflected power. *Proceedings of the IRE*, 36(10):1196–1204, 1948.

[6] P. Nikitin. Leon Theremin (Lev Termen). *Antennas and Propagation Magazine, IEEE*, 54(5):252–257, 2012.

[7] F. Fuschini, C. Piersanti, F. Paolazzi, and G. Falciasecca. Analytical approach to the backscattering from uhf rfid transponder. *Antennas and Wireless Propagation Letters, IEEE*, 7:33 –35, 2008.

[8] Alfred R. Koelle, Steven W. Depp, and R.W. Freyman. Short-range radio-telemetry for electronic identification, using modulated rf backscatter. *Proceedings of the IEEE*, 63(8):1260–1261, 1975.

[9] Michael Buettner. *Backscatter Protocols and Energy-Efficient Computing for RF-Power Devices*. PhD thesis, University of Washington, 2012.

[10] Ieee standard for local and metropolitan area networks– part 15.4: Low-rate wireless personal area networks (lr-wpans) amendment 2: Active radio frequency identification (rfid) system physical layer (phy). *IEEE Std 802.15.4f-2012 (Amendment to IEEE Std 802.15.4-2011)*, pages 1–72, 2012.

[11] V. Chawla and Dong-Sam Ha. An overview of passive rfid. *Communications Magazine, IEEE*, 45(9):11–17, 2007.

[12] S. Preradovic and N.C. Karmakar. Chipless rfid: Bar code of the future. *Microwave Magazine, IEEE*, 11(7):87–97, 2010.

[13] G. De Vita and G. Iannacone. Design Criteria for the RF Section of UHF and Microwave Passive RFID Transponders. *IEEE Transactions on Microwave Theory and Techniques*, 53(9):2978–2990, September 2005.

[14] U. Karthaus and M. Fischer. Fully integrated passive UHF RFID transponder IC with 16.7- μ W minimum RF input power. *Solid-State Circuits, IEEE Journal of*, 38(10):1602 – 1608, oct. 2003.

[15] A.P. Sample, D.J. Yeager, P.S. Powledge, and J.R. Smith. Design of a Passively-Powered, Programmable Sensing Platform for UHF RFID Systems. In *RFID, 2007. IEEE International Conference on*, pages 149 –156, 26-28 2007.

[16] K. Kurokawa. Power waves and the scattering matrix. *Microwave Theory and Techniques, IEEE Transactions on*, 13(2):194–202, 1965.

[17] R. Chakraborty, S. Roy, and V. Jandhyala. Revisiting rfid link budgets for technology scaling: Range maximization of rfid tags. *Microwave Theory and Techniques, IEEE Transactions on*, 59(2):496 –503, 2011.

[18] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[19] J.D. Griffin and G.D. Durgin. Gains For RF Tags Using Multiple Antennas. *Antennas and Propagation, IEEE Transactions on*, 56(2):563 –570, feb. 2008.

[20] S. Roy, V. Jandhyala, J.R. Smith, D.J. Wetherall, B.P. Otis, R. Chakraborty, M. Buettner, D.J. Yeager, You-Chang Ko, and A.P. Sample. Rfid: From supply chains to sensor nets. *Proceedings of the IEEE*, 98(9):1583 –1592, sept. 2010.

[21] O.B. Akan, M.T. Isik, and B. Baykal. Wireless passive sensor networks. *Communications Magazine, IEEE*, 47(8):92–99, 2009.

[22] Michael Buettner, Ben Greenstein, Alanson Sample, Joshua R Smith, and David Wetherall. Revisiting smart dust with rfid sensor networks. *Proceedings of the 7th ACM Workshop on Hot Topics in Networks (HotNets-VII)*, 2008.

[23] D.J. Yeager, J. Holleman, R. Prasad, J.R. Smith, and B.P. Otis. Neuralwisp: A wirelessly powered neural interface with 1-m range. *Biomedical Circuits and Systems, IEEE Transactions on*, 3(6):379–387, 2009.

[24] D. Yeager, Fan Zhang, A. Zarrasvand, and B.P. Otis. A 9.2 microamp gen 2 compatible uhf rfid sensing tag with -12dbm sensitivity and 1.25 microvolt rms input-referred noise floor. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2010 IEEE International*, pages 52–53, 2010.

[25] S.J. Thomas, R.R. Harrison, A. Leonardo, and M.S. Reynolds. A battery-free multichannel digital neural/emg telemetry system for flying insects. *Biomedical Circuits and Systems, IEEE Transactions on*, 6(5):424–436, 2012.

[26] D.J. Yeager, P.S. Powledge, R. Prasad, D. Wetherall, and J.R. Smith. Wirelessly-Charged UHF Tags for Sensor Data Collection. In *RFID, 2008 IEEE International Conference on*, pages 320 –327, 16-17 2008.

[27] EPC Global. *EPC Global RFID Protocol Class 1 Generation 2 UHF*, January 2005.

[28] Dimitri Bertsekas and Robert Gallager. *Data networks (2nd ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1992.

[29] Michael Buettner and David Wetherall. An Empirical Study of UHF RFID Performance. In *MobiCom*, 2008.

[30] T. Le, K. Mayaram, and T. Fiez. Efficient far-field radio frequency energy harvesting for passively powered sensor networks. *Solid-State Circuits, IEEE Journal of*, 43(5):1287–1302, 2008.

[31] V. Tarokh, N. Seshadri, and A.R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *Information Theory, IEEE Transactions on*, 44(2):744 –765, mar 1998.

[32] Lizhong Zheng and D.N.C. Tse. Diversity and multiplexing: a fundamental tradeoff in multiple-antenna channels. *Information Theory, IEEE Transactions on*, 49(5):1073–1096, 2003.

[33] M. Simon and D. Divsalar. Some Interesting Observations for Certain Line Codes with Application to RFID. *Communications, IEEE Transactions on*, 54(4):583 – 586, april 2006.

[34] Stewart Thomas and Matthew S. Reynolds. QAM Backscatter for Passive UHF RFID Tags. In *RFID, 2010 IEEE International Conference on*, pages 210 –214, 14-16 2010.

[35] C. Boyer and S. Roy. Coded qam backscatter modulation for rfid. *Communications, IEEE Transactions on*, 60(7):1925–1934, 2012.

[36] B. Masnick and J. Wolf. On linear unequal error protection codes. *Information Theory, IEEE Transactions on*, 13(4):600 – 607, October 1967.

[37] Impinj - Speedway Reader documentation, 2008.

[38] V. Pillai. Impedance matching in rfid tags: to which impedance to match? In *Antennas and Propagation Society International Symposium 2006, IEEE*, pages 3505–3508, 2006.

[39] T. Bauernfeind, K. Preis, G. Koczka, S. Maier, and O. Biro. Influence of the nonlinear uhf-rfid ic impedance on the backscatter abilities of a t-match tag antenna design. *Magnetics, IEEE Transactions on*, 48(2):755–758, 2012.

[40] R.C. Hansen. Relationships between antennas as scatterers and as radiators. *Proceedings of the IEEE*, 77(5):659–662, 1989.

[41] J. Kimionis, A. Bletsas, and J.N. Sahalos. Bistatic backscatter radio for tag readrange extension. In *RFID-Technologies and Applications (RFID-TA), 2012 IEEE International Conference on*, pages 356–361, 2012.

[42] G. Vannucci, A. Bletsas, and D. Leigh. A software-defined radio system for backscatter sensor networks. *Wireless Communications, IEEE Transactions on*, 7(6):2170–2179, 2008.

[43] S.J. Thomas, E. Wheeler, J. Teizer, and M.S. Reynolds. Quadrature amplitude modulated backscatter in passive and semipassive uhf rfid systems. *Microwave Theory and Techniques, IEEE Transactions on*, 60(4):1175–1182, 2012.

[44] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[45] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1997.

[46] U. Madhow. *Fundamentals of Digital Communication*. Cambridge, 2008.

[47] R.H. Morelos-Zaragoza and Shu Lin. Qpsk block-modulation codes for unequal error protection. *Information Theory, IEEE Transactions on*, 41(2):576 –581, March 1995.

[48] J.D. Griffin and G.D. Durgin. Multipath fading measurements at 5.8 ghz for backscatter tags with multiple antennas. *Antennas and Propagation, IEEE Transactions on*, 58(11):3693 – 3700, Nov. 2010.

[49] J.D. Griffin. *High-frequency modulated-backscatter communication using multiple antennas*. PhD thesis, Georgia Institute of Technology, 2009.

[50] A. Ingram, M. Demirkol, and D. Kim. Transmit diversity and spatial multiplexing for rf links using modulated backscatter. In *Signals, Systems, and Electronics, 2001 International Symposium*, pages 1–4, July 2001.

[51] C. Angerer, R. Langwieser, G. Maier, and M. Rupp. Maximal ratio combining receivers for dual antenna rfid readers. In *Wireless Sensing, Local Positioning, and RFID, 2009. IMWS 2009. IEEE MTT-S International Microwave Workshop on*, pages 1 –4, sept. 2009.

[52] R. Langwieser, C. Angerer, and A.L. Scholtz. A uhf frontend for mimo applications in rfid. In *Radio and Wireless Symposium (RWS), 2010 IEEE*, pages 124 –127, jan. 2010.

[53] Do-Yun Kim, Han-Shin Jo, Hyungoo Yoon, Cheol Mun, Byung-Jun Jang, and Jong-Gwan Yook. Reverse-link interrogation range of a uhf mimo-rfid system in nakagami-fading channels. *Industrial Electronics, IEEE Transactions on*, 57(4):1468 –1477, april 2010.

[54] Chen He and Z.J. Wang. Closed-form ber analysis of non-coherent fsk in miso double rayleigh fading/rfid channel. *Communications Letters, IEEE*, 15(8):848 –850, august 2011.

[55] Chen He and J. Wang. Impact of the correlation between forward and backscatter channels on rfid system performance. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 3540–3543, May 2011.

[56] Chen He and Z.J. Wang. Gains by a space-time-code based signaling scheme for multiple-antenna rfid tags. In *Electrical and Computer Engineering (CCECE), 2010 23rd Canadian Conference on*, pages 1 –4, may 2010.

[57] G. Smietanka and J. Gotze. Modeling and simulation of miso diversity for uhf rfid communication. In *Computer Science and Information Systems (FedCSIS), 2012 Federated Conference on*, pages 813–820, Sept 2012.

[58] S.M. Alamouti. A simple transmit diversity technique for wireless communications. *Selected Areas in Communications, IEEE Journal on*, 16(8):1451 –1458, oct 1998.

[59] M. Uysal. Diversity Analysis of Space-Time Coding in Cascaded Rayleigh Fading Channels. *Communications Letters, IEEE*, 10(3):165 – 167, mar 2006.

[60] Chen He, Xun Chen, Zhen Jane Wang, and Weifeng Su. On the performance of mimo rfid backscattering channels. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):357, 2012.

[61] Feng Zheng and Thomas Kaiser. A space-time coding approach for rfid mimo systems. *EURASIP Journal on Embedded Systems*, 2012(1):9, 2012.

[62] Feng Zheng and T. Kaiser. On the transmit signal design at the reader for rfid mimo systems. In *RFID Technology (EURASIP RFID), 2012 Fourth International EURASIP Workshop on*, pages 59–64, Sept 2012.

[63] Chen He and Z.J. Wang. Ser of orthogonal space-time block codes over rician and nakagami-m rf backscattering channels. *Vehicular Technology, IEEE Transactions on*, 2014. to appear.

[64] D. Arnitz and M.S. Reynolds. Wireless power transfer optimization for nonlinear passive backscatter devices. In *RFID (RFID), 2013 IEEE International Conference on*, pages 245–252, April 2013.

[65] A. Hasan, Chenming Zhou, and J.D. Griffin. The phase difference method for transmit diversity in monostatic rfid systems. In *Antennas and Propagation Society International Symposium (APSURSI), 2012 IEEE*, pages 1–2, July 2012.

[66] E. Denicke, M. Henning, H. Rabe, and B. Geck. The application of multiport theory for mimo rfid backscatter channel measurements. In *Microwave Conference (EuMC), 2012 42nd European*, pages 522–525, Oct 2012.

[67] S. Loyka and A. Kouki. On mimo channel capacity, correlations, and keyholes: analysis of degenerate channels. *Communications, IEEE Transactions on*, 50(12):1886 – 1888, dec 2002.

[68] G. Levin and S. Loyka. On the outage capacity distribution of correlated keyhole mimo channels. *Information Theory, IEEE Transactions on*, 54(7):3232–3245, July 2008.

[69] D. Chizhik, G.J. Foschini, and R.A. Valenzuela. Capacities of multi-element transmit and receive antennas: Correlations and keyholes. *Electronics Letters*, 36(13):1099 –1100, jun 2000.

[70] P. Almers, F. Tufvesson, and A.F. Molisch. Keyhole Effect in MIMO Wireless Channels: Measurements and Theory. *Wireless Communications, IEEE Transactions on*, 5(12):3596 –3604, december 2006.

[71] A. Leon-Garcia. *Probability, Statistics, and Random Processes for Electrical Engineering*. Pearson Prentice Hall, 2008.

[72] G. Casella and R.L. Berger. *Statistical Inference*. Duxbury, 2002.

[73] T. Cover and J. Thomas. *Elements of Information Theory*. John Wiley & Sons, 2006.

[74] J. Salo, H.M. El-Sallabi, and P. Vainikainen. The distribution of the product of independent rayleigh random variables. *Antennas and Propagation, IEEE Transactions on*, 54(2):639 – 643, feb. 2006.

[75] V. Tarokh, H. Jafarkhani, and A.R. Calderbank. Space-time block codes from orthogonal designs. *Information Theory, IEEE Transactions on*, 45(5):1456 –1467, jul 1999.

[76] E. Larsson and P. Stocia. *Space-Time Block Coding for Wireless Communications.* Cambridge University Press, 2008.

[77] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products.* Academic Press, 1980.

[78] NIST Digital Library of Mathematical Functions. http://dlmf.nist.gov/, Release 1.0.6 of 2013-05-06.

[79] S.J. Thomas and M.S. Reynolds. A 96 mbit/sec, 15.5 pj/bit 16-qam modulator for uhf backscatter communication. In *RFID (RFID), 2012 IEEE International Conference on*, pages 185 –190, april 2012.

[80] Sheng Yang and J. C Belfiore. Diversity-multiplexing tradeoff of double scattering mimo channels. *Information Theory, IEEE Transactions on*, 57(4):2027–2034, 2011.

[81] Emre Telatar. Capacity of multi-antenna gaussian channels. *European Transactions on Telecommunications*, 10(6):585–595, 1999.

[82] N. Bleistein and R. Handlesman. *Asymptotic Expansion of Integrals.* Dover Publications, 1975.

[83] A. Tulino and S. Verdu. *Random Matrix Theory and Wireless Communication.* Now Publishers, 2004.

[84] R. Narasimhan. Finite-snr diversity multiplexing tradeoff for correlated rayleigh and rician mimo channels. *Information Theory, IEEE Transactions on*, 52(9):3965–3979, 2006.

[85] You-Chang Ko, S. Roy, J.R. Smith, Hyong-Woo Lee, and Choong-Ho Cho. Rfid mac performance evaluation based on iso/iec 18000-6 type c. *Communications Letters, IEEE*, 12(6):426 –428, june 2008.

[86] Inwhee Joe. A low-power hybrid arq scheme for the rfid system. HPCC'06, pages 535–541, Berlin, Heidelberg, 2006. Springer-Verlag.

[87] Pengyu Zhang, Jeremy Gummeson, and Deepak Ganesan. Blink: a high throughput link layer for backscatter communication. MobiSys '12, pages 99–112, New York, NY, USA, 2012. ACM.

[88] Jue Wang, Haitham Hassanieh, Dina Katabi, and Piotr Indyk. Efficient and reliable low-power backscatter networks. SIGCOMM '12, pages 61–72, New York, NY, USA, 2012. ACM.

[89] Grzegorz Smietanka and Juergen Goetze. Performance of bch codes and transmit diversity for miso uhf rfid communication. In *Smart Objects, Systems and Technologies (SmartSysTech), Proceedings of 2012 European Conference on*, pages 1–6, June 2012.

[90] G. Smietanka, J. Geldmacher, and J. Gotze. Error detection based on correlation analysis for bch encoded uhf rfid communication. In *Industrial Technology (ICIT), 2013 IEEE International Conference on*, pages 1666–1670, Feb 2013.

[91] A. Schantin and C. Ruland. Retransmission strategies for rfid systems using the epc protocol. In *RFID-Technologies and Applications (RFID-TA), 2013 IEEE International Conference on*, pages 1–6, Sept 2013.

[92] A. Schantin. Iterative decoding of baseband and channel codes in a long-range rfid system. In *Industrial Technology (ICIT), 2013 IEEE International Conference on*, pages 1671–1676, Feb 2013.

[93] R. Morelos-Zaragoza. Unequal error protection with crc-16 bits in epc class-1 generation-2 uhf rfid systems. In *Information Theory and its Applications (ISITA), 2012 International Symposium on*, pages 36–40, Oct 2012.

[94] E. Soijanin, N. Varnica, and P. Whiting. Punctured vs rateless codes for hybrid arq. In *ITW '06 Punta del Este. IEEE*, pages 155 –159, march 2006.

[95] C. Lott, O. Milenkovic, and E. Soljanin. Hybrid arq: Theory, state of the art and future directions. In *2007 IEEE ITW*, pages 1 –5, july 2007.

[96] D. Chase. Code combining–a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets. *Communications, IEEE Transactions on*, 33(5):385 – 393, may 1985.

[97] Jr. Costello, D.J., J. Hagenauer, H. Imai, and S.B. Wicker. Applications of error-control coding. *Information Theory, IEEE Transactions on*, 44(6):2531 –2560, oct 1998.

[98] S. Wilson. *Digital Modulation and Coding*. Prentice Hall, 1996.

[99] Sheldon M. Ross. *Stochastic Processes*. Wiley, 2 edition.

[100] C. Floerkemeier. Transmission control scheme for fast rfid object identification. In *PerCom Workshops 2006.*, pages 6 pp. –462, march 2006.

[101] U. Erez, M.D. Trott, and G.W. Wornell. Rateless coding for gaussian channels. *Information Theory, IEEE Transactions on*, 58(2):530 –547, feb. 2012.

[102] D.N. Rowitch and L.B. Milstein. On the performance of hybrid fec/arq systems using rate compatible punctured turbo (rcpt) codes. *Communications, IEEE Transactions on*, 48(6):948 –959, jun 2000.

[103] G. Caire, G. Taricco, and E. Biglieri. Bit-interleaved coded modulation. *Information Theory, IEEE Transactions on*, 44(3):927 –946, may 1998.

[104] Aditya Gudipati and Sachin Katti. Strider: automatic rate adaptation and collision handling. SIGCOMM '11, pages 158–169, New York, NY, USA, 2011. ACM.

[105] Jonathan Perry, Peter A. Iannucci, Kermin E. Fleming, Hari Balakrishnan, and Devavrat Shah. Spinal codes. *SIGCOMM Comput. Commun. Rev.*, 42(4):49–60, August 2012.

[106] C. Floerkemeier and M. Wille. Comparison of transmission schemes for framed aloha based rfid protocols. In *Applications and the Internet Workshops, 2006. SAINT Workshops 2006. International Symposium on*, pages 4 pp.–97, 2006.

[107] L. Barletta, F. Borgonovo, and M. Cesana. Performance of dynamic-frame-aloha protocols: Closing the gap with tree protocols. In *Ad Hoc Networking Workshop (Med-Hoc-Net), 2011 The 10th IFIP Annual Mediterranean*, pages 9–16, 2011.

[108] Haifeng Wu, Yu Zeng, Jihua Feng, and Yu Gu. Binary tree slotted aloha for passive rfid tag anticollision. *Parallel and Distributed Systems, IEEE Transactions on*, 24(1):19–31, 2013.

[109] Y. Maguire and R. Pappu. An optimal q-algorithm for the iso 18000-6c rfid protocol. *Automation Science and Engineering, IEEE Transactions on*, 6(1):16–24, 2009.

[110] Lei Zhu and T.P. Yum. Optimal framed aloha based anti-collision algorithms for rfid systems. *Communications, IEEE Transactions on*, 58(12):3583–3592, 2010.

[111] C. Floerkemeier. Bayesian transmission strategy for framed aloha based rfid protocols. In *RFID, 2007. IEEE International Conference on*, pages 228–235, 2007.

[112] Gianluigi Liva. Graph-based analysis and optimization of contention resolution diversity slotted aloha. *Communications, IEEE Transactions on*, 59(2):477–487, 2011.

[113] C. Stefanovic, P. Popovski, and D. Vukobratovic. Frameless aloha protocol for wireless networks. *Communications Letters, IEEE*, 16(12):2087–2090, 2012.

[114] S. Verdu. *Multiuser Detection*. Cambridge University Press, 1998.

[115] G. Caire and D. Tuninetti. The throughput of hybrid-arq protocols for the gaussian collision channel. *Information Theory, IEEE Transactions on*, 47(5):1971–1988, 2001.

[116] P. Minero, M. Franceschetti, and D.N.C. Tse. Random access: An information-theoretic perspective. *Information Theory, IEEE Transactions on*, 58(2):909–930, 2012.

[117] M. Honig, U. Madhow, and S. Verdu. Blind adaptive multiuser detection. *Information Theory, IEEE Transactions on*, 41(4):944–960, 1995.

[118] Lorne Applebaum, Waheed U. Bajwa, Marco F. Duarte, and Robert Calderbank. Asynchronous code-division random access using convex optimization. *Physical Communication*, 5(2):129 – 147, 2012.

[119] Xiao Li, A. Rueetschi, A. Scaglione, and Y.C. Eldar. Compressive link acquisition in multiuser communications. *Signal Processing, IEEE Transactions on*, 61(12):3229–3245, 2013.

[120] Alyson K. Fletcher, Sundeep Rangan, and Vivek K. Goyal. On-off random access channels: A compressed sensing framework. *CoRR*, abs/0903.1022, 2009.

[121] V. Shah-Mansouri, Suyang Duan, Ling-Hua Chang, V.W.S. Wong, and Jwo-Yuh Wu. Compressive sensing based asynchronous random access for wireless networks. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 884–888, 2013.

[122] M.K. Tsatsanis, Ruifeng Zhang, and S. Banerjee. Network-assisted diversity for random access wireless networks. *Signal Processing, IEEE Transactions on*, 48(3):702–711, 2000.

[123] Ruifeng Zhang, N.D. Sidiropoulos, and M.K. Tsatsanis. Collision resolution in packet radio networks using rotational invariance techniques. *Communications, IEEE Transactions on*, 50(1):146–155, 2002.

[124] Young-Han Nam, Praveen Kumar Gopala, and H. El-Gamal. Resolving collisions via incremental redundancy: Arq diversity. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 285–293, 2007.

[125] Rukun Mao and Husheng Li. A novel multiple access scheme via compressed sensing with random data traffic. In *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, pages 1–6, 2010.

[126] Yuzhe Jin, Young-Han Kim, and B.D. Rao. Limits on support recovery of sparse signals via multiple-access communication techniques. *Information Theory, IEEE Transactions on*, 57(12):7877–7892, 2011.

[127] C. Angerer, Robert Langwieser, and M. Rupp. Rfid reader receivers for physical layer collision recovery. *Communications, IEEE Transactions on*, 58(12):3526–3537, 2010.

[128] K. Fyhn, R.M. Jacobsen, P. Popovski, A. Scaglione, and T. Larsen. Multipacket reception of passive uhf rfid tags: A communication theoretic approach. *Signal Processing, IEEE Transactions on*, 59(9):4225–4237, 2011.

[129] A. Bletsas, J. Kimionis, A.G. Dimitriou, and G.N. Karystinos. Single-antenna coherent detection of collided fm0 rfid signals. *Communications, IEEE Transactions on*, 60(3):756–766, 2012.

[130] J.G. Andrews. Interference cancellation for cellular systems: a contemporary overview. *Wireless Communications, IEEE*, 12(2):19–29, 2005.

[131] C. Schlegel and D. Truhachev. Multiple access demodulation in the lifted signal graph with spatial coupling. *Information Theory, IEEE Transactions on*, 59(4):2459–2470, 2013.

[132] U. Niesen, U. Erez, D. Shah, and Gregory W. Wornell. Cth05-2: Rateless codes for the gaussian multiple access channel. In *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, pages 1–5, 2006.

[133] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[134] T. Moon. *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley Interscience, 2005.

[135] S.D. Howard, A.R. Calderbank, and S.J. Searle. A fast reconstruction algorithm for deterministic compressive sensing using second order reed-muller codes. In *Information Sciences and Systems, 2008. CISS 2008. 42nd Annual Conference on*, pages 11–15, 2008.

[136] Lei Zhang, Jun Luo, and Dongning Guo. Neighbor discovery for wireless networks via compressed sensing. *Performance Evaluation*, 70(78):457 – 471, 2013.

[137] Lei Zhang and Dongning Guo. Neighbor discovery in wireless networks using compressed sensing with reed-muller codes. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2011 International Symposium on*, pages 154–160, 2011.

[138] S. Searle, S. Howard, and Calderbank. Sparse signal processing applications of reed-muller code sets. In *DASP 2011 Workshop*, 2011.

[139] V. Aggarwal, L. Applebaum, A. Bennatan, A.R. Calderbank, S.D. Howard, and S.J. Searle. Enhanced cdma communications using compressed-sensing reconstruction methods. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, pages 1211–1215, 2009.

[140] Feng Ge and Ken Choi. Novel design and implementation for highly sensitive baseband protocol of class-1 generation-2 uhf rfid system. In *Electro/Information Technology (EIT), 2010 IEEE International Conference on*, pages 1–5, 2010.

[141] C. Angerer and M. Rupp. Advanced synchronisation and decoding in rfid reader receivers. In *Radio and Wireless Symposium, 2009. RWS '09. IEEE*, pages 59–62, 2009.

[142] Xiaolin Jia, Quanyuan Feng, and Lishan Yu. Stability analysis of an efficient anti-collision protocol for rfid tag identification. *Communications, IEEE Transactions on*, 60(8):2285–2294, 2012.

[143] J.K. Cavers. An analysis of pilot symbol assisted modulation for rayleigh fading channels [mobile radio]. *Vehicular Technology, IEEE Transactions on*, 40(4):686–693, 1991.

[144] R. Raheli, A. Polydoros, and Ching-Kae Tzou. Per-survivor processing: a general approach to mlse in uncertain environments. *Communications, IEEE Transactions on*, 43(234):354–364, 1995.

[145] Weiyu Xu and B. Hassibi. Efficient compressive sensing with deterministic guarantees using expander graphs. In *Information Theory Workshop, 2007. ITW '07. IEEE*, pages 414–419, 2007.

[146] Sina Jafarpour, Weiyu Xu, B. Hassibi, and R. Calderbank. Efficient and robust compressed sensing using optimized expander graphs. *Information Theory, IEEE Transactions on*, 55(9):4299–4308, 2009.

[147] I. Dumer. Soft-decision decoding of reed-muller codes: a simplified algorithm. *Information Theory, IEEE Transactions on*, 52(3):954–963, 2006.

[148] I. Dumer and K. Shabunov. Soft-decision decoding of reed-muller codes: recursive lists. *Information Theory, IEEE Transactions on*, 52(3):1260–1266, 2006.

[149] O. Kapralova and I. Dumer. Spherically punctured reed-muller codes. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 1047–1051, 2013.

Appendix A

RFID BACKSCATTER LINK BUDGET

To relate the physical properties of the tag and reader during backscatter, we briefly restate and summarize the key developments regarding the RFID link budget from [17]. The electric field at the tag $\mathcal{E}_{Tag}$ depends on the reader's effective isotropic radiated power (EIRP) P_{EIRP}, the distance between the tag and reader l, and the polarization match coefficient ρ. P_{EIRP} is related by the raw transmit power P_{rdr} and reader antenna gain G_{rdr}: $P_{rdr}G_{rdr} = P_{EIRP}$

$$\mathcal{E}_{Tag} = \sqrt{\frac{Z_0 \rho P_{EIRP}}{4\pi l^2}} \tag{A.1}$$

where Z_0 is free-space impedance. The open circuit voltage V_{OC} is proportional to the $\mathcal{E}_{Tag}$ and given by

$$V_{OC} = \alpha_{Tag}\mathcal{E}_{Tag} = \alpha_{Tag}\sqrt{\frac{Z_0 \rho P_{EIRP}}{4\pi l^2}} \tag{A.2}$$

The α_{Tag} variable relates the physical geometry of the antenna to the incident electric field. The electric field generated by the tag's antenna is proportional to current passing through the antenna, I_{ant}. At the reader's antenna, the reflected electric field is

$$\mathcal{E}_{rdr} = I_{ant}\frac{Z_0 \alpha_{Tag}}{2\lambda l} \tag{A.3}$$

where λ is the wavelength of the carrier wave. $\mathcal{E}_{rdr}$ induces a voltage V_{rdr} across the reader's antenna, which is measured by the receiver to make a bit decision.

$$V_{rdr} = \alpha_{rdr}\mathcal{E}_{rdr} = \alpha_{rdr} I_{ant}\frac{Z_0 \alpha_{Tag}}{2\lambda l} \tag{A.4}$$

Lastly, the reader's receiver is subject to noise from thermal affects and from phase noise leakage from the transmitter circuit. It is understood in RFID applications that phase noise dominates thermal noise [2]. A typical observed phase noise power is $-115\frac{\text{dBc}}{\text{Hz}}$. We assume there is a power reduction factor of 15dB between the receiver and transmitter

circuits, and a further reduction of 50dB for the conversion of phase noise to amplitude noise. ($-115\frac{\text{dBc}}{\text{Hz}} - 15\text{dB} - 50\text{dB} = -180\frac{\text{dBc}}{\text{Hz}}$)

$$N_0 = -180\frac{\text{dBc}}{\text{Hz}} \tag{A.5}$$

Relating the above parameters to the observed SNR at the reader SNR $= \frac{E}{N}$ for 2-QAM modulation is achieved by noting that the instantaneous signal power

$$\begin{aligned}
E &= \left(\frac{|x_1 - x_0|}{2}\right)^2 \\
&= \left(\frac{1}{2\sqrt{2}}|V_{rdr,0} - V_{rdr,1}|\right)^2 \\
&= \left(\frac{1}{2\sqrt{2}}|I_{ant,0} - I_{ant,1}|\frac{|\alpha_{rdr}||\alpha_{Tag}|Z_0}{2\lambda l}\right)^2 \\
&= \left(\frac{1}{2\sqrt{2}}|\eta_0 - \eta_1|\frac{|V_{OC}|}{R_{ant}}\frac{|\alpha_{rdr}||\alpha_{Tag}|Z_0}{2\lambda l}\right)^2 \\
&= |\eta_0 - \eta_1|^2\left[\frac{1}{2\sqrt{2}}\frac{1}{R_{ant}}|\alpha_{Tag}|\sqrt{\frac{Z_0\rho P_{EIRP}}{4\pi l^2}}\right]^2 \\
&\quad \left[\frac{|\alpha_{rdr}||\alpha_{Tag}|Z_0}{2\lambda l}\right]^2
\end{aligned} \tag{A.6}$$

and noise power

$$N = P_{rdr}BN_0 \tag{A.7}$$

This equation for SNR can be used to calculate the uplink BER by (2.6) the antipodal signaling BER.

The system parameter values discussed and defined by the previous are listed here:

$$\begin{array}{ccc}
P_{rdr} = 1\text{W} & G_{rdr} = 6\text{dB} & B = 640\text{KHz} \\
|\alpha_{Tag}| = 0.1 & |\alpha_{rdr}| = 0.12 & \lambda = 0.32\text{m} \\
l = 25\text{m} & Z_0 = 377\Omega & \rho = 0.5 \\
R_{ant} = 76\Omega & &
\end{array}$$

Appendix B

GEOMETRIC INTERPRETATION OF P_{LOSS} MINIMIZATION

Define the feasible y region as

$$D = \{y : \ |y - 0.75| \leq d\}, \tag{B.1}$$

and $|y-0.75| = d$ occurs when two symbols lay on the boundary of W (e.g. $|\eta_i - 0.75| = 0.25$).

Without loss of generality, assume $|\eta_1 - \eta_2| = 2\sqrt{E_A}$, $|\eta_0 - \eta_1| = 2\sqrt{E_B}$, and that the boundary condition for W is met by ether η_1 and η_2, or η_0 and η_1. From the rectangular geometry of the constellation and symmetry of the region D, there are two candidates for d:

$$d_1 = \sqrt{0.25^2 - E_A} - \sqrt{E_B}, \text{ if } |\eta_1 - 0.75| = 0.25, \ |\eta_2 - 0.75| = 0.25 \tag{B.2}$$

or

$$d_2 = \sqrt{0.25^2 - E_B} - \sqrt{E_A}, \text{ if } |\eta_0 - 0.75| = 0.25, \ |\eta_1 - 0.75| = 0.25. \tag{B.3}$$

Equality between d_1 and d_2 occurs when either $E_A = E_B$ or $E_A + E_B = 0.25^2$. The latter occurs when all four constellation points lay on the boundary of W, which results in the trivial case of $d = 0$.

The remaining step to prove is for what values of E_A and E_B does $d_1 > d_2$ hold true. Perform a change of variables such that $a = \sqrt{E_A}$ and $b = \sqrt{E_B}$, and define the two functions

$$f_1(x) = \sqrt{0.25^2 - x^2} - b \tag{B.4}$$

$$f_2(x) = \sqrt{0.25^2 - b^2} - x. \tag{B.5}$$

$$\tag{B.6}$$

Assume that b is kept fixed and x varies with respect to $b \leq x \leq \sqrt{0.25^2 - b^2}$. It is easy to verify the following: 1) $f_1(x)$ is a concave function with respect to x, 2) $f_2(x)$ is a linear

function with respect to x, and 3) $f_1(b) = f_2(b)$ and $f_1(\sqrt{0.25^2 - b^2}) = f_1(\sqrt{0.25^2 - b^2})$. Then by the definition of convexity and linearity, the following is true

$$f_1(\theta b + (1-\theta)\sqrt{0.25^2 - b^2}) \geq \theta f_1(b) + (1-\theta) f_1(\sqrt{0.25^2 - b^2}) \tag{B.7}$$

$$= \theta f_2(b) + (1-\theta) f_2(\sqrt{0.25^2 - b^2}) \tag{B.8}$$

$$= f_2(\theta b + (1-\theta)\sqrt{0.25^2 - b^2}). \tag{B.9}$$

Therefore if $a > b$, then $f_1(a) > f_2(a)$ is true. Correspondingly, this implies

$$\{y : \ |y - 0.75| \leq \sqrt{0.25^2 - E_A} - \sqrt{E_B}, \ E_A > E_B\} \tag{B.10}$$

Appendix C

PROOFS FOR CHAPTER 4

C.1 Proof of Lemma 4.2.3

Following [32] [80], recall that $p(\boldsymbol{\alpha}|\boldsymbol{\beta})$ follows the Wishart eigenvalue distribution defined by (4.11) for the $N \geq L$, under the change of variables $a_i = \mathrm{SNR}^{-\alpha_i}$ and $b_i = \mathrm{SNR}^{-\beta_i}$. The RHS product term in (4.11) approximates to

$$\prod_{i<j}^{L} \frac{(\mathrm{SNR}^{-\alpha_i} - \mathrm{SNR}^{-\alpha_j})}{(\mathrm{SNR}^{-\beta_i} - \mathrm{SNR}^{-\beta_j})} \doteq \mathrm{SNR}^{-\sum_{i<j}^{L}(\alpha_i - \beta_i)} \tag{C.1}$$

due to the identity $\mathrm{SNR}^{-\alpha_i} - \mathrm{SNR}^{-\alpha_j} \doteq \mathrm{SNR}^{-\alpha_i}$ for $\alpha_i < \alpha_j$ [80]. Due to lemma 1 in [80], the determinant term in (4.11) approximates to

$$\exp\left[-\sum_{i=1}^{L} \mathrm{SNR}^{-(\alpha_i - \beta_i)}\right] \mathrm{SNR}^{-\sum_{i<j}^{L}(\alpha_i - \beta_j)^+} \tag{C.2}$$

for $\alpha_i \geq \beta_i,\ \forall i$; otherwise $\mathrm{Det}[\exp(-\mathrm{SNR}^{-(\alpha_j - \beta_i)})] \to 0$ exponentially in SNR. Combining all the terms in (4.11) and $p(\boldsymbol{\beta})$ in (4.19) gives the form stated in the lemma.

C.2 Proof of Lemma 4.2.4

Following [32] [80], recall that $p(\boldsymbol{\alpha}|\boldsymbol{\beta})$ follows the Wishart eigenvalue distribution defined by (4.12) for the $N < L$. Performing a change of variables $a_i = \mathrm{SNR}^{-\alpha_i}$ and $b_i = \mathrm{SNR}^{-\beta_i}$, the RHS product terms are approximated in the same fashion as the $N \geq L$ case. Applying lemma 2 in [80], the determinant term is approximated as

$$\begin{aligned}\mathrm{Det}[\boldsymbol{\Xi}] \doteq & \exp\left[-\sum_{i=1}^{N} \mathrm{SNR}^{-\alpha_i + \beta_i}\right] \mathrm{SNR}^{-\sum_{i<j}^{N}(\alpha_i - \beta_j)^+} \\ & \cdot \mathrm{SNR}^{-\sum_{i=1}^{N+1}(L-N-1)\beta_i} \mathrm{SNR}^{-\sum_{i=N+2}^{L}(L-i)\beta_i} \\ & \cdot \mathrm{SNR}^{-\sum_{i=1}^{N}\sum_{N+1}^{L}(\alpha_i - \beta_j)^+}\end{aligned} \tag{C.3}$$

if $\alpha_i \geq \beta_i \forall 1 \leq i \leq N$, otherwise the determinant term decreases exponentially in SNR. Combining all the terms in (4.12) and $p(\boldsymbol{\beta})$ in (4.19) gives the form stated in the lemma.

VITA

Colby Boyer was born in California, USA 1987. He obtained his B.S. in Electrical Engineering and Computer Science from the University of California at Berkeley in 2009. He is a student member of IEEE and a member of HKN. Starting in 2009 he began his PhD in Electrical Engineering at the University of Washington, and in 2014 he defended his dissertation. His research interests are in designing MAC and PHY systems for wireless networks. Currently, he is a systems engineer at Fastback Networks in San Jose, CA.

Journals

- Boyer, C.; Roy, S.; ,"Backscatter Communication and RFID: Coding, Energy, and MIMO Analysis," *Communications, IEEE Transactions on*, (invited) to appear

- Boyer, C.; Roy, S.; ,"Space Time Coding for Backscatter RFID," *Wireless Communications, IEEE Transactions on*, vol.12, no.5, pp.2272-2280, May 2013 doi: 10.1109/TWC.2013.031313.120917

- Boyer, C.; Roy, S.; ,"Coded QAM Backscatter Modulation for RFID," *Communications, IEEE Transactions on*, vol.60, no.7, pp.1925-1934, July 2012 doi: 10.1109/TCOMM.2012.051012.110317

Conferences

- Boyer, C.; Dastangoo, S.; ,"Random Access with Imperfect State Information: An Information Theoretic Approach," *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on* pp.1817-1822, 8-11 Sept. 2013 doi: 10.1109/PIMRC.2013.6666438

- Boyer, C.; Roy, S.; ,"Hybrid ARQ Improvements for RFID Systems," *RFID 2013, IEEE International Conference on*, pp.223-230, April 30 2013-May 2 2013

- Phoebus Chen, Parvez Ahammad, Colby Boyer, Shih-I Huang, Leon Lin, Edgar Lobaton, Marci Meingast, Songhwai Oh, Simon Wang, Posu Yan, Allen Yang, Chuohao Yeo, Lung-Chung Chang, Doug Tygar, and Shankar Sastry. "CITRIC: A low-bandwidth wireless camera network platform." *International Conference on Distributed Smart Cameras (ICDSC)*, 2008
  http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-50.pdf