Error Control for Radio Frequency Identification

DISSERTATION
zur Erlangung des Grades eines Doktors
der Ingenieurwissenschaften

vorgelegt von
Dipl.-Ing. Andreas Schantin

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen 2014

1. Gutachter: Prof. Dr. rer. nat. Christoph Ruland
2. Gutachter: Prof. Dr.-Ing. Jürgen Götze

Tag der mündlichen Prüfung: 24.03.2015

urn:nbn:de:hbz:467-9250

UNIVERSITÄT SIEGEN

Institut für
Digitale Kommunikationssysteme

Forschungsberichte

Herausgeber: Univ.-Prof. Dr. Christoph Ruland

Band 31

Andreas Schantin

Error Control for Radio Frequency Identification

SHAKER
VERLAG

Aachen 2015

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: Siegen, Univ., Diss., 2015

Printed in Germany.

ISBN 978-3-8440-3742-5
ISSN 1614-0508

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

To Mariana

Acknowledgements

This thesis was created during my employment at the Chair for Data Communications Systems at the University of Siegen. First and foremost, I would like to express my sincere gratitude to the chairholder and my adviser Prof. Dr. rer. nat. Christoph Ruland, for it was he who enabled and encouraged me to pursue this endeavour in the first place, provided me with the necessary means for my research and the required time next to my regular tasks at the chair.

Further, I would also like to express my gratitude to Prof. Dr.-Ing. Jürgen Götze, who without hesitation agreed to act as the second reviewer for this thesis, Prof. Dr.-Ing. Elmar Griese for his participation in the examination commission, as well as Prof. Dr.-Ing. Horst Bessai for chairing the commission.

I would also like to thank all of my colleagues, past and present. Thanks to Romeo Ayemele Djeujo, Robin Fay, André Groll, Jan Holle, Thomas Koller, Stephan Meyer, Obaid Ur-Rehman, Rainer Schick, Jinsuh Shin, Amir Tabatabaei, Donatus Weber, Tao Wu and Natasa Živić, for the continuous support and valuable discussions. I would also like to specially express my gratitude to Markus Dunte, who encouraged me to join the chair and pursue a doctorate degree in the first place.

Last but not least, I would like to thank my wife Mariana, for the valuable help proofreading this thesis and finding many of the errors, typos and misplaced punctuation marks.

Abstract

This work proposes to use Forward Error Correction (FEC) for the tag-to-reader communication in long-range Radio Frequency Identification (RFID) systems, specifically in the widely used EPCglocal Class-1 Generation-2 (EPCc1g2) systems, to overcome some of the limitations that occur at extended communication distances.

In a long-range RFID system, the reader must be able to identify a great number of tagged items in a reliable fashion, and in a limited amount of time. Cluttered propagation paths, difficult channel conditions, and the low power of the backscattered tag signals, make this task difficult. Especially, when the communication distances are high, as for battery-assisted tags, the signal-to-noise ratio at the reader's receiver is low, and the bit-error probability is high. Nevertheless, the widely used standard protocols offer only rudimentary error-control mechanisms. The use of FEC is not yet established in practical systems and has not been widely researched in the context of RFID.

A model was developed, to describe the impact of errors on the duration of the tag identification, and it is subsequently used, to evaluate the use of convolutional codes and their benefits in terms of the improved identification time, related to the tag-reader distance. For a chosen code, it was shown that the range of the system can be extended by approximately 4 m without increasing the identification time. Furthermore, the application of an iterative decoder architecture is presented that is used to decode the concatenated code, which is formed by the proposed FEC code and the standard EPCc1g2 baseband codes. An improvement of the coding gain of up to 2.8 dB could be shown in simulations, comparing the iterative decoder to a non-iterative one. Additionally, the convergence behavior of the decoder is analyzed and its limitations are discussed.

Since a great number of RFID systems have already been deployed worldwide, standards and standard compatibility should be a great concern, when

proposing extensions for an existing system. It is shown in this work, how a standard EPCc1g2 system may be improved, by using combining at the reader to minimize the number of retransmissions that are caused by frame-errors. Furthermore, it is shown how the standard may be extended with Hybrid Automatic Repeat Request (HARQ) functionality, without losing backward compatibility. Simulations of the identification time as well as the system's throughput show significant improvements when using the proposed schemes.

Finally, tag collisions are another concern in EPCc1g2 systems. It is investigated whether the use FEC can increase the frequency of captures: i.e., recoverable tag collisions. To this end, a capture model is proposed, and the relations for some important channel models are derived. This capture model is then used to evaluate the throughput without and with FEC coding. An increased capture probability can be demonstrated when using FEC, however, the improvement of the system's throughput is not very pronounced, when only regarding the influence of captures.

Zusammenfassung

Die vorliegende Arbeit untersucht die Eignung von Vorwärtsfehlerkorrekturverfahren (FEC) zur Verbesserung der Kommunikation zwischen Transpondern und Lesegerät in einem long-range Radio Frequency Identification (RFID) System. Der Fokus liegt dabei auf den EPCglocal Class-1 Generation-2 (EPCc1g2) Systemen.

Das Lesegerät in einem long-range RFID System muss in der Lage sein eine große Anzahl markierter (getaggter) Gegenstände in einer limitierten Zeit zuverlässig zu identifizieren. Die in diesen Systemen vorherrschenden Ausbreitungsbedingungen, stark gestörte Übertragungskanäle und die geringe Leistung der empfangenen Transpondersignale stehen dieser Aufgabe entgegen. Besonders wenn die Kommunikationsentfernungen, wie bei semi-passiven Transpondern, groß sind, ist der Signal-Rauschabstand der Tagsignale niedrig und die Fehlerwahrscheinlichkeit hoch. Obwohl diese Probleme hinreichend bekannt sind, gibt es in den zurzeit standardisierten Protokollen nur sehr einfache Fehlerschutzmechanismen. FEC wird in der Praxis noch nicht eingesetzt und ihr Einsatz in RFID Systemen ist noch nicht ausreichend erforscht.

In dieser Arbeit wird zunächst ein Modell beschrieben, welches den Einfluss von Übertragungsfehlern auf die Identifikationszeit der Transponder beschreibt. Anhand dieses Modells wird der Einsatz von FEC evaluiert und es werden die Auswirkungen auf die maximale Lesereichweite des Systems aufgezeigt. Für einen ausgewählten Code konnte gezeigt werden, dass sich so die Reichweite des System um 4 m erhöhen lässt, ohne dass dabei die Identifikationszeit steigt. Weiter wird ein iterativer Decoder vorgestellt der dazu verwendet werden kann, den verketteten Code zu decodieren, welcher sich aus dem vorgeschlagenen Faltungscode und den standardisierten EPCc1g2 Leitungscodes ergibt. Dabei konnte in Simulationen eine Verbesserung des Codiergewinns um bis zu 2.8 dB, beim Vergleich des iterativen mit einem nicht-iterativen Decoder, erzielt werden. Zudem wird das Konvergenzverhalten des Decoders untersucht.

Da sich weltweit bereits viele RFID Systeme im Einsatz befinden, sollten mögliche Erweiterungen stets mit Hinblick auf deren Kompatibilität zu den heutigen Standards entworfen werden. Es wird in dieser Arbeit gezeigt, wie sich die bereits in Gebrauch befindlichen EPCc1g2 Systeme mit Combining-Verfahren erweitern lassen, um die Anzahl von Wiederholungen gestörter Pakete zu minimieren. Weiter wird aufgezeigt, wie sich der existierende Standard um Hybrid Automatic Repeat Request (HARQ) Funktionalität erweitern lässt, ohne die Kompatibilität zu älteren Standardsystemen einzubüßen. In Simulationen konnte für die vorgeschlagenen Erweiterungen eine signifikante Reduktion der Identifikationszeit bzw. eine Erhöhug des Datendurchsatzes demonstriert werden.

Ein letzter Aspekt sind die in EPCc1g2 unvermeidlichen Kollisionen von Transponderantworten. Es wird untersucht, ob FEC dazu beitragen kann die Häufigkeit von *Captures*, d. h. die spontane Auflösung von Kollisionen, zu erhöhen. Dazu wurde ein Capture-Modell entwickelt und die Zusammenhänge für einige wichtige Kanalmodelle hergeleitet. Anhand dieses Modells wird schließlich untersucht, ob der Einsatz von FEC den Durchsatz des Systems steigern kann. Dabei ist beim Einsatz von FEC zwar eine Erhöhung der Capturewahrscheinlichkeit nachweisbar, jedoch ergibt sich daraus nur eine geringe Verbesserung des Durchsatzes im System.

Contents

List of Figures

Acronyms

ALOHA A statistical multiplexing and media access control scheme developed from 1968 on at the University of Hawaii.

APP A Posteriori Probability.

ARQ Automatic Repeat Request.

ASK Amplitude Shift Keying.

auto-id Automatic Identification.

AWGN Additive White Gaussian Noise.

BAP Battery-assisted passive tags.

BCH Bose Chaudhuri Hocquenghem (code), named after the inventors.

BER Bit-Error Rate.

BNetzA Bundesnetzagentur.

BPSK Binary Phase Shift Keying.

BSC Binary Symmetric Channel.

CC Chase Combining.

CDF Cumulative Distribution Function.

CRC Cyclic Redundancy Check.

CW Continuous Wave, referrers to an unmodulated carrier signal.

DBC Differential Phase Shift Keying.

DPSK Differential Phase Shift Keying.

DTMC Discrete Time Markov Chain.

EEPROM Electrically Erasable Programmable Read-Only Memory.

EPCc1g2 EPCglocal Class-1 Generation-2.

ERP Effective Radiated Power.

EXIT chart Extrinsic Information Transfer chart.

FCC Federal Communications Commission.

FEC Forward Error Correction.

FM Frequency Modulation.

FSA Framed Slotted ALOHA.

FSM Finite-State Machine.

HARQ Hybrid Automatic Repeat Request.

HF High Frequncy, frequency range from 3 MHz to 30 MHz.

IC Integrated Circuit.

IR Incremental Redundancy.

ISO International Organization for Standardization.

LBT Listen Before Talk.

LF Low Frequncy, frequency range from 30 kHz to 300 kHz.

LLR Log-Likelihood Ratio.

MAC Media Access Control.

MIMO Multiple input, multiple output.

MLSE Maximum likelihood sequence estimation.

MR combining Maximum Ratio combining.

N-LOS non line-of-sight.

OFD Optimum Free Distance (code). A convolutional code with the maximum free disctance for a given constraint length.

PDF Probability Density Function.

PIE Pulse-Interval Encoding.

PMF Probability Mass Function.

PSK Phase Shift Keying.

R-to-T reader to tag.

RADAR Radio detection and ranging.

RCPC Rate Compatible Punctured Convolutional.

RCS Radar Cross-Section.

RF Radio Frequency.

RFID Radio Frequency Identification.

RSC Recursive Systematic Convolutional.

SAW Surface Acoustic Wave.

SISO Soft Input Soft Output.

SNR Signal-to-Noise Ratio.

T-to-R tag to reader.

UHF Ultra High Frequncy, frequency range from 300 MHz to 3000 MHz.

VLSI Very-Lage-Scale Integration.

WGN White Gaussian Noise.

Nomenclature

λ	Wavelength.
ν	Memory of a convolutional coder.
$P_{K,i}(m)$	Probability of finding m slots occupied by one tag.
$P_{K,i}(m, R)$	Probability of finding m slots occupied by R tags.
ρ	Complex reflection coefficient.
σ	Radar cross-section of the tag.
a	Maximum number of retransmissions.
A_r	Random variable describing the real part of the channel coefficient H_r.
d	The distance between a tag and the reader.
d_f	Free distance of a convolutional code.
d_h	Hamming distance of a block code.
$E[\cdot]$	Expected value of a ramdom variable
E_b	Bit energy.
E_s	Symbol energy.
$F_A(a)$	Cumulative distribution function of the random variable A.
$f_A(a)$	Probability density function of the random variable A.
G_r, G_t	Gain of the reader antenna and the tag antenna.
$\mathbf{H}$	Random vector of independent channel coefficients $(H_1, \ldots, H_r)$.

h_f, h_b	Complex channel coefficient of the backward and forward link in a backscatter system
H_r	Random variable describing the channel coefficient of the r^{th} tag.
K	Number of slots in an inventory round.
K_α	Modified Bessel function of the second kind
L_p, l_p	Path-loss and path-loss in dB
L_s, l_s	Free-space path loss and Free-space path loss in dB
N	Size of the tag population.
N_0	Noise (power) spectral density.
$\mathbf{P}$	Transition matrix of a Discrete Time Markov Chain
p_b	Bit-error probability.
$P_c(R)$	Capture probability in a slot with R tags
p_f	Frame-error probability.
$p_{i,j}$	Transition probabilities from state s_i to s_j of a discrete-time Markov chain.
p_s	Slot-error probability: The probability that any of the frames in a slot of the inventory round contains errors.
p_{sym}	Symbol-error probability.
P_{th}, p_{th}	Turn-on or threshold power of a tag in W or dBm.
R	Occupancy number of a slot in an inventory round.
$\mathcal{S}$	Set of states of a Discrete Time Markov Chain
t	Error-correction capability of a block code.
T, T_c	Frame-local throughput without and with consideration of the capture effect.

1 Introduction

1.1 Motivation

The term Radio Frequency Identification (RFID) describes a wide range of technologies, systems and devices that exist in the context of the Automatic Identification (auto-id) world. The common task in any such system is the identification of a marked object with the help of a Radio Frequency (RF) communication link. The complexity of the system and amount of data transported from the so-called transponder or tag to the reader varies greatly, from simple one-bit theft-protection transponders to sophisticated sensor tags that allow the wireless collection of sensor data related to the marked object.

The technological foundations of RFID date back to the 1940s. Stockman is usually credited as one of the first to investigate the possibility of transferring information by changing the reflective properties of a target in a Radio detection and ranging (RADAR) system [Sto48], thus enabling the target to communicate with the RADAR receiver. In years to come, this lead to the development of long-range transponder systems for the use aboard aircrafts for the "identification, friend, or foe" [Lan05]. Important progress was made in the 60s when companies brought first simple anti-theft systems to the market. Subsequent developments and the advent of the microchip era, made for decreasing transponder sizes and eventually the wide variety of systems of today.

Nowadays RFID technologies can be found in contact-less smart cards, animal transponders, and ultralight and flexible smart tags for item identification in the supply chain. The most widely used system, when it comes to logistic applications, is the EPCglocal Class-1 Generation-2 (EPCc1g2) standard, with its most recent release [EPC13]. The International Organization for Standardization (ISO) has specified a compatible, mostly identical system based on the EPCc1g2 specification as ISO/IEC 18000-63:2013 in [ISO13].

Especially in logistics applications, readers have to identify a great amount of items in a short period of time. At the same time the propagation environment in places like warehouses is far from benign. Difficult channel conditions, noise and interference lead to an increased probability of frame errors, which in turn limit the throughput of the system. Current standards for RFID systems, however, do specify only basic mechanisms for error-control. The protocol elements in EPCc1g2 systems, for example, only include Cyclic Redundancy Check (CRC) checksums in some of the frame types, and the standard does not make any statements on how to recover from transmission errors altogether. No standard for long-range systems at all so far proposes the use of channel codes for Forward Error Correction (FEC) to mitigate the effects of transmission errors.

State of the art channel codes allow the reliable exchange of information close to the capacity limit of a channel [BG96, CFRU01] and FEC is ubiquitously used in modern wireless communication systems. Nevertheless, few work has been published on their use in an RFID systems.

In this work, the use of different error-control schemes in the scope of long-range RFID systems is investigated. While many of the channel codes and protocols are well established; the integration into an existing system poses a number of new questions, especially in RFID systems. RFID systems are fairly limited in complexity of their communication protocols for a simple reason: Tags are present in large numbers and should thus be small, cheap and must not require much power to operate. So while the tags cannot support the complex operations that are required to *decode* a channel code, it is possible to implement channel coder in a tag [JC09] or simply store the tag data pre-coded. The RFID readers on the other hand posses far better processing capabilities and are thus suited carry out the channel decoding. Also, readers are more readily modified than tags, because often the signal processing is already done in software.

1.2 Organization

The work is organized as follows. The next chapter gives a short overview of the research concerned with long-range RFID systems, and summarizes approaches and results of the publications that are relevant for this work.

Chapter 3 introduces the basic components of a long-range RFID system. The limitations of systems due to the power constraints of the tags and the readers are discussed, and an example for estimation of the read range of a system is given, based on a common propagation model. Furthermore, all channel models that are used in the work are introduced briefly. The last sections of the chapter offer an in-depth description of the EPCc1g2 protocol and procedures. The achievable throughput of the system and the error-behavior is discussed, and a Markov model of the tag identification process is presented [Vog02b, BDVA11], which is extended to consider transmission errors. The extended model was previously published in [Sch12].

Chapter 4 discusses the addition of FEC codes to improve an EPCc1g2 system. First, the Markov model from Chapter 3 is applied to determine the improvement that the use of an FEC code on the tag to reader (T-to-R) link has on the duration of the inventory process. Parts of this analysis were published in [Sch12], as well. Furthermore, the properties of the standard EPCc1g2 baseband (line) codes are discussed. Finally, an iterative decoder is presented, which is designed to decode the concatenation of a convolutional code and the standard baseband codes. It is shown that the iterative decoder can achieve superior decoding results, compared to a setup with a separate decoder for the baseband and the convolutional code. The convergence behavior of the iterative decoder is then analyzed using the Extrinsic Information Transfer chart (EXIT chart) of the decoder [Sch13].

Chapter 5 investigates advanced retransmission strategies for the EPCc1g2 system, and it is shown how combining can be used in the reader to improve the performance of the standard EPCc1g2 protocol. An extension is then proposed that adds Hybrid Automatic Repeat Request (HARQ) functionality to the system; Rate Compatible Punctured Convolutional (RCPC) codes are proposed as the class of underlying channel codes. The different retransmission mechanism are finally compared by monte carlo simulation. Originally the results presented in Chapter 5 were published in [SR13]. Another work proposes a very similar modification: [BR13] was published when [SR13] was accepted but not yet published, and also proposes the use of HARQ with RCPC codes in an EPCc1g2 system (Later [Boy14] acknowledges the existence of [SR13] and its independent development). The coding schemes of both works are compared in the last part

of Chapter 5, and an improved choice of codes is presented, which is superior to both previous schemes.

Chapter 6 presents a bit-error based capture model. With this model the capture probability for the collision of R tags replies can be calculated, when a certain number t of bit-errors are tolerable in the tag replies. Different channel models are supported by the proposed model and the specific relations for the Additive White Gaussian Noise (AWGN) channel, the Rayleigh channel and the dyadic backscatter channel are given in this work. The capture model is lastly used to evaluate the impact of channel coding on the capture probability, by considering the example of a Bose-Chaudhuri-Hocquenghem (BCH) code.

2 Related Work

2.1 Performance of RFID systems

A major goal of this work is to show that the performance of an RFID system can be improved by using advanced error-control methods in the system. This poses the question: which are the common performance indicators for RFID systems and what are the current approaches to improve system performance? [Vog02b, NR06, RD06, BW08] consider the read or identification rate as one of the most important characteristics of a system, regardless of the frequency and classification of the system. It is the most relevant performance indicator from the applications standpoint, and characterizes the overall system consisting of the reader and multiple tags. The read or identification rate, is closely related to the concept of the throughput of a communication system. Works concerned with optimizing the tag estimation process (e.g. [Vog02a, Vog02b, Flo07, BDVA11]) often give the frame-local throughput of the system, which relates only to the number of attempts and the number of collisions. However, in these works transmission errors are usually neglected for simplicity. This simplification cannot be made when the impact of error control and FEC is studied in this work.

2.2 Error control and RFID

It is well established that FEC codes can drastically improve the reliability of the communication in a noisy or otherwise impaired channel. Terrestrial wireless communication system deal with particularly bad conditions, compared to satellite or wire channels. In passive and semi-passive RFID systems the situation is aggravated because backscatter modulation is used, which leads to a steeper slope of the path loss compared to other wireless systems. These basic properties make it particularly hard to design and deploy efficient backscatter

RFID systems. Using FEC codes to overcome some challenges in these type of systems has been proposed by various works in the past.

In 2006, [Joe06] proposed to use BCH codes and a HARQ scheme in an RFID system to improve the low efficiency of the existing Automatic Repeat Request (ARQ) based protocols. Joe, however, does not elaborate the specific relevance of his proposed system to an RFID application. It is not clear in the paper which part of the communication is considered – presumably the T-to-R communication – nor does he address the issue of deploying coders or decoder in the resource limited tags. The author supposes further, that energy efficiency is a relevant factor, which seem counterintuitive at least for passive tags that are fully powered from the reader-generated field. Disregarding the shortcomings of the paper, using HARQ in an RFID system is a desirable approach that is later explored in [BR13] as well as in Chapter 5 of this work.

[JKC08] first proposed the use of convolutional codes to add FEC capabilities to the T-to-R link in an EPCc1g2 system. The authors carefully explain how such an extension could be integrated into the existing standard and analyze its merits, mainly by demonstrating the improvement of the identification time of a tag population. Later the same group describes a proof-of-concept for the Very-Lage-Scale Integration (VLSI) implementation of an FEC enabled tag [JC09]. The authors demonstrate that the addition of the coder logic in the tag only increases the number of gates and the chip area by approximately 5.4 %. They conclude that thus adding FEC capabilities to the tag is a viable possibility. Another, though not a very comprehensive work, that advocates the use of FEC is [DHK+11]. In the paper the author proposes to use Reed-Solomon codes as an improvement for Surface Acoustic Wave (SAW) tags. However, the authors do not offer any quantitative, nor qualitative statement on the expected performance improvements.

Though often overlooked, many protocols already have a built-in coding scheme that can potentially be exploited for FEC. In the case of EPCc1g2 systems many protocol elements carry a CRC checksum. Morelos-Zaragoza point out in his works [MZ11, MZ12] that already in standard systems single-bit errors can be corrected. The author proposes to use syndrome decoding for the CRC-16 in a standard EPCc1g2 reply. Syndrome decoding with a look-up table is, however, only feasible in this case, because the original code – an expurgated variant of the (32767,32752) hamming code – is severely shortened in practical

usage. The author shows by way of simulation that this results in a coding gain of approximately 1.25 dB in the AWGN channel. In a consecutive publication the author extends the original idea and describes a decoding procedure that corrects most (79 %) of the two-bit errors. This yields a slightly increased ($\approx$ 1.5 dB) coding gain. An issue that is not mentioned by Morelos-Zaragoza is that the suggested error-correction scheme effectively disables the otherwise outstanding error-detection capabilities of the CRC code, leaving the reader unable to detect remaining bit errors.

Additional work on the use of FEC in EPCc1g2 systems has been published by Smietanka et. al. in the recent years [SG12, SGG13, MSG14]. The authors suggest the use of BCH codes instead of the CRC in the EPCc1g2 protocol elements. For up to k=239 bits a standard frame could then use a $(255, 239)$ code with a hamming distance of $d_h = 5$ and an error-correction capability of $t = 2$, instead of only $t = 1$ for the CRC-16. Again, simply correcting all frames with an algebraic decoder, deprives the decoder of the capability to detect decoding errors. Smietanka proposes to overcome this by correlating the received codeword with a re-encoded version of the recovered data [SGG13]. By incorporating a chase decoder into the proposed system [MSG14] the authors can improve the system further. Unfortunately, in all of the publications, the proposed scheme is only evaluated in the terms of the Bit-Error Rate (BER) and the achieved coding gain. The authors never relate this to the throughput of the range or reliability of the system. Furthermore, only the transmission over a Rayleigh channel (or Dyadic Backscatter Channel (DBC) in [SGG13]) with uncorrelated fading amplitudes at rates *faster* than the data rate (independent fades at the sub-carrier rate) is considered. This assumption my not be sustainable at the relatively low bit rates and low mobility of the tags in an RFID system.

2.3 Open tasks

A number of works that propose FEC capabilities to long-range RFID have been published in the past decade. All of these works agree that the system will benefit from this extension but many leave issues unaddressed.

Evaluation of the improvements: The existing works advocate the use of FEC and argue that the system will benefit from it. E.g. [SG12, SGG13, MSG14, MZ11, MZ12] successfully demonstrated an improvement in the BER when FEC

is used. While these results are promising, they do not show how the performance at application level is affected, because neither the improvement of the read rates nor the throughput are evaluated. Only [JKC08, JC09] provide an estimation of the overall improvement of the inventory duration, which ultimately determines the performance with regard to the application.

Consideration of standard compatibility: The EPCc1g2 is the most widespread standard for long-range RFID up to this point. While an extension can probably be integrated fairly easily into the readers [MZ11, MZ12], a modification or replacement of all tags is undesirable. Any proposed extension to the standard should be evaluated having this in mind. Again, [JKC08, JC09] are the only ones considering the issue of (backward) compatibility in their work.

Having in mind these two major points, this work aims to propose mechanisms that will improve the system performance of a (EPCc1g2) long-range RFID system. Where possible these mechanisms shall be standard compatible or at least backward compatibility shall be preserved. The merits of the proposed changes shall be evaluated with respect to the performance indicators that are critical to application: the operational range of the system and the tag identification rate.

3 Basic technologies

3.1 System architecture

The term Radio Frequency Identification (RFID) is used to describe a wide variety of wireless communication systems. A common characteristic for all of these system is their use in an Automatic Identification (auto-id) scenario. The initiating instance of the communication is called *reader* or *interrogator*. The reader's task is to identify a number of *tags* (also: *transponders*) in its vicinity, by establishing a communication link to each tag and reading its unique identifier. The systems can be categorized based on the use case and the properties of the system components. International standards have been created for some of the most widespread RFID applications. The basic classifications and the most relevant standards are presented in the following two sections. The operational range of a single RFID reader is, as in any wireless communication system, limited, because the transmitted signal is attenuated along the propagation path. Section 3.4 summarizes the limiting factors and introduces the propagation models that are relevant in the latter parts of this work. Aside from the large scale-changes of the propagation environment, the system performance is also limited by small-scale fading and noise. These random variations are described by channel models. Section 3.5 summarizes the most important, general channel models and a fading model that was specifically proposed to model backscatter systems.

Besides the limitation due path-loss, fading, and noise, RFID systems are also subjected to the multiple access problem. In most scenarios there are multiple tags in the range of one reader and possibly multiple readers at one site. Simultaneous access of several contending stations to the shared channel then leads to collisions. The different types of possible collisions and the related media access protocols are presented in Section 3.6.

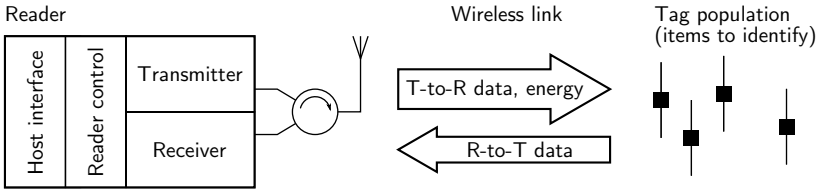


Figure 3.1: Model of the reader-tag communication system.

3.2 RFID Technologies and standards

3.2.1 Technologies and frequency bands

The term RFID refers to a wide range of wireless communication systems which differ in the designated operational range, the method of powering the tags and their frequency range in which the system operates. A rough classification according to the range of the system is given by [Fin08]

Close coupling systems operate at reading distances of a few centimeters and rely on inductive coupling to exchange information, as well as providing energy to the tags. *Remote coupling systems* also use inductive coupling, but provide reading distances from one up to a few meters. These kinds of systems are used for many applications; including the tagging of animals, entry control, wireless chip cards, etc. *Long range systems* do not require an immediate coupling between the reader and the tag. The tag is situated in the far field region of the reader antenna. Reading distances thus only depend on the field strength that is required to power the tag, and on the quality of the communication link. Thus operational ranges of a few meters to well over ten meters are possible, depending on the specific system, the propagation environment and required reliability. Long range systems are preferably used in logistics scenarios, e.g., automatic inventory, tracking of goods on a container or item level. However, the technologies that are developed for long range RFID systems, are also used in other systems, such as passive wireless sensor networks.

The classification according to the operational distances of the systems is closely related to the frequency range that is being used, and the coupling methods. *Low Frequncy (LF)* systems are usually low-cost, close coupling systems. Reader and tag are connected by magnetically coupled inductors (coils),

which serve for the communications as well as for providing power to the tag. Due to the magnetic coupling, a change in the impedance connected to the inductor in the tag, has an effect on the current in the reader's inductor. This *load modulation* is used to implement the tag to reader (T-to-R) communication in LF systems. The commonly used bands for LF systems are at 125 kHz and 135 kHz.

High Frequncy (HF) systems also mostly rely on inductive coupling for the reader to tag (R-to-T) and T-to-R communication, as well as the tag power supply. These systems are usually remote coupling systems with reading distances of up to a few meters. HF systems also use load modulation for the T-to-R communication. The most prominent band in use for HF systems is at 13.56 MHz.

Systems that operate in the *Ultra High Frequency (UHF)* regime of the spectrum can offer higher reading distances than the ones in the LF and HF bands, because they do not require a direct coupling between tag and reader. Instead UHF systems use the method of *modulated backscatter* for the T-to-R communication. [Sto48] is usually credited as the first scientific work describing this method of communication. In his work, Stockman describes the general idea of modulating the reflective properties of a radar target to transmit information from the target back to the radar station. He demonstrates the feasibility of implementing both amplitude and phase modulation by changing the reflective properties of the target. In RFID systems the tag antenna is equivalent to the radar target. The tag can perform a simple binary Amplitude Shift Keying (ASK) or Binary Phase Shift Keying (BPSK) modulation by switching the complex load connected to the tag antenna. Even though the communication method in UHF systems enables longer reading distances, passive tags still need to be powered from the electromagnetic field created by the reader, so the systems usually offer reading distances of about 10 m. There are several bands in the UHF region that can be used for RFID applications: 433.050 MHz to 434.790 MHz [ETS12] the 2.446 MHz to 2.454 MHz [ETS10] and, most important for the EPCglocal Class-1 Generation-2 (EPCc1g2) systems, the bands at 865 MHz to 865 MHz [ETS11] and 915 MHz to 921 MHz [ETS14].

3.2.2 Tag types

Another way of classifying RFID systems is given e.g. in [EPC07] and [Fin08]. RFID systems differ in the ways of providing energy to the tags, and in the way the T-to-R communication link is implemented.

Passive tags are tags that do not have a power source of their own. They are also not capable of generating a Radio Frequency (RF) field and thus use methods like load modulation of modulated backscatter on the T-to-R link. Passive tags can only operate when powered by the reader. [EPC07] defines the classes 1 and 2, which are passive tags.

Battery-assisted passive tags (BAP) include a power source that allows them to operate independent of a reader generated field, e.g. to record sensor values for an extended time span. Just like passive tags, these tags are not able to generate an RF field and use the same mechanisms as class 1 tags to communicate with the reader. However, since they have a power source of their own, they require less received power when communicating with the reader. The EPC classification for battery-assisted passive tags is class 3.

Active tags include a power source and a transmitter, which allows them to generate an RF field for communication, independent of a reader. This type of tags form the EPC class 4. Active tags are out of scope of this work, because of their significantly different characteristics compared to the classes 1-3.

3.2.3 Standardization efforts

Standardization is an important aspect when it comes to facilitating the adoption of a technology. Although first proprietary RFID systems were patented and developed in the 1970s [CP73, Wal73, Wal74], it took some time before international standards were developed.

[MIT02] and [MIT03] were first standard proposals for UHF systems created by the Auto-ID center at the Massachusetts Institute of Technology. Both documents specify simple systems using passive tags; neither became an international standard, however. The Auto-ID center's task were later taken over

by the EPCglobal inc. which peruses the international standardization of systems and protocols, and the Auto-ID labs that continue the research on auto-id systems.

Today, a wide number of standards exists, covering all aspects of an RFID system. This includes standards for data storage, information system, mechanical properties etc. The most important standards for this work are the standards that define the properties and protocols of the air interface of the long range systems that operate in the UHF range. The International Organization for Standardization (ISO) released the first standards for systems with application to item management in the ISO/IEC 18000 series in 2004. Today, the ISO 18000 series includes air interface definitions for frequency bands at 135 kHz, 13.56 MHz, 433 MHz, 860 – 960 MHz and 2.45 GHz. Most relevant for this work are the standards ISO 18000-61 – ISO 18000-64, which describe four different air interface definition at 860 – 960 MHz, called types A-D. ISO 18000-63 (type C) [ISO13] is mostly equal to the *EPC^{TM} Radio-Frequency Identity Protocols Generation-2 UHF RFID Specification for RFID Air Interface Protocol for Communications at 860 MHz — 960 MHz* issued by the EPCglobal [EPC08, EPC13], which is commonly referred to as EPCc1g2. This type of system seems to make up for most of the currently available systems in the UHF range. In the early versions the standards EPCc1g2 (before 2013) and ISO 18000-6 (before 2010) only described systems with passive tags. Later versions also include additions for BAP tags. Neither the EPC nor the ISO standard include active tags.

3.3 EPCglobal system components

3.3.1 Tags

Most of the tag Integrated Circuits (ICs) for the UHF range that are available from the semiconductor manufacturers today, are passive class 1 tags according to the EPCc1g2 standard. The ICs are fixed on a flexible carrier together with an antenna to form a so called inlay, which can then be integrated into flexible paper tags. Figure 3.2 shows a scan of a UHF inlay. All active components of the tag are contained on the chip that is shown in the center of the image.

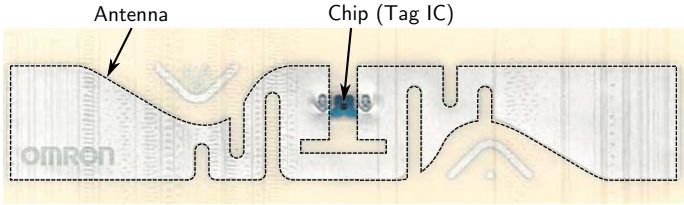


Figure 3.2: Scan of a passive UHF inlay manufactured by Omron (to scale).

Figure 3.3 shows a simplified sketch of the RF front end of a passive RFID tag. There are three main components [Dob08]:

Envelope detector: Passive and BAP tags do not contain a local oscillator or mixer circuits, but use a simple envelope detector to demodulate the R-to-T signal instead. This keeps the power consumption and the complexity of the tag low, but on the downside, disables the tag from obtaining any information that might be contained in the carrier phase of the R-to-T signal.

Modulator: The modulator for the T-to-R link is essentially a switching element, e.g. a transistor that is used to change the impedance connected to the tag antenna. By varying the resistance, the tag can cause an ASK modulation of the backscattered signal. Tags generally implement more intricate modulation circuits than shown in the figure, to realise the modulation while ensuring that the tag receives enough power to supply its circuits. By changing the load reactance connected to the tag antenna, a tag can change the phase of the backscattered signal and thus use Phase Shift Keying (PSK) modulation [Dob08, 5.4]. [BR12] even proposes to implement higher order modulation schemes such as QAM in the tag.

Power supply: A passive tag utilizes the power of the received reader signal to power its circuits. Since the voltage at the antenna port is usually too low to supply an IC directly, multistage voltage multipliers are used to generate the necessary level.

The minimum field strength required to operate a passive tag depends mainly on the power consumption of the tag. It can be specified in terms of the *turn-on*

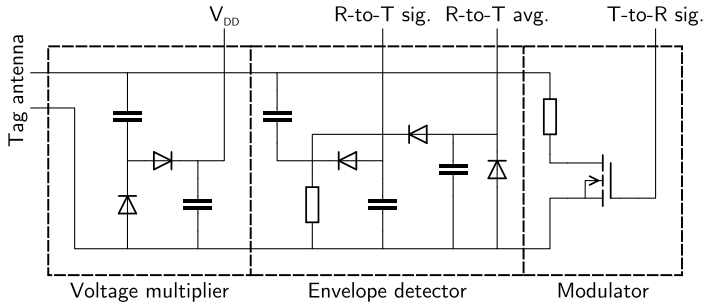


Figure 3.3: Simplified circuit of the RF front-end of a passive tag [Dob08, Fig. 5.6, 5.13].

or *threshold* power of the tag p_{th}. For BAP tags, the required power is lower than for passive tags, because a BAP tag does not rely on the reader signal for its power supply. Generally, the power requirements for passive tags are higher, when write operations have to be carried out by the tag, because the tag requires additional energy to erase and reprogram the Electrically Erasable Programmable Read-Only Memory (EEPROM).

Table 3.1: Tag sensitivity values according to manufacturer [NXP13, Imp12d, Imp14a, Imp14b, Int10].

Manufacturer	Product Series	Type	Read sensitivity p_{th}
NXP	UCODE 7	passive	$-18\,\mathrm{dBm}$[1]
”	”	BAP	$-27\,\mathrm{dBm}$[1]
Impinj	Monza 4	passive	$-17.4\,\mathrm{dBm} \ldots -19.9\,\mathrm{dBm}$[1]
”	Monza 5	passive	$-20\,\mathrm{dBm}$
”	Monza X	passive	$-17\,\mathrm{dBm} \ldots -19.5\,\mathrm{dBm}$[1]
Alien Technology	Higgs 4 SOT	passive	$-18.5\,\mathrm{dBm}$[1]
Intelleflex	XC3	BAP	$-40\,\mathrm{dBm}$[2]

[1] datasheet, [2] markting claim

Table 3.1 lists the read sensitivity for commercially available state of the art UHF tags. All of the passive tags have a read sensitivity of about $-17\,\mathrm{dBm}$ to $-20\,\mathrm{dBm}$, according to the claims of the manufacturers. BAP tags like the XC3 series by Intelleflex, which implements the BAP mode as described in ISO-

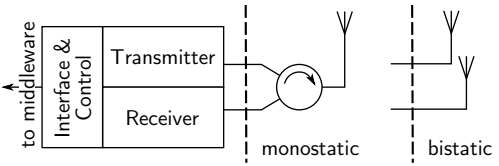


Figure 3.4: Mono- and bistatic reader configurations.

18000-63, can operate at much lower power levels, and allow reading distances in the order of 100 m, according to the claims of the manufacturer.

3.3.2 Reader

The reader or interrogator of an RFID system can be divided into the components depicted in Figure 3.4. The reader consists of a transmitter, a receiver and one or several antennas and the software which provides an interface to the middleware. The reader software and the middleware are not considered in this work.

Unlike in most other communication systems, the transmitter does not only encode and send the R-to-T data, but also provides the power for the tags. In consequence, the transmitter has to remain active, even when the reader is not sending data. The unmodulated (CW) signal which the reader transmits causes a self-interference at the reader receiver during the backscatter phase, which is aggravated by the fact that the T-to-R signals are essentially received on the same frequency as the transmitter carrier. The receiver has to be able to mitigate the effect of self-interference to achieve a good sensitivity for the reception of the R-to-T signals. Simultaneous transmission and reception on the same frequency also pose special requirements for the connection of the antennas at the reader. In a *bistatic* setup, the transmitter and the receiver use separate antennas. In a *monostatic* setup, only one antenna is used, and the reader connects to the antenna using either a *directional coupler* or a *circulator* to separate the transmit and receive branches.

The read range that a reader can achieve depends on the transmit power of the reader on the one hand, and the receiver sensitivity on the other hand. The maximum transmit power is usually limited by national or international regula-

tions, enforced by the regulatory organizations such as the Bundesnetzagentur (BNetzA) and the Federal Communications Commission (FCC). In Germany, the maximum transmit power for indoor RFID applications is limited to 2 W ($\approx$ 33 dBm) Effective Radiated Power (ERP) [Bun10] in the 865 MHz band. The reader sensitivity as specified by the manufacturer consists of two values: The Listen Before Talk (LBT) sensitivity and the sensitivity during backscatter. The former is relevant for the reader to be compliant with the regulatory requirements, but only the latter is meaningful for the read range. Impinj specifies the backscatter sensitivity of their reader chips to be in the order of -65 dBm down to -95 dBm [Imp12c, Imp12a, Imp12b], depending on the specific chip and the magnitude of the self-interference power.

3.4 Propagation models and Interrogation Range

The read range of an RFID system does not only depend on the sensitivity of both tags and the reader, but also on the spatial setup of the system. A prediction of the approximate read range of a system can be obtained based on the ratings of the components, using a suitable path loss model.

[NR06] briefly summarizes the most important parameters that limit the range of an RFID system and gives an estimation of the read range for systems with passive tags in a free-space propagation environment. Free-space propagation is present when there is a line-of-sight between the transmitter and the receiver and there are no additional propagation paths caused by: reflections, diffraction, etc. In this case, the path loss is described by the Friis equation [PS08]

$$P_{r,tag} = P_{t,reader} G_r G_t \frac{\lambda^2}{(4\pi d)^2} \tag{3.1}$$

where $P_{r,tag}$ is the received power at the tag and $P_{t,reader}$ is the power transmitted by the reader. G_r and G_t are the antenna gains of the reader and the tag antennas, respectively. The remaining factor is called free-space path loss for two stations separated by a distance of d.

$$L_s = \left(\frac{\lambda}{4\pi d}\right)^{-2} \tag{3.2}$$

(3.1) can be applied straightforward, to calculate the received power at the tag on the R-to-T link.

The received power on the T-to-R link in a backscatter system is described by the monostatic radar equation

$$P_{r,reader} = \frac{P_{t,reader} G_r^2 \lambda^2}{(4\pi)^3 d^4} \sigma \tag{3.3}$$

Here, σ is the Radar Cross-Section (RCS) which describes the fraction of the power of an incidental wave that is reflected by a (radar) target towards the origin, e.g. an RFID reader. The RCS of an antenna generally depends on its effective aperture, which in turn is related to the antenna gain G_t and the load connected to the antenna. In an RFID system the tag modulates the antenna load to transmit information on the T-to-R link. This generally means that the RCS changes between two states σ_1 and σ_2, when a binary transmission scheme is realised. [NRM07] elaborates on this relations and notes that only the change of the RCS is relevant to determine the power of the modulated signal received by the reader. The mean power $P_{r,reader} \propto \frac{\sigma_1+\sigma_2}{2}$ constitutes the DC-component of the signal and does not contribute to the transmission of the tag information. The differential backscattered power of the *modulated* tag signal can be described in terms of the differential RCS

$$\Delta\sigma = \frac{\lambda^2 G_t^2}{4\pi} |\rho_1 - \rho_2|^2 \tag{3.4}$$

where ρ_1 and ρ_2 are the reflection coefficients at the interface between tag antenna and input, for the two backscatter states. Using (3.3) and (3.4) the power of the modulated tag signal at the reader is given as [LGS09]

$$\begin{aligned} P_{mod,reader} &= P_{t,reader} G_r^2 G_t^2 \left(\frac{\lambda}{4\pi d}\right)^4 |\rho_1 - \rho_2|^2 \\ &= P_{t,reader} \frac{G_r^2 G_t^2}{L_s^2} |\rho_1 - \rho_2|^2 \end{aligned} \tag{3.5}$$

(3.5) is often also given in its logarithmic form, as

$$p_{mod,reader} = p_{t,reader} + 2g_r + 2g_t - 2l_s + 20 \log |\rho_1 - \rho_2| \tag{3.6}$$

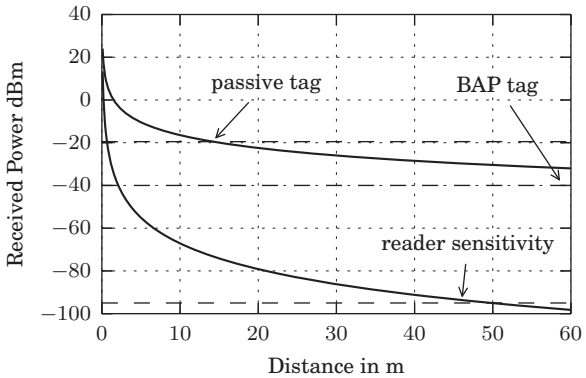


Figure 3.5: Received power at the tag (top) and the reader (bottom) in a backscatter RFID system with free-space propagation conditions.

An estimation of the maximal range of an RFID system can be calculated using (3.1) and (3.5). The system ceases to work reliably when the received power at the tag falls below the specified read sensitivity of the tag or when the power of the tag reply received by the reader falls below the reader sensitivity. Consider an example based on the values given in the previous subsections 3.3.1 and 3.3.2. The transmit power of the reader $P_{t,reader}$ was chosen to satisfy the limits imposed by the BNetzA, i.e., $P_{t,reader} \cdot G_r = 2\,\mathrm{W}$ (33 dBm ERP). The reader and tag antennas have a gain of 9 dBi and 1.72 dBi, which corresponds to the gain of a patch antenna at the reader and a dipole antenna at the tag, respectively. The system operates at a frequency of 866 MHz so $\lambda \approx 35\,\mathrm{cm}$. The differential reflection coefficient was chosen as $-10.2\,\mathrm{dB}$ according to the results presented in [NRM07]. Figure 3.5 shows the plots of (3.1) (solid line, top) and (3.5) (solid line, bottom) over the tag-reader-distance d for the selected parameters. Additionally, the dashed lines depict (from top to bottom) the read sensitivity of a passive tag (e.g. Impinj Monza 5), a BAP tag (Intelleflex XC3) and the sensitivity of the reader as specified by Impinj for the R2000 chip (Table 3.1, [Imp12b]).

In the example $P_{r,tag}$ the received power at the tag falls below the sensitivity threshold of a passive tag ($-20\,\mathrm{dBm}$) at a distance of 14 m. At this, point the

received power at the reader is well above the specified sensitivity of the reader. So, in the case that passive tags are used, the range of the system is limited to approximately 14 m, due to the tag's power requirements. If BAP tags are used in the system, the conditions change. A system with BAP tags can support considerably longer reading distances, as far as the R-to-T link is concerned. In the example, the received power at the tag has not yet fallen below the read sensitivity threshold of the BAP tag at 60 m. However, on the T-to-R link the received power at the reader falls below the reader sensitivity at about 46 m. The read distance of the system is then limited by the reader sensitivity, regardless of the sensitivity of the BAP tag. The distinction between systems with passive and BAP tags is important when the use of Forward Error Correction (FEC) on the T-to-R link is analyzed in the Chapters 4 and 5. Coding mechanisms can only be effective if the tags still receive enough power to operate. Thus, this techniques are most valuable in systems with BAP tags, where the T-to-R link is the limiting factor.

RFID systems are often deployed indoors, and even if a line of sight is present between the reader and the tags, the surroundings might be cluttered with obstacles that cause reflections and scattering. In this cases, it is clear that the assumption of free-space propagation becomes invalid, since multiple propagation paths are available between the reader and any tag. Propagation models for indoor channel have been investigated in the context of indoor mobile communications for some decades. [Has93] sums up many important properties and references. [LGS09, LGSV09, GD09] propose some known path loss models specifically for the use RFID applications.

For l_p the large scale path loss, the sources propose empirical models that are based on measurements. The most simple model for non-free-space propagation is the one-slope model. In this model, the parameter n can be adjusted to match the propagation environment.

$$l_p = -20 \log \frac{\lambda}{4\pi} + n10 \log d \tag{3.7}$$

For $n = 2$ (3.7) is the logarithmic version of the free-space path loss given in (3.2). For an indoor propagation environment the parameter n can vary between $n = 1.5 \ldots 4$ (see [Has93] and references). Values smaller than two, i.e. a path loss that is lower than in the free-space case, are reported when the environment

has a waveguide like structure, like hallways or tunnels. Larger values indicate a higher path loss, often due to destructive interference of multiple paths or objects obstructing the line of sight. E.g. [ETS07, eq. D.3] includes an indoor propagation model in which the path loss exponent is $n = 3.5$,

$$l_{ETSI} = 50.2 + 35 \log \frac{d}{10} + M_{\text{wall}} \tag{3.8}$$

which is applicable to distances $d = 10\,\text{m} \ldots 500\,\text{m}$. M_{wall} models the attenuation of the signal by walls in the propagation path. The report does not specify an estimate for M_{wall} though.

3.5 Channel models

3.5.1 Model of the discrete communication system

The communication channel in an RFID system can be described using the model of a discrete communication system. Figure 3.6 shows the structure of the T-to-R link of the system, that will be used throughout most of this work. The system has to be designed in a way that the source (tag) can transmit information to the sink (reader) via a noisy channel, while minimizing the probability of errors in the information that is passed to the sink. Channel models are used to quantify the error behavior in a discrete communication system. Such a model can be used to determine the performance and the limitations of the system, but unlike the deterministic path-loss models that were presented in the previous section, channel models are stochastic models. This means, channel models also give stochastic results, e.g. the error probability of the communication under a given set of circumstances. This section will introduce the most common channel models and their relevance for this work.

3.5.2 Basic channel models

The Binary Symmetric Channel (BSC) is a simple model of a discrete memoryless channel. In the model of the discrete communication system, the modulator takes a bit stream at its input and the demodulator output are bits as well. If

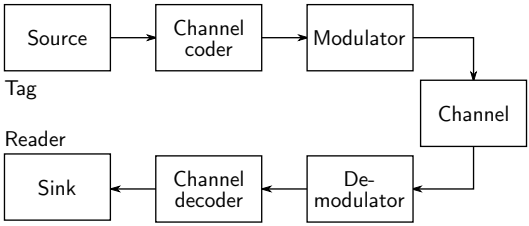


Figure 3.6: Model of a discrete communication system, e.g. the T-to-R link in an RFID system.

the channel and the modulation scheme are memoryless, then the entire branch of the system can characterized completely by the cross-over probabilities

$$p\,(0 \text{ received} \mid 1 \text{ sent}) \text{ und } p\,(1 \text{ received} \mid 0 \text{ sent})\,.$$

When the error probability is equal for both binary symbols, the channel is called symmetric with a bit-error probability of p_b. The BSC model is applicable

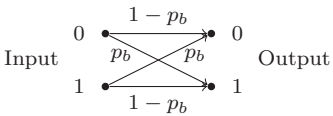


Figure 3.7: Binary symmetric channel with error probability p.

whenever discrete binary values are passed to the channel decoder. This is the case when either the demodulator can only provide binary decisions, or a *hard decision* channel decoder is used.

The most essential model for a discrete-input continuous-output channel is the time discrete Additive White Gaussian Noise (AWGN) channel model. Consider a modulator that maps the output of the channel coder to one symbol from a set of M discrete values. The sequence of symbols x_i is energy constrained and satisfies the condition $E\left[\|x_i\|^2\right] = E_s$, where the expected values E_s is called the symbol energy. In the channel, noise is added to the transmitted waveforms. The noise sequence n_i is a zero-mean complex Gaussian random process with variance σ^2, where $\sigma^2 = E\left[\|n_i\|^2\right]$ for a zero-mean process. The

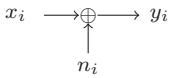


Figure 3.8: Time discrete AWGN channel.

time discrete AWGN channel relates to the time continuous AWGN channel in the way that the variance of the random process is related to the power spectral density N_0 of a time continuous process $\sigma^2 = N_0$ [PS08].

3.5.3 Channel models for mobile stations

While the simplicity of the AWGN channel model makes it of great value for the analysis of codes, modulation schemes, etc., it does not apply to many communication systems that are found in practice. In most mobile or indoor scenarios, there are multiple propagation paths, as pointed out in Section 3.4, and because of the mobility of the participants, the channel is time variant. The fading of the channel is then modeled by a complex random channel coefficient. Thus, also the Signal-to-Noise Ratio (SNR) is no longer constant, but becomes a random value.

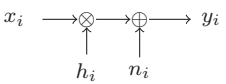


Figure 3.9: Time discrete fading channel.

Figure 3.9 shows a symbol rate model of a fading channel. The properties of the channel are determined by the probability distribution of the fading amplitudes and the average SNR. When there is a large number of independent propagation paths and no dominant (e.g. line of sight) component is present, the Rayleigh channel model is applicable. The distribution of the signal amplitudes in this model is Gaussian with zero mean, in both the real and the imaginary part. The magnitude of the channel coefficient $|h_i|$, and thus the signal amplitude $|y_i|$ is Rayleigh distributed [PS08]. If there is a multitude of propagation paths, but there is a line of sight component, or there are other

dominant propagation paths, then the signal magnitude is described by the Rice channel model.

There are further distinctions between slow and fast, and flat-frequency and frequency selective fading. In RFID systems with moderate transmission rates such as the EPCc1g2 system, slow and flat fading is predominantly applicable [FS09].

3.5.4 The dyadic backscatter channel

The fading mechanisms in backscatter systems differ from those in regular one way communication systems, due to the dependency of the backscattered signal from the incidental CW signal. Both the forward CW and the modulated backward signal are subject to fading processes. As depicted in Figure 3.10,

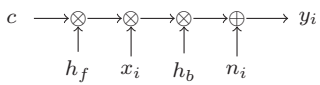


Figure 3.10: Dyadic backscatter channel.

the complex amplitude of the received signal is

$$y_i = c\,h_f\,x_i\,h_b + n_i = \underbrace{c\,h_f h_b}_{h} \cdot x_i + n_i \tag{3.9}$$

Since the carrier signal c is a constant, the amplitude distribution of y_i is determined by the product of the two random fading processes h_f and h_b and an additive noise term. Modeling the fading in backscatter systems by the product of two fading processes was proposed in [KIS03]. Later, Griffin coined the term *dyadic backscatter channel* [GD07, GD08]. It should be noted however, that this channel model is not only applicable to backscatter systems. In the context of Multiple input, multiple output (MIMO) systems, so-called keyhole or pinhole channels are known, which can be modelled the same way [CFV00], and single-relay amplify and forward systems show such a behaviour, as well.

The overall fading statistics in the dyadic backscatter channel depend on the propagation environment and the reader setup. Under non line-of-sight (N-LOS) conditions, both forward and backward link are subjected to Rayleigh

fading, hence the overall fading process is characterized by the product of two Rayleigh random variables. In the most general case, the two Rayleigh processes are correlated. For a monostatic reader setup, the forward and backward links are identical due to the reciprocity principle of electromagnetism. Thus, h_f and h_b are *fully correlated* in this case. For a bi- or multistatic reader setup the correlation of the channel coefficients depends on the spatial configuration of the reader antennas. If those are sufficiently separated, the forward and backward link can be regarded as independent. Below, the Probability Density Functions (PDFs) of the channel coefficients in the dyadic backscatter channel are given for the case of an $1 \times 1 \times 1$ (Single-input single-output) configuration. For the case of independent h_f and h_b [GD08, Table I] the PDF of $h = h_f h_b$ is

$$f(h) = \frac{4h}{\Omega^2} \mathrm{K}_0 \left(\frac{2h}{\Omega} \right) \tag{3.10}$$

and for the fully-correlated case [GD08, Table II] it is

$$f(h) = \frac{1}{\Omega} \exp \left(\frac{-h}{\Omega} \right). \tag{3.11}$$

Ω is the parameter of the parameter of the Rayleigh distribution and is assumed to be identical for the forward and the backward link. The complete PDFs of the dyadic backscatter channel are given in [GD08].

3.6 Tag arbitration and Media Access Control

3.6.1 Media Access Control for RFID

In an RFID system, the reader faces the challenge of identifying a tag population of an a priori unknown size. In the course of this process commonly called *tag identification*, *tag inventory* or *tag arbitration*, there is an unknown number of active tags in the coverage area of the reader. Controlling the access of each individual tag to the shared T-to-R channel requires a suitable Media Access Control (MAC) scheme. [ZY11] offers a comprehensive overview of the anti-collision schemes that are used in RFID systems today. There are two basic MAC and tag inventory schemes in RFID systems: *tree based* and *random access*.

Tree based methods organize the tag population in a binary tree based on the unique identifier or a temporary random identifier. The reader then traverses the tree, prompting only the tags in a given sub-tree to reply. The reader must be able to reliably distinguish between a single or multiple tag replies. If multiple tags reply, the reader divides the remaining selected tags according to the binary tree and proceeds until a leaf of the tree is reached, i.e., a single tag reply. The obsolete [MIT03] standard for EPCglobal Class-0 systems specifies a tree based approach for the tag arbitration. Of the UHF standards currently in force, only ISO 18000-62 (Type B) uses a tree based tag arbitration method. All others, including EPCc1g2, use a random based approach.

3.6.2 Random based tag arbitration

Random based tag arbitration methods are built on the idea of allocating the medium to the contenders in a communication system, based on a random process. The ALOHA protocol [Abr70] is generally considered one of the first examples of this approach. Pure ALOHA uses fixed packet length and allows all contending stations to transmit at any moment, which eventually leads to collisions between two or more packets. When the intervals between channel uses follow a negative exponential distribution, this leads to the well known maximum throughput of ALOHA of 0.184. In slotted ALOHA channel uses are synchronized, i.e., all contending stations start their transmission at the beginning of a definite time interval. For slotted ALOHA the throughput increases to 0.368 [Rob75]. Framed Slotted ALOHA (FSA) is a variant of slotted ALOHA [Sch83], in which the channel time is divided into frames with a fixed number of slots. Each participating station is only allowed to occupy one slot per frame with its transmissions. Figure 3.11 illustrates the principle of FSA. In the figure, the beginning of a frame is marked by SOF, which also marks the beginning of the first slot. The SOF command may contain additional information, e.g., the number of available slots in a frame. At the beginning of each frame, the contending stations independently choose a slot at random. Consecutive slots are marked by a delimiter(DEL) command.

EPCc1g2 uses a form of FSA for the tag arbitration. According to the EPCc1g2 standard, the tag identification is carried out in several so-called inventory rounds, which essentially correspond to FSA frames. The frame delimiters

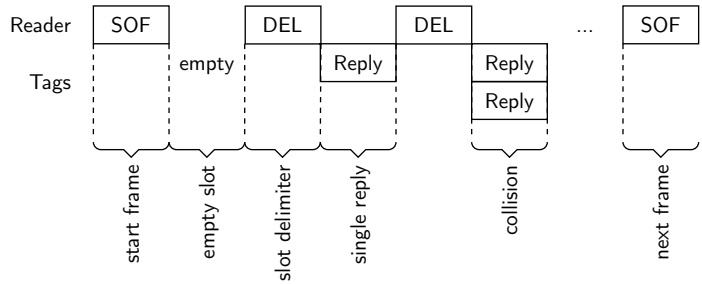


Figure 3.11: Framed slotted ALOHA.

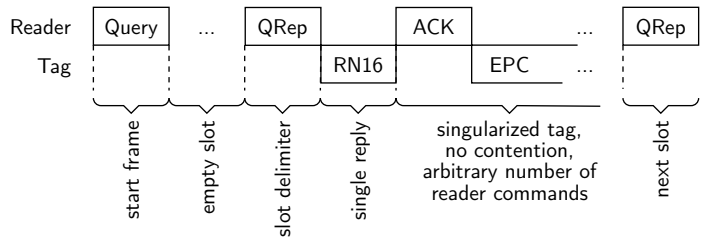


Figure 3.12: Framed slotted ALOHA in an EPCglobal system.

shown above in Figure 3.11 correspond to the reader commands: *Query* and *QueryAdjust*, which mark the beginning of a new round. *QueryRep* corresponds to the delimiter and marks the beginning of a new slot. Figure 3.12 shows the lightly modified scheme. Slots have a variable duration in the EPCc1g2 system, unlike in classical FSA systems, and the tags do not immediately respond with the tag data – their EPC identifier. Instead, they only transmit a short, temporary, random identifier called *RN16*. Only if the reader successfully receives the *RN16* of one tag, will it acknowledge the tag by issuing an *ACK* command. This exchange may fail, if a tag collision happens, i.e., several tags that reply in the same slot (not shown in Fig. 3.12). If there is no collision, the acknowledgement succeeds and the slot is successfully allocated to one specific tag. The tag then sends its full identifier (*EPC*). During the duration of the slot, the reader and *one* tag communicate in half-duplex mode. All remaining tags are silenced. The reader may send an arbitrary number of commands to the tag

before continuing the inventory process by issuing another *QueryRep*, *Query* or *QueryAdjust* command. So, the duration of the slots and the frame duration are variable in the EPCc1g2 system. The collision domain, however, is fixed and corresponds to the time interval that starts after the *QueryRep* command and has the duration of the *RN16* tag response.

Though the structure of the EPCc1g2 protocol differs slightly from the classic FSA, the throughput that can be achieved in the EPCc1g2 tag arbitration is modelled the same way in both cases. Consider a system where N is the number of tags in the coverage area of the reader and K is the number of slots in an inventory round. The random assignment of the tags to the slots corresponds to the random placement of N balls in K bins, which is known as the classical occupancy problem [Fel50]. Let R be the occupation number, i.e. the number of tag replies in a particular slot. The occupation numbers of each slot in the frame can be described by independent identically distributed random variables that follow a binomial distribution [Vog02b] with the well-know Probability Mass Function (PMF)

$$\Pr(R) = \binom{N}{R} \left(\frac{1}{K}\right)^R \left(1 - \frac{1}{K}\right)^{N-R} \tag{3.12}$$

Any tag chooses a particular slot independently of the other tags with equal probability $\frac{1}{R}$. So $\left(\frac{1}{K}\right)^R$ is the probability that R tags choose a certain slot, while $(1 - \frac{1}{K})^{N-R}$ is the probability that the remaining $N - R$ tags choose other slots. There are $\binom{N}{R}$ possibilities to choose R tags out of N, thus (3.12) expresses the probability that R arbitrary tags choose a particular slot. Since the slots are indistinguishable, (3.12) is valid for all K slots.

The occupation number is $R = 0$ for an empty slot. A slot with $R = 1$ is called a singleton slot. Any occupation number $R > 1$ constitutes a tag collision, and the slot is called a collision slot. The expected number of slots with an occupation number of R is given by [Vog02b]

$$S(R) = K \binom{N}{R} \left(\frac{1}{K}\right)^R \left(1 - \frac{1}{K}\right)^{N-R} \tag{3.13}$$

Clearly, collisions are not desirable, neither are empty slots, since they occupy the channel without contributing to the transfer of information. The system should be optimized, so the number of slots that carry meaningful information

is maximized, while at the same time minimizing the total duration of an FSA frame. It is generally sought to maximize the frame-local throughput T, which is defined as the ratio of the number of useful slots in a frame, and K the total number of slots. A slot is considered useful, when a tag is identified in the given slot. It is commonly assumed that only singleton slots contribute to the identification of tags and can thus be considered useful. If a system can, however, separate tag replies in a collision slot with $1 < R \leq M$, then any slot with occupation number $1 < R \leq M$ can be considered useful as well. E.g., [ALR10] proposes a MIMO RFID system that can separate the tag replies in a collision slot. The frame-local throughput is then

$$T = \sum_{R=1}^{M} \frac{S(R)}{K} = \sum_{R=1}^{M} \binom{N}{R} \left(\frac{1}{K}\right)^{R} \left(1 - \frac{1}{K}\right)^{N-R} \tag{3.14}$$

according to [ALR10], where M marks the number of colliding tag replies that can be tolerated, while still being able to recover the tag data.

3.6.3 Capture Effect

In a real system, the collision of two tag replies does not generally mean that the reader will receive erroneous data. If one tag signal is dominant, a reader may receive the correct reply, even without taking measures to resolve the collision. This case is called a *capture*. Captures have been identified as an important aspect of ALOHA systems from the beginning and have been addressed early on in the literature. [Rob75] suggests that a capture may occur when the signal power of one of the stations involved in the collision exceeds the signal power of the other transmission by a certain factor. This relation can be expressed by the capture ratio

$$CR = 10 \log \frac{P_d}{P_i} \tag{3.15}$$

which lies in the order of 1.5 dB . . . 3 dB for Frequency Modulation (FM) receivers, according to Roberts. In (3.15), P_d is the power of the dominant signal and P_i the power of the interference. Either the power of the second strongest station in the slot, or the sum of the power of all interfering stations in the current slot is usually chosen for P_i [LL90].

The capture effect affects the frame-local throughput in an FSA system. Captures happen with a certain *capture probability* that depends of various parameters like: number of collisions R, spatial distribution of the contenders or power ratio, properties of the transmission channel, etc. Let $P_c(R)$ denote the probability that the information transmitted by any of the R tags in a collision slot can be recovered without errors. $P_c(R)$ must be independent of the slot number, since all slots are indistinguishable and occupied with the same probability. With (3.12), the conditional probability that a slot with R tags is useful, i.e., leads to the identification of a tag, can be written as

$$P\left(\text{slot useful} \mid R\right) = P_c\left(R\right) \Pr\left(R\right) \tag{3.16}$$

with $P_c(0) = 0$. For singleton slots $R = 1$, the identification cannot fail as long as only collisions are considered, so $P_c(1) = 1$. Summing over all possible occupation numbers $R = [0, N]$ yields the total probability that a tag is identified in a slot, when captures are considered.

$$P\left(\text{slot useful}\right) = \sum_{R=0}^{N} P_c\left(R\right) \Pr\left(R\right) \tag{3.17}$$

When it is assumed that at most one tag can be identified in a slot, then the result from (3.17) is equal to the expected number of tags identified in each slot. (3.17) also gives the frame-local throughput, when the capture effect is considered. It can be expanded and written as

$$T_c = \sum_{R=1}^{N} P_c(R) \binom{N}{R} \left(\frac{1}{K}\right)^R \left(1 - \frac{1}{K}\right)^{N-R} \tag{3.18}$$

For a system that can resolve collisions of up to M tags perfectly but fails for collision slots with $R > M$ [ALR10]

$$P_c(R) = \begin{cases} 1 & 1 \leq R \leq M \\ 0 & \text{otherwise} \end{cases} \tag{3.19}$$

(3.18) becomes equal to (3.14)

In the context of RFID, the capture effect has to be considered during the inventory process while estimating the number of backlogged stations, and when calculating the optimal number of slots in an FSA frame [FS09, BDVA11].

3.7 The EPCglobal protocol

3.7.1 Protocol elements and procedures

The EPCc1g2 standard divides the communication of the reader and the tags into three phases: The selection phase, the inventory phase and the access phase, where the inventory phase essentially corresponds to FSA.

In the *selection phase*, a reader can choose the subset of tags that will be participating in the tag inventory process, by issuing one or multiple *Select* commands. The reader may select tags based on their identifier (*EPC*), Memory content and various other flags (see [EPC13, 6.3.2.12.1.1]). *Select* commands are received and processed by all tags in the range of the reader. The communication is strictly one-way, from the reader to the tags only, during this phase.

During the following phase, called *inventory phase*, the reader carries out the FSA protocol as described above. In the terms of the EPCglobal, an FSA frame is called *inventory round*; multiple inventory rounds are generally required to identify all tags. Unlike in classical FSA, however, the inventory rounds do not have a fixed duration and may be interrupted at any time by the reader. The reader issues a *Query* command to mark the beginning of a new round, which at the same time specifies the number of available slots. The end of the *Query* command marks the beginning of the first slot. To mark the beginning of the next slot, the reader transmits a *QueryRep* command. The reader may at any point issue a new *Query* or a *QueryAdjust* command, interrupting the current round and beginning a new round. This allows the reader to update the number of slots, without having to complete the current round first. Tags that did not reply up to this instant remain unidentified.

Tags reply during an inventory round, when their chosen slot is called, i.e, a tag that has chosen the first slots replies with *RN16* to the reader right after receiving the *Query* or *QueryAdjust* command. A tag that has chosen the second slot replies after receiving the first *QueryRep* command; the tag that has chosen the ith slot replies after receiving the $(i-1)^{\text{th}}$ *QueryRep*. The tag transmits

its data in two steps. The first reply (*RN16*) is a temporary 16 bit random identifier. This is *not* yet the unique identifier of the tag, which is exchanged in an additional step. If several tags reply in the same slot, this is called a tag collision. The reader may or may not be able to resolve the collision. In the latter case (see Fig. 3.13a) it will proceed by transmitting a *NAK* command or resume the inventory round by issuing the next *QueryRep*, leaving colliding tags unidentified. When a tag replies without a collision (see Fig. 3.13b) or the collision can be resolved, the reader acknowledges the tag by returning the value of *RN16* in an *ACK* command. The tag then replies with its unique identifier *EPC*. At this point the tag is regarded as identified, and the reader may proceed with the inventory round by issuing the next *QueryRep* command or start a new round by sending *Query* or *QueryAdjust*. Upon receiving any of these commands, the identified tag sets an internal flag that will silence it for further inventory rounds. The reader may reactivate silent tags by setting a corresponding flag in the *Query* command.

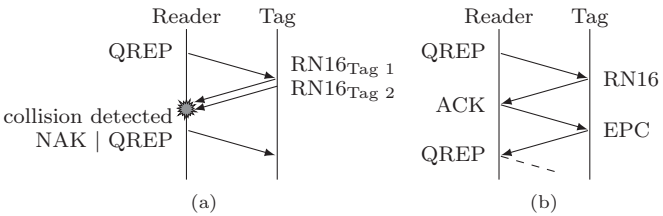


Figure 3.13: Sequence of reader commands and tag replies during one slot of the inventory process. With two colliding tag replies (a) and without a collision (b).

However, the reader may inquire additional information from a tag, once it has been identified. This constitutes the *access phase*, in which the reader may issue an arbitrary number of commands to the previously singled-out tag, before continuing the inventory round. Every standard compliant tag supports at least the *Read*, *Write*, *Lock* and *Kill* commands, which the reader may use in the access phase. The command set can be extended by additional commands, which shall not be discussed here, for the sake of brevity. The duration of the access phase, or whether it is carried out at all, depends on the application and

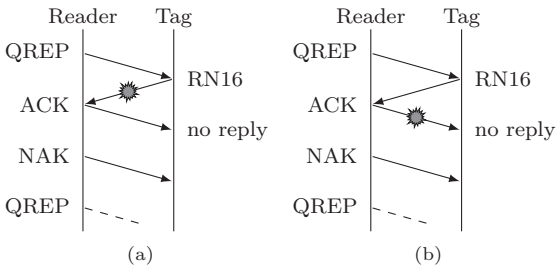


Figure 3.14: Non-recoverable transmission errors

so does the duration of the process. Thus, only the inventory phase is considered in the latter, which is more critical when it comes to the fast identification of tags in a large tag population.

3.7.2 Error behavior of the protocol

The inventory phase as described above may not only fail because of tag collisions. Both, the R-to-T and the T-to-R channels are wireless, and the transmitted information is subjected to noise and other impairments that may cause errors. The EPCc1g2 protocol specifications do not explicitly describe how the reader must recover from transmission errors. Some, but not all protocol elements, include parity check sums (Cyclic Redundancy Checks (CRCs)) that allow the receiving tag or the reader to detect transmission errors (e.g. *Query*, *EPC*). Other protocol elements, especially ones with a short frame length (e.g. *QueryAdjust*, *ACK*, *RN16*), do not include a check sum, and rely on other mechanisms, such as a unique frame length, to allow the detection of transmission errors. In any case, a frame that contains (detectable) errors is ignored by the tag, which will usually remain in its current state, awaiting further commands.

The possibilities to recover from a transmission error depend on whether a tag has yet been singled out by the reader and can be individually addressed. EPCc1g2 is a reader-talks-first system, which means tags only reply to the reader when addressed directly or in the randomly chosen slot during the inventory phase. When a transmission error occurs in the inventory phase before a tag is acknowledged (Fig. 3.14), neither the reader nor the tag is able to detect

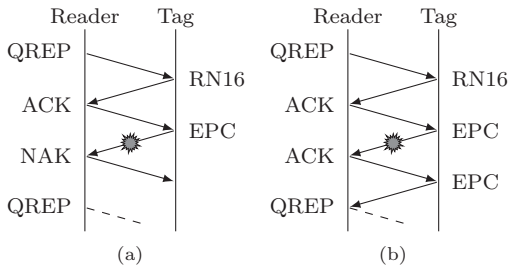


Figure 3.15: Recoverable transmission errors

the error or recover the lost data. The temporary identifier *RN16* transmitted by the tag does not include a parity check sum, so if errors are present in it, the reader replies with an *ACK* that contains an erroneous identifier. In consequence a tag must assume that the mismatching *ACK* was directed to another tag replying in the same slot. The tag ceases to transmit until the next inventory round starts. If an error occurs on the R-to-T link during the transmission of the *ACK*, the consequences are the same. Again, the tag has no possibility of detecting or correcting an error and will cease to participate in the current inventory round.

Only after the successful exchange of *RN16* and *ACK*, transmission errors can usually be detected, because all the corresponding protocol elements include parity check sums. When an error is detected, a retransmissions can be requested by the reader. Consider the examples shown in Figure 3.15. At the point when a tag transmits its *EPC* (unique identifier), it has already been singled out, and all other tags are silenced. If an error occurs in the transmission of the *EPC* frame, the reader has the option to either send a *NAK* tag, or to resend the *ACK* frame. In the first case the tag remains marked as unidentified and will take part in the next inventory round. In the second case, the tag attempts a retransmission of the *EPC* frame, which might then be received error free.

3.7.3 Markov model of the tag inventory process

The FSA based tag inventory process as described in 3.6.2 consists of several rounds, where each round is a random trial. In each round a certain number of tags compete for the slots in the FSA frame. Some of the tags are identified successfully, others remain unidentified, their replies being lost due to tag collisions or transmission errors. The probability of identifying a certain number of tags during one round can be described, in the most simple case, by a function of the number of participating tags and the number of available slots. The outcomes of consecutive rounds are then dependant, because the number of participating tags in a given round depends of the number identified in the previous round. Such systems with turn-based trials, where the probability of the trial results depends on one or few (state) variables, can be described using a Discrete Time Markov Chain (DTMC). Modeling the tag inventory process in an RFID system as a DTMC, has been proposed by [Vog02b, BDVA11], among others.

Any DTMC consists of a set of states $\mathcal{S} = \{s_0, s_2 \dots s_N\}$ and the transition probabilities $\mathcal{P} = \{p_{i,j} | i, j \in [0, N]\}$, where $p_{i,j}$ is the probability of going from state i to a state j. Furthermore, for any state, the sum of the probabilities of changing to *any* state must be one, so $\sum_{j=0}^{N} p_{i,j} = 1$ must hold true [GS97]. In general, the number of states N does not need to be finite, it is, however, for a practical system as presented here. Let N be the number of (unidentified) tags in the system and suppose the total number of tags does not increase or decrease throughout the entire inventory process. The system is then modelled by a DTMC with $N + 1$ states (Fig. 3.16). The state s_i represents the number of tags that have not been identified, i.e., the number of active tags in the given inventory round. The system starts in the state s_N, i.e., none of the N tags are identified. An inventory round that is carried out by the reader causes a transition to another state, depending on the number of tags identified in that particular round. Suppose the system is in the state s_i, then the transition $s_i \to s_i$ happens if no tags are identified, $s_i \to s_{i-1}$ if one tag is identified, etc. In general, a transition $s_i \to s_j$ occurs when $m = i - j$ tags are identified. The maximum number of tags that can be identified in one inventory round is $m_{\max} = \min(i, K)$, where i is the number of participating tags and K is the number of slots. Transitions $s_i \to s_j$ with $j > i$ cannot occur in an EPCc1g2

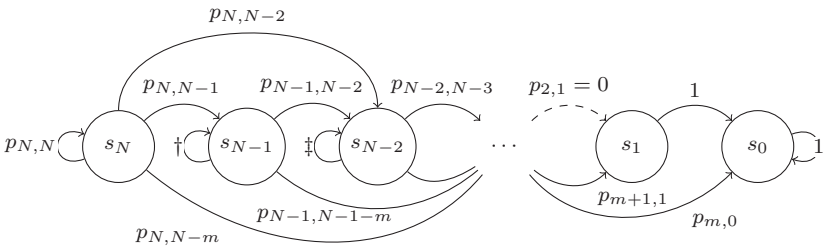


Figure 3.16: Discrete time Markov chain for the EPCglobal tag inventory, for a tag population of N.

system, because the tags only need to be identified once and are silenced after the identification.

So, finding the probability $p_{i,j}$ for a transition $s_i \rightarrow s_j$ means finding the probability that $m = i - j$ tags are identified in an inventory round with i participating tags and K slots. Without consideration of the capture effect and transmission errors, the number m of identified tags is equal to the number of slots with occupancy number $R = 1$. The stated problem is another aspect of the classical occupancy problem (see Section 3.6.2) and the solutions for the probability, as described above is found in the literature [Fel50, WW99, Vog02b, BDVA11]. According to [Fel50], the probability of finding m bins occupied by exactly R balls, when randomly placing i balls in K bins is given by

$$P_{K,i}(m,R) = \frac{(-1)^m K!i!}{m!K^i} \sum_q \frac{(-1)^q (K-q)^{i-qR}}{(q-m)!(K-q)!(i-qR)!(R!)^q} \tag{3.20}$$

with $q \geq m$, $q \leq K$ and $q \leq qR$. The evaluation of (3.20) becomes, however, difficult for large i and K. In [WW99] the authors propose an alternative way to recursively calculate the required probability, specifically for singleton slots ($R = 1$). The probability of having $m = i - j$ singleton slots, when there are K slots and i contending tags is

$$P_{K,i}(m) = \frac{S(m,i,K)}{K^i} \tag{3.21}$$

where K^i is the number of all possibilities, when placing i balls into K bins and $S(m,i,K)$ is an auxiliary function to calculate the cases that will result in m slots occupied by exactly one tag. For $i > K$

$$S(m,i,K) = \begin{cases} 0 & m = K \\ \binom{i}{m}\binom{K}{m} m!\, G(K-m, i-m) & 0 \le m < K \end{cases} \tag{3.22}$$

and for $i \le K$

$$S(m,i,K) = \begin{cases} 0 & i < m = K \vee m = i-1 \\ \binom{K}{m} m! & m = i \\ \binom{i}{m}\binom{K}{m} m!\, G(K-m, i-m) & 0 \le m < K \end{cases} \tag{3.23}$$

where

$$G(\mu,\nu) = \mu^\nu - \sum_{l=1}^{\min(\mu,\nu)} S(l,\mu,\nu)\,. \tag{3.24}$$

With (3.21) the transition probabilities of the DTMC are

$$p_{i,j} = \begin{cases} P_{K,i}(i-j) & 0 < i-j \le \min(i,K) \\ 0 & \text{otherwise.} \end{cases} \tag{3.25}$$

These can be written in matrix form

$$\mathbf{P} = \begin{pmatrix} p_{N,N} & p_{N,N-1} & p_{N,N-2} & \cdots & p_{N,2} & p_{N,1} & p_{N,0} \\ 0 & p_{N-1,N-1} & p_{N-1,N-2} & \cdots & p_{N-1,2} & p_{N-1,1} & p_{N-1,0} \\ 0 & 0 & p_{N-2,N-2} & \cdots & p_{N-2,2} & p_{N-2,1} & p_{N-2,0} \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \cdots & p_{2,2} & 0 & p_{2,0} \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix} \tag{3.26}$$

where $\mathbf{P}$ is called the transition matrix of the DTMC. The transition matrix can be used to easily calculate the probability of finding the system in a certain

state [GS97]. Let the $\mathbf{u}^{(l)}$ be a row vector with the elements $(u_N^{(l)} \dots u_0^{(l)})$, and let $u_i^{(l)}$ denote the probability that the system is in state i (tags left unidentified) after l trials (inventory rounds). The system starts with no known tags, so

$$u_i = u_i^{(0)} = \begin{cases} 1 & i = N \\ 0 & \text{otherwise} \end{cases}. \tag{3.27}$$

The probability distribution of the states after l inventory rounds is [GS97]

$$\mathbf{u}^{(l)} = \mathbf{u}\mathbf{P}^l. \tag{3.28}$$

With respect to the EPCc1g2 inventory process, $u_0^{(l)}$ is most interesting, because it contains the probability of having identified all N tags after l inventory rounds.

The state s_0 is an absorbing state, i.e., a state which cannot be left anymore, once entered. In general, a state s_i is an absorbing state when

$$p_{i,j} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \tag{3.29}$$

For a DTMC with absorbing states, the expected time, as in number of trials, to absorption can be calculated using the fundamental matrix $\mathbf{N}$ [GS97]. When $\mathbf{P}$ is given in canonical form

$$\mathbf{P} = \left(\begin{array}{c|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbf{I} \end{array} \right) \tag{3.30}$$

the fundamental matrix is

$$\mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}. \tag{3.31}$$

Note that for the EPCc1g2 inventory process (3.26), $\mathbf{P}$ has only one absorbing state, so the size of $\mathbf{I}$ as shown in (3.30) is 1×1 in this case. The fundamental matrix $\mathbf{N}$ contains the elements $n_{i,j}$, where $n_{i,j}$ is the expected number of rounds that the system will be in state j, if the system starts in state i. With $\mathbf{c}$ an all-one $N \times 1$ column vector and $\mathbf{t} = (t_N, \dots t_0)$

$$\mathbf{t} = \mathbf{N}\mathbf{c} \tag{3.32}$$

t_i is the expected time to absorption when the system starts in state i. So for the model of the tag inventory process, t_N contains the expected number of inventory rounds that is required to identify all tags, when the capture effect is neglected and frame errors do not occur.

Consideration of transmission errors in the Markov model.

The model presented above offers a rather coarse-grained vision of the tag inventory process. Every transition of the DTMC corresponds to an entire inventory round, with the corresponding probabilities depending only on the slot occupation probabilities. However, the model is easily extended, to also account for transmission errors that may occur during the inventory process in singleton slots [Sch12]. A complete model would have to account for errors in both the R-to-T and the T-to-R transmissions. The quality of the R-to-T channel can be expected to be better than the T-to-R channel, since the received power at the tag is considerably higher than at the reader (see: 3.4). So, in the model below, only errors in the T-to-R channel are considered.

A tag has to successfully transmit two frames in its slot of the inventory round: the *RN16* and its unique identifier: the *EPC*. When one of the frames contains transmission errors, the tag identification fails. The reader may request retransmissions of the *EPC* frame, under certain conditions, which will be discussed in detail in Chapter 5. For now, the assumption is that the reader does not request a retransmission in the same round. Let p_s be the probability of failure in a slot, which depends on the frame-error probability for the transmission of both tag replies. Under these assumptions, the DTMC model of the inventory process is modified in the following manner: The transition from state s_i to s_j represents the successful identification of $i - j$ tags. If there are $m = \hat{m} + f$ singleton slots (non-collision replies) in the inventory round, let $\hat{m}$ be the number of successfully identified tags, and let f be the number of failed tag replies that are lost due to errors. Now, let $\hat{P}_{K,i}(\hat{m})$ be the probability of identifying $\hat{m}$ tags when i tags compete in an inventory round with K slots, where p_s is the slot-error probability as defined above. With (3.21), $\hat{P}_{K,i}(\hat{m})$ is

$$\hat{P}_{K,i}(\hat{m}) = \sum_{f=0}^{K-\hat{m}} P_{K,i}(\hat{m}+f) \binom{\hat{m}+f}{f} p_s^f (1-p_s)^{\hat{m}} \tag{3.33}$$

where $P_{K,i}\left(\hat{m}+f\right)$ is the probability of having $m=\hat{m}+f$ singleton slots, and $p_s^f\left(1-p_s\right)^{\hat{m}}$ is the probability that f slots fail while $\hat{m}$ slots are error-free. The binomial coefficient gives the number of possibilities of having f erroneous replies in $m=\hat{m}+f$ replies overall. The modified transition matrix $\hat{\mathbf{P}}$ of the DTMC is then composed of the values given by (3.33) in the same way as shown in the previous section.

$$\hat{p}_{i,j}=\begin{cases}\hat{P}_{K,i}(i-j) & 0<i-j\leq\min\left(i,K\right)\\ 0 & \text{otherwise}\end{cases} \tag{3.34}$$

4 Forward error correction in EPCglobal systems

4.1 The introduction of FEC and its benefits

4.1.1 One-bit error correction

In this chapter, it is proposed to improve the performance of an EPCc1g2 system, by using FEC on the T-to-R link. FEC codes add redundancy to the transmitted information, to improve the probability that a receiver can recover from errors that are introduced in a noisy channel. For the system described in the previous chapter this means that all tag replies are encoded by the tag using a channel coder. After the transmission over the T-to-R channel and the decoding in the reader, this leads to an improvement of the frame-error probability p_f of the tag replies. This reduces the slot-error probability p_s defined in Section 3.7.3 which determines the transition probabilities of the DTMC model of the EPCc1g2 system. The average time to absorption, i.e., the expected number of inventory rounds to identify a given tag population, is reduced as a consequence. So, it can be expected that using FEC leads to a reduction of the identification time in the system.

Consider a simple example, based on the standard EPCc1g2 protocol elements. The standard *EPC* reply contains a cyclic redundancy check (CCITT-CRC-16) and is a (shortened) codeword of a systematic, cyclic block code with a minimum distance of $d_h = 4$. This allows the correction of all single bit errors, which has been pointed out by [MZ11] specifically for the EPCc1g2 system. Assume further that the T-to-R channel can be modeled by a BSC with the bit-error probability p_b. Then the frame-error probability when the CRC code is used for error detection only is $p_{f,r} = 1 - (1 - p_b)^{n_r}$ for the *RN16* frame, and $p_{f,e} = 1 - (1 - p_b)^{n_e}$ for the *EPC* frame, where n_r is the length of the *RN16* frame, n_e is the length of the *EPC* frame. Note that both, *RN16* and *EPC* have to be transmitted error-free over the channel, for the identification to succeed. If the transmission of the *RN16* fails, the *EPC* is not sent by the tag. So the

slot-error probability is the probability that either the *RN16* contains errors, or the *RN16* is correct, but the *EPC* contains errors

$$\begin{aligned} p_s &= p_{f,r} + (1 - p_{f,r})p_{f,e} \\ &= p_{f,r} + p_{f,e} - p_{f,e}p_{f,r} \\ &= 1 - (1 - p_{f,r})(1 - p_{f,e}) \\ &= 1 - (1 - p_b)^{n_e + n_r} \end{aligned} \tag{4.1}$$

When a decoder for error-correction is used in the reader, e.g. syndrome based [MZ11], the transmission succeeds, even if there is a single-bit error in the *EPC* frame. This slightly reduces the frame-error probability for the *EPC* frame

$$\hat{p}_{f,e} = 1 - \left[(1 - p_b)^{n_e} + n_e(1 - p_b)^{n_e - 1}\right] \tag{4.2}$$

and in consequence the slot-error probability

$$\hat{p}_s = 1 - \left[(1 - p_b)^{n_r}\left((1 - p_b)^{n_e} + n_e p_b(1 - p_b)^{n_e - 1}\right)\right] \tag{4.3}$$

The expected number of inventory rounds required to identify a tag population of a given size can be calculated for both variants using the relations presented in Section 3.7.3. The length of the *RN16* is fixed $n_r = 16$ for EPCc1g2 systems and an *EPC* length of 96 bit was chosen, which implies $n_e = 128$. The number of rounds can then be plotted as a function af the bit-error probability p_b of the BSC. Figure 4.1 shows a plot for an example of a tag population with $N = 32$ tags and a fixed number of slots $K = 32$. The expected number of inventory rounds to identify all tags was calculated using the characteristic matrix of the DTMC model with the transition probabilities from (3.33), and plotted for a range of $10^{-4} < p_b < 10^{-1}$. It can be seen from the plot, that for low bit-error probabilities, both variants do not differ in their performance, because for $p_b \to 0$ (3.33) becomes equal to (3.21). In the error-free case, an average of 6.63 rounds are required to identify all tags. For higher p_b, the variant employing error correction proves superior to the other variant. At a bit-error probability of $p_b = 10^{-2}$, the expected number of rounds required to identify all tags is $t_1 \approx 85.5$, when error-detection is used only. With the variant that can correct single-bit errors, the expected number of rounds is only $\hat{t}_1 \approx 21.19$. Note that this improvement can be achieved without any changes to the original protocol.

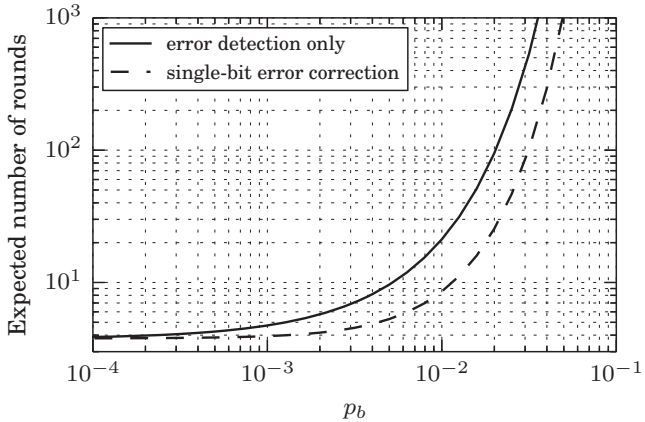


Figure 4.1: Expected number of rounds required to identify a tag population of N=32 tags, fixed number of slots $K = 32$ and transmission over a BSC.

The BSC based modelling of the system behavior gives some insight for the analysis of the system. The analysis can be extended for the AWGN channel model, by substituting the bit-error probability in the AWGN channel, at a certain SNR for p_b. If the baseband-coding scheme is taken into consideration, then the bit-error probability of the T-to-R channel is given by [SD06, eq. 9][1]

$$p_b = \operatorname{erfc}\left(\sqrt{\frac{E_s}{N_0}}\right)\left[1 - \frac{1}{2}\operatorname{erfc}\left(\sqrt{\frac{E_s}{N_0}}\right)\right] \tag{4.4}$$

for the FM0 code. The results for the AWGN channel are obtained by using (4.4) in (4.1) and (4.3). Figure 4.2 shows the expected number of rounds to identify all $N = 32$ tags in the system, for varying ratios of E_s/N_0, where E_s is the symbol energy and N_0 is the noise power spectral density of the AWGN channel. The two variants may then be compared for a specific number of expected rounds. E.g., for 10 inventory rounds, the variant using error correction is able to perform at $E_s/N_0 \approx 5$, while the error detection variant requires $E_s/N_0 \approx 6$.

[1] $\operatorname{erfc}(\cdot) = 2\mathrm{Q}\left(\sqrt{2}\,\cdot\right)$

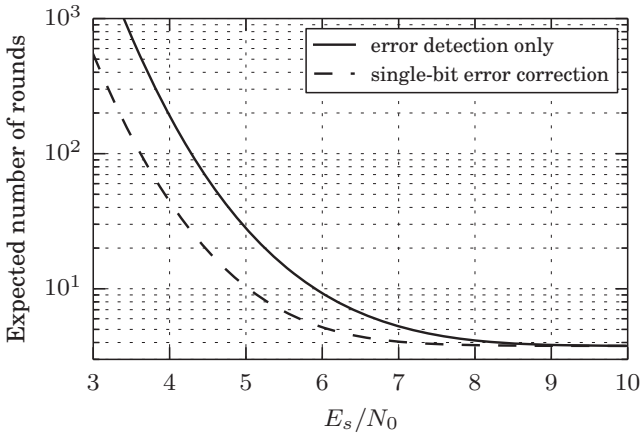


Figure 4.2: Average number of rounds required to identify a tag population of N=32 tags, fixed number of slots $K = 32$ and transmission over an AWGN channel.

This improvement of about one decibel can be achieved exploiting the standard protocol elements.

4.1.2 Extended range system with convolutional codes

The previous example illustrates the usefulness of using FEC in an EPCc1g2 system. However, the coding gain of the hamming block code is low, so different channel codes should be taken into consideration. Of the two classes of channel codes: block and convolutional codes, the latter one has the advantage that low complexity soft-decision decoders are available. The use of a convolutional code is evaluated below, using the method described above. Furthermore, an estimate for expected increase of the read range of the system is given, based on the free-space propagation model from 3.4.

Two contradictory facts have to be taken into consideration, when choosing the convolution code to be used. The asymptotic coding gain of a convolutional code depends on the code rate and the free distance of the code. The free distance is generally higher for lower code rates and greater memory orders. A

high memory order implies a high complexity of the decoder though. Additionally, when zero termination is used to produce a finite-length code words, a high memory order leads to a rate loss, and thus decreases the actual coding gain [Fri96]. Here, a code with constant length $\nu = 6$ and $d_f = 8$ was chosen, which has the generator polynomials 53_8 and 75_8 (octal notation). The upper bound for the coding gain with soft-decision decoding is $G_a \approx 6\,\text{dB}$ [Fri96, (9.5.4)]. For the analysis of the coded system with the DTMC model, the frame-error rates at the output of the channel decoder are required. For convolutional codes there are some bounds for the bit-error probability at the output of the Viterbi decoder, but no closed form description of the frame-error probability for the terminated code is available. Instead, here, the frame-error probabilities were obtained by monte carlo simulation.

To illustrate the effects of the coding on the read range of a system in [Sch12] the performance of the system is expressed in terms of the average number of inventory rounds for the identification of a fixed-size tag population, depending on the average tag-reader distance. The frame-error rates in the system are obtained using a free-space propagation model (see 3.4), together with a set of common tag and reader parameters (see 3.3), to calculate the signal-to-noise ratio of the tag signals at the reader's receiver. An AWGN channel model is used, and the slot-error probability for a system without FEC is calculated as shown in (4.1) using the relation from (4.4). The slot-error probability for the system using the convolutional code, is obtained by simulation. In addition to the analytical results based on the DTMC model, the system behavior is also simulated, and the simulation results are presented for comparison.

In both, analysis and simulations, a tag population of $N = 100$ tags and a reader that uses a fixed frame size of $K = 2^7 = 128$ is considered. The condition for a successful identification of a tag is: The tag must reply in a singleton slot and both, the *RN16* and the *EPC* frames, must be received error-free at the reader (no slot-error). In the case of the variant with convolutional coding, both frames are zero-terminated codewords, encoded with the $R_c = 1/2, \nu = 6$ encoder, which is described above. For a fair comparison, the encoded frames are transmitted over a channel with reduced signal-to-noise ratio, guaranteeing that the equal energy E_b is spent for transmitting each information bit, in both cases. For $R = 1/2$, this implies a 3 dB reduction of the symbol energy E_s, when coding is used. In practice, this could be achieved by doubling the

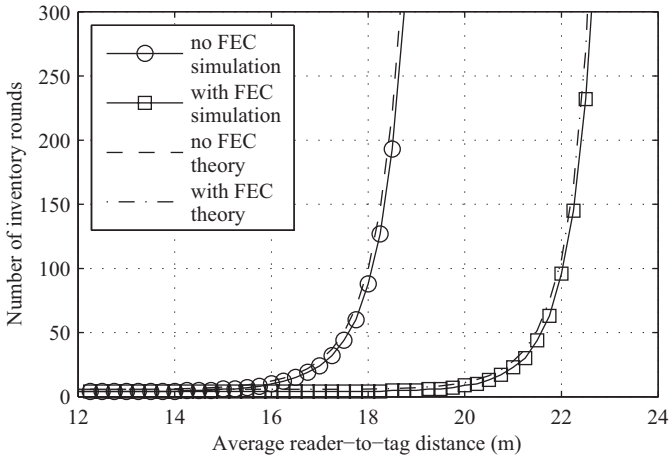


Figure 4.3: Number of inventory rounds to read all tags [Sch12]
(ITG-Fachbericht 235: Smart SysTech, 12-13 June 2012, Munich, VDE VERLAG, Berlin, Offenbach, 2012)

symbol rate of the tag transmissions for the encoded frames. This implies that the average duration of an inventory round remains equal, which is important, because only the number of inventory rounds is compared in the latter, not taking into account their duration.

Figure 4.3 shows the expected number of inventory rounds required to identify all $N = 100$ tags, plotted against the average tag-reader distance. Figure 4.4 shows the number of tags that can be identified by the reader, when the number of inventory rounds is limited to ten. The Figures show the analytical results with dashed lines and the simulation results with solid lines with additional markers. As can be seen from Figure 4.3, for good channel condition, i.e., short distances between the reader and the tags, there is no difference between the variant without (dashed/circles) and the variant with coding (dot-dashed/squares). Up to a distance of about 14 m, the curves do not differ. In this region, the bit-error probability is low and thus frame-errors rarely occur. The duration of the inventory process is then limited due to tag collisions. The probability of tag collisions depends only on the number of tags in the system

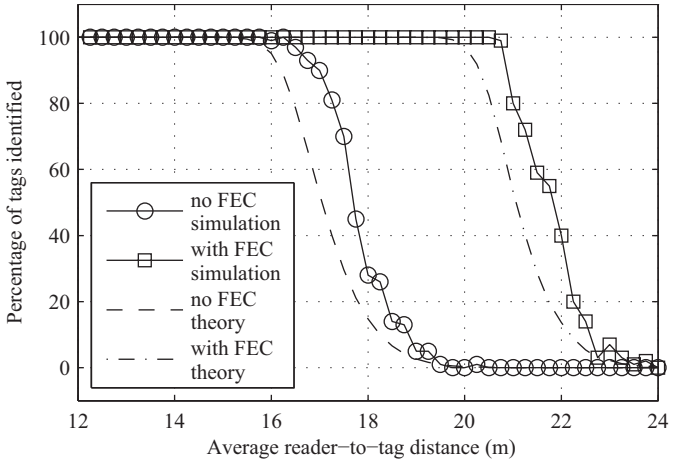


Figure 4.4: Percentage of tags read after 10 rounds [Sch12]
(ITG-Fachbericht 235: Smart SysTech, 12-13 June 2012, Munich, VDE VERLAG, Berlin, Offenbach, 2012)

and the number of slots. For the selected fixed FSA frame-size of $K = 128$ and $N = 100$ tags, it takes an average of five rounds to identify all tags.

When the T-to-R distance increases, the SNR in the T-to-R channel decreases according to the $1/d^4$ slope of the received tag-power at the reader that is given in (3.5). With decreasing SNR, the bit- and frame-error probabilities increase rapidly. Consequently, the required number of inventory rounds grows exponentially. At a distance of $d \approx 17.5\,\text{m}$, the average number of required inventory rounds in the uncoded case increases ten-fold, at a distance of $d \approx 18\,\text{m}$ it grows twenty-fold, compared to the error-free case. Given that tags, in a practical scenario, do not necessarily remain within same distance of the reader for a longer period of time, the likelihood that those tags are not identified by the reader increases drastically. Figure 4.4 presents the results in a different manner, to illustrate this claim. Consider a case where the identification time is limited, e.g. only 10 inventory rounds can be carried out by the reader to identify the entire tag population. The reader is then only able to do so, if the T-to-R distance is below 15 m. At higher distances, the fraction of successfully

identified tags decreases rapidly and at a distance of 20 m, the reader is not able to identify any tags within the limited amount of rounds.

The curves for the system using FEC on the T-to-R link, show a similar behaviour, as the standard system without coding. However, the curves are shifted to higher distances. Up to a distance of ≈ 18 m, the number of inventory rounds to identify all tags does not increase significantly. At a distance of $d \approx 21.5$ m, the required number of rounds increases ten-fold, at $d \approx 22$ m twenty-fold, compared to the error-free case. If a round limit of ten is imposed, the system with coding can operate at distances of up to $d \approx 20$ m, while still being able to identify all tags with high probability. In conclusion, using FEC on the T-to-R link can significantly increase the operational range of a system. For the chosen code and the given system parameters, the range can be extended by approximately four meters.

Discussion of the results

The results as presented above, are based on a model of the system, for which some important assumptions about the properties of the system are made. First and foremost, it is assumed that the tags will received a sufficient amount of energy to operate over the entire range of distances that was analyzed. If this is not the case, then the system might not be limited by the quality of the T-to-R link, but by the power available at the tags. This constraint is satisfied, when BAP tags are used instead of passive tags [CYX+09] (see also 3.4). Note also that there are systems on the market that implement a separation of the concerns: power and communication [Rob09]. In these systems, the power constraint is also satisfied at high (communication) distances.

Other aspects of the model are: free-space propagation and an AWGN channel. In the most practical cases, the actual propagation environment is cluttered with obstacles in or near the line-of-sight, so that there is no free-space propagation [GD09, LGS09, LGSV09] and [KIS03, GD07, GD08] suggest that severe fading might be occurring in the T-to-R channel. While this reduces the overall range of the system, coding gains are usually higher in fading channels, so it can be expected that the improvements can still be found in more realistic scenarios.

Lastly, all quantitative results presented above were obtained using one specific set of parameters (transmission power, reader and tag sensitivity, etc.), so

all absolute results will vary, when the parameters change, e.g. because a regulatory instance imposes a different constraint regarding the maximum transmit power or different tags are used.

4.2 Optimal decoding of the baseband codes

4.2.1 Description of the coding rules and coders

It has been implied in the first part of the chapter that the reader is able to achieve the optimum bit-error probability given in [SD06], when decoding the tag signals. The following sections contain a more in-depth description of baseband code (or line codes) used on the T-to-R link, the corresponding coding rules and the decoding methods. The EPCc1g2 standard specifies one baseband coding scheme for R-to-T communication and two distinct baseband coding schemes for the T-to-R communication. As throughout the work, this section will focus on the properties of the T-to-R link.

The two baseband codes proposed in the standard for the T-to-R communication are called *FM0* and *Miller modulated subcarrier*. The FM0 code, also denoted as *bi-phase space*, is a differential baseband code that maps the binary information bits to one out of four baseband code symbols shown in Figure 4.5c. The encoding is defined as follows in [EPC13, 6.3.1.3.2.1]:

> FM0 inverts the baseband phase at every symbol boundary; a data-0 has an additional mid-symbol phase inversion.

The dependency of successive symbols can be expressed in a state diagram as shown in Figure 4.5a. Any transition corresponds to the current input value, the state marks the output waveform of the baseband coder (Figure 4.5c). The notation as shown in Figure 4.5 is taken from the EPCglobal standard [EPC13]. The coding rules of the FM0 code can, however, be expressed differently. From the verbal description of the coding rules, it is obvious that the coder possesses only a memory of one (symbols only depend on the phase at the end of the previous symbol), thus two states are sufficient to describe the coder. The states are reduced when translating the Finite-State Machine (FSM) shown in Figure 4.5a from the Moore machine representation to a Mealy machine. The Mealy machine representation also immediately gives the trellis representation

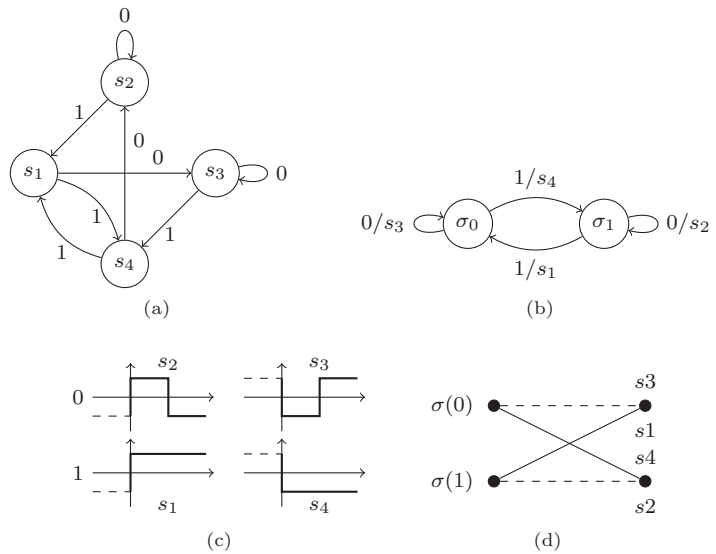


Figure 4.5: FM0 generator state diagram in Moore (a) [EPC13, Fig 6.8] and Mealy representation (b), the corresponding waveforms (c) and the code trellis (d)

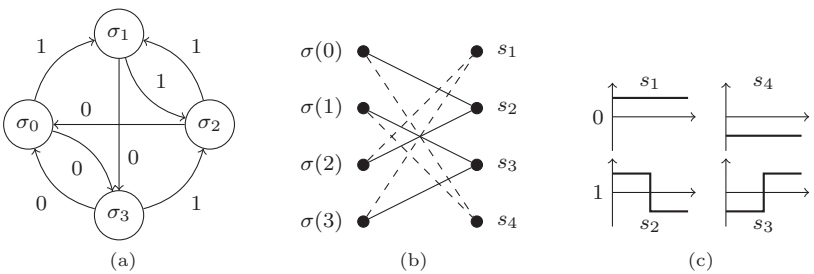


Figure 4.6: State machine [EPC13] and trellis representation [SD06] of baseband Miller and the associated waveforms.

of the FM0 code 4.5d. The trellis describes a non-linear convolutional code. The possibility of describing the FM0 code in this way, has previously been proposed by [SD06], moreover, the authors provide the bit-error probabilities for the optimal decoding of FM0. They further point out that the FM0 code is essentially equal to differential Manchester. This becomes apparent if the observation interval (symbol bounderies) is shifted by $T_s/2$ at the decoder. The code then only consists of two antipodal bi-phase symbols. Differential Manchester in turn, has an equal bit-error probability as differentially encoded BPSK. [SD06] further points out that it is sufficient to observe two successive symbols to decode while achieving an optimal error probability.

The second baseband code used in EPCc1g2 systems is the Miller code. A verbal description of the coding rule is given in [EPC13, 6.3.1.3.2.3]

> Baseband Miller inverts its phase between two data-0s in sequence. Baseband Miller also places a phase inversion in the middle of a data-1 symbol.

The Miller coder has a memory of two, i.e., four states to store the value of the previous information bit and the symbol phase at the end of the previous symbol. The state diagram description from the standard can directly be translated into the elementary Trellis segment as given in Figure 4.6b

While both baseband codes can be decoded by making symbol-by-symbol decisions, this does not yield optimal results, as pointed out in [SD06]. A decoder must take previous symbols into account, in order to achieve better results.

4.2.2 Decoding and bit-error probability of FM0

Obtaining the decision rule and the bit-error performance of the FM0 code is straightforward, when it is taken into account that FM0 is equal to differential Manchester. Differential Manchester uses a differential coder at the transmitter and uses a set of two antipodal signals, s_2 and $s_3 = -s_2$ for the transmission of the encoded symbols. If $d_i = a_i \oplus d_{i-1}$ is the differentially coded sequence, the receiver recovers the information by decoding $\hat{a}_i = \hat{d}_{i-1} \oplus \hat{d}_i$. So a bit error in $\hat{a}_{i-1}$ occurs if either $\hat{d}_{i-1}$ (x)or $\hat{d}_i$ contains a symbol error. If both $\hat{d}_{i-1}$ *and* $\hat{d}_i$

contain symbol errors, the errors cancel out, and $\hat{a}_i$ is decoded correctly. Hence, the exact bit error probability of FM0 and differential Manchester is

$$p_b = p_{sym}(1 - p_{sym}) + (1 - p_{sym})p_{sym} = 2p_{sym}(1 - p_{sym}) \tag{4.5}$$

where p_{sym} is the symbol-error probability. For antipodal signalling over an AWGN channel the symbol-error probability is, independent of the actual shape of the pulse [PS08],

$$p_{sym} = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_s}{N_0}}\right). \tag{4.6}$$

Substituting (4.6) in (4.5) yields the result previously given in (4.4), which is equal to the bit-error probability of coherently detected, differentially coded BPSK. Clearly, as also pointed out by [SD06], the FM0 trellis can also be used to decode FM0, by carrying out a Maximum likelihood sequence estimation (MLSE) using Viterbi's algorithm. The bit-error probability after the Viterbi decoder, however, is not improved compared to differential decoder.

FM0 can also be decoded on a symbol-by-symbol basis. The decoder must simply compare the decisions on each half of the symbol. Equal phases are decodes as "1", distinct phases as "0". Then, if one of the half-symbols is received with an error, this will result in a bit-error. If both half-symbols are inverted, the decoder will give the correct result, despite the transmission errors. The bit error probability can then be described by (4.5) as well, but the meaning of p_{sym} changes. While for the previous case, p_{sym} was the error-probability of antipodal signalling of a symbol of the duration T_s, in the latter case it refers to the half-symbols with duration $T_s/2$. The half-symbols are still antipodal signals, they do, however, have only half of the symbol-energy. So the symbol-error probability for the half symbol becomes

$$p_{sym} = \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_s}{2N_0}}\right) \tag{4.7}$$

and the bit-error probability for symbol-by-symbol decoding is

$$p_b = \operatorname{erfc}\left(\sqrt{\frac{E_s}{2N_0}}\right)\left[1 - \frac{1}{2} \operatorname{erfc}\left(\sqrt{\frac{E_s}{2N_0}}\right)\right] \tag{4.8}$$

which corresponds to a 3 dB loss compared to (4.4).

4.2.3 Decoding and bit-error probability of baseband Miller

Baseband Miller, just like FM0, can be decoded on a symbol-by-symbol basis or using MLSE. The symbol-by-symbol decoder for baseband Miller is equal to the one of FM0, except for the fact that the representations for 0 and 1 are interchanged. Thus, the symbol-by-symbol decoder for baseband Miller has the same bit-error probability as given in (4.8). Clearly, the symbol-by symbol decoder does not yield the optimum decoding result, because it ignores the inherent dependency of successive symbols in the Miller code (E.g. waveform s_1 will never be followed by s_3). Better decoding results can be achieved with a Viterbi decoder that performs a MLSE. There is, however, no straightforward way of giving a closed-form expression for the bit-error probability for the Viterbi decoder. [SD06] argues that the Viterbi decoder asymptotically performs just like the symbol-by-symbol decoder, based on the Miller trellis and its free distance.

4.2.4 Bit-error-rate plots for FM0 and baseband Miller

Figure 4.7 shows a plot of the bit-error probabilities for the different decoders described above. Additionally, the plots show the Bit-Error Rates (BERs) that were obtained by computer simulation. The dashed line shows the analytical results from (4.8) for the symbol-by-symbol decoding of FM0 and baseband Miller. The solid line shows the bit-error probability which was given in (4.4) for the decoder, which takes into account the equality of differential Manchester and FM0. The 3 dB improvement of the latter one, over the symbol-by-symbol decoder, is clearly visible, also in the simulation results. The symbol-by-symbol decoders of FM0 ($+$) and baseband Miller ($\square$) yield results predicted by (4.8), while the differential Manchester decoder ($\times$) is superior in performance.

The simulations of the Viterbi decoder for baseband Miller ($\lozenge$) show that for high signal-to-noise ratios, the bit-error rate approaches the curve of the symbol-by-symbol decoder. It can be concluded, that using the Viterbi algorithm to decode baseband Miller is only useful at low signal-to-noise ratios. At high signal-to-noise ratios, the coding gain vanishes, and the extra complexity required in the Viterbi decoder is not justifiable.

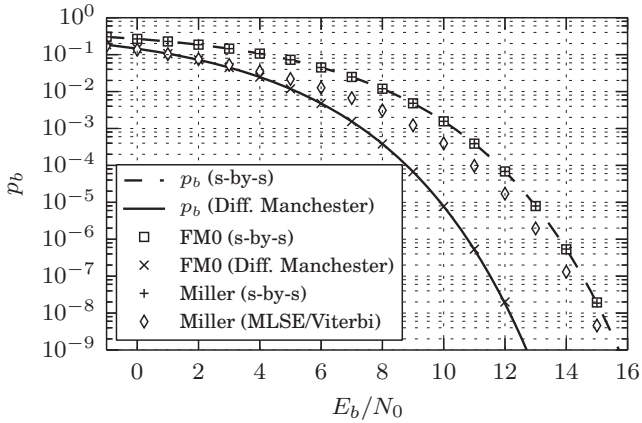


Figure 4.7: Bit-error probabilities and simulated bit-error rates for FM0 and baseband Miller in the AWGN channel for different decoders.

4.3 The concatenation of baseband and channel codes

4.3.1 Concatenated codes

All previous evaluations of the decoding methods, and the resulting bit-error probabilities, are based on the premise that the input values of the baseband coder are independent binary random values. Dependencies of successive baseband-code symbols are considered according to the encoding rules of the baseband code. In this work however, it is proposed that a channel code should be used in the system. The channel code introduces redundancy to the source information, to enable the detection and correction of transmission errors. Figure 4.8 shows the general structure of a system with a channel coder and decoder, and an inner baseband code. As pointed out in [Bos98, ch. 9], two different paradigms exist for approaching a system with concatenated coders (though Bosserts mainly refers to schemes of concatenated channel codes). The inner (baseband) coder, together with the channel and the corresponding decoder, can be seen as a *super channel*, whose properties are defined by the actual channel and the properties of the inner code. Alternatively, the output of the concatenated coders can be regarded as *one* code, and the process of recovering

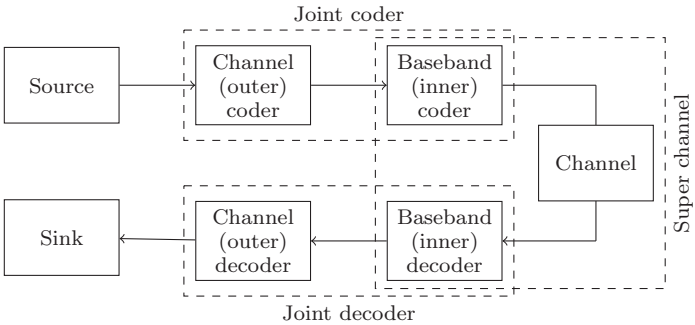


Figure 4.8: General structure of a communication system with concatenated codes.

the source information in the receiver can take place in a decoder that *jointly* decodes the concatenated code. Note that in this work, the inner codes are *no* channel codes in the sense that they do not add redundancy to the overall code.

4.3.2 A posteriori probability decoding of FM0

In a system which uses FM0 as baseband code and an additional outer channel code, the FM0 coder, the channel and the FM0 decoder can then be regarded as a super channel, over which the outer channel code values are transmitted. If the baseband decoder takes hard-decisions on each received symbol, and outputs the differentially decoded result, the super channel will have a bit-error probability as described in (4.5). However, some information is lost in the baseband decoder, when taking the hard-decisions. In a concatenated system it is preferable that the baseband decoder outputs soft-decisions, e.g. the a posteriori Log-Likelihood Ratios (LLRs) for the decoded symbols, which can then be used by an outer Soft Input Soft Output (SISO), to improve the decoding result.

[BCJR74] describes a decoding algorithm widely know as the BCJR algorithm or as A Posteriori Probability (APP) decoder. The algorithm is optimal, in the sense that it minimizes the symbol-error probability, i.e. the error probability for every trellis segment. The BCJR algorithm, or one of its variants [RVH95],

can be used to implement an APP decoder for FM0. The decoder shall calculate the a posteriori LLR of the decision on the t^{th} bit

$$L\left(u_t\right) = \log \frac{p\left(u_t = 1|\mathbf{y}\right)}{p\left(u_t = 0|\mathbf{y}\right)} \tag{4.9}$$

based on the entire received sequence of baseband code symbols $\mathbf{y}$. The calculation is based on the trellis (Figure 4.5d) of the code [RVH95, BG96].

$$L\left(u_t\right) = \log \frac{\sum\limits_{\sigma_t} \sum\limits_{\sigma_{t+1}} \gamma_0\left(y_t, \sigma_t, \sigma_{t+1}\right) \cdot \alpha_{t-1}\left(\sigma_{t-1}\right) \cdot \beta_t\left(\sigma_t\right)}{\sum\limits_{\sigma_t} \sum\limits_{\sigma_{t+1}} \gamma_1\left(y_t, \sigma_t, \sigma_{t+1}\right) \cdot \alpha_{t-1}\left(\sigma_{t-1}\right) \cdot \beta_t\left(\sigma_t\right)} \tag{4.10}$$

Recall that every trellis segment represents one possible input of the coder. The output of the coder depends on the previous state of the coder and the input u_t, which shall be determined by the decoder. The APP decoder calculates the probability for every possible transition in the trellis. Each transition goes from a state σ_{t-1} to another state σ_t, where $\sigma_t, \sigma_{t-1} \in [0, 2^\nu - 1]$, for a binary code with memory ν. In (4.10) $\alpha_{t-1}(\sigma_{t-1})$ is the probability that the coder is in the state σ_{t-1} at the instant $t-1$, and $\beta_t(\sigma_t)$ is the probability that the coder is in the state σ_t at the instant t. $\alpha_{t-1}(\sigma_{t-1})$ is calculated in the forward recursion and includes the dependencies of the current symbol from all prior symbols $(1 \ldots t-1)$. $\beta_t(\sigma_t)$ is calculated in the backward recursion, and depends on all symbols after the current one $(t+1 \ldots N$, where N is the length of $\mathbf{y})$. Furthermore, the probability of the transition depends on the channel observation y_t and the a priori probability $P\left(u_t\right)$, which are combined in the $\gamma\left(y_t, \sigma_t, \sigma_{t+1}\right)$ terms. The decoder sums up the probabilities for all transitions caused by a zero-input to the coder (numerator) and all transitions caused by a one-input (denominator). Finally, the output $L\left(u_t\right)$ is the a posteriori LLR for the t^{th} bit u_t.

For a simple differential code, such as FM0/differential Manchester, (4.10) can be simplified, taking into consideration the calculating of α_{t-1} and β_t. Specifically in the FM0 trellis, two branches end in each state. So, the forward recursion calculates

$$\alpha_{t-1}\left(0\right) = \alpha_{t-2}(0) \frac{P(u_{t-1} = 0) p(y_{t-1}|x_{t-1} = s_2)}{p(y_t)} +$$

$$\alpha_{t-2}(1)\frac{P(u_{t-1}=1)p(y_{t-1}|x_{t-1}=s_2)}{p(y_t)} \tag{4.11}$$

$$\alpha_{t-1}(1)=\alpha_{t-2}(0)\frac{P(u_{t-1}=1)p(y_{t-1}|x_{t-1}=s_3)}{p(y_t)}+$$
$$\alpha_{t-2}(1)\frac{P(u_{t-1}=0)p(y_{t-1}|x_{t-1}=s_3)}{p(y_t)} \tag{4.12}$$

Under the condition that the a priori probabilities are either equal or unknown, the decoder will have to assume $P(u_{t-1}=1)=P(u_{t-1}=0)=1/2$, and (4.11) and (4.12) can be simplified

$$\alpha_{t-1}(0)=\frac{p(y_{t-1}|x_{t-1}=s_2)}{2p(y_{t-1})}\underbrace{[\alpha_{t-2}(0)+\alpha_{t-2}(1)]}_{=1} \tag{4.13}$$

$$\alpha_{t-1}(1)=\frac{p(y_{t-1}|x_{t-1}=s_3)}{2p(y_{t-1})}\underbrace{[\alpha_{t-2}(0)+\alpha_{t-2}(1)]}_{=1} \tag{4.14}$$

For the backward recursion, $\beta_t(\sigma_t)$ is

$$\beta_t(0)=\beta_{t+1}(0)\frac{P(u_{t+1}=0)p(y_{t+1}|x_{t+1}=s_2)}{p(y_{t+1})}+$$
$$\beta_{t+2}(1)\frac{P(u_{t+1}=1)p(y_{t+1}|x_{t+1}=s_3)}{p(y_{t+1})} \tag{4.15}$$

$$\beta_t(1)=\beta_{t+1}(0)\frac{P(u_{t+1}=1)p(y_{t+1}|x_{t+1}=s_2)}{p(y_{t+1})}+$$
$$\beta_{t+2}(1)\frac{P(u_{t+1}=0)p(y_{t+1}|x_{t+1}=s_3)}{p(y_{t+1})} \tag{4.16}$$

If $P(u_{t+1}=1)=P(u_{t+1}=0)=1/2$ holds, then $\beta_t(0)=\beta_t(1)$. Substituting (4.11)-(4.12) in (4.10), all β_t terms, $p(y_t)$, as well as the factor $1/2$ get canceled.

$$\begin{aligned}L(u_t)&=\log\frac{\gamma_0(y_t,0,0)\,\alpha_{t-1}(0)\cdot\beta_t(0)+\gamma_0(y_t,1,1)\,\alpha_{t-1}(1)\cdot\beta_t(1)}{\gamma_1(y_t,0,1)\,\alpha_{t-1}(0)\cdot\beta_t(1)+\gamma_1(y_t,1,0)\,\alpha_{t-1}(1)\cdot\beta_t(0)}\\&=\log\frac{p(y_t\mid x_t=s_2)\,p(y_{t-1}\mid x_{t-1}=s_2)+p(y_t\mid x_t=s_3)\,p(y_{t-1}\mid x_{t-1}=s_3)}{p(y_t\mid x_t=s_3)\,p(y_{t-1}\mid x_{t-1}=s_2)+p(y_t\mid x_t=s_2)\,p(y_{t-1}\mid x_{t-1}=s_3)}\\&=\log\frac{1+\frac{p(y_t|x_t=s_3)}{p(y_t|x_t=s_2)}\frac{p\left(y_{t-1}|x_{t-1}=s_3\right)}{p\left(y_{t-1}|x_{t-1}=s_2\right)}}{\frac{p(y_t|x_t=s_3)}{p(y_t|x_t=s_2)}+\frac{p\left(y_{t-1}|x_{t-1}=s_3\right)}{p\left(y_{t-1}|x_{t-1}=s_2\right)}}\end{aligned} \tag{4.17}$$

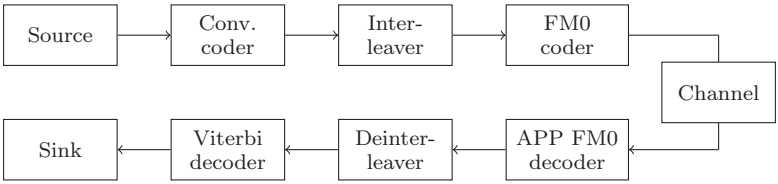


Figure 4.9: Model of the T-to-R link for the system with an outer channel code.

Observe that $s_2 = -s_3$. So, with the definition of the LLR

$$L(x_t) = \log \frac{p(y_t \mid x_t = s_3)}{p(y_t \mid x_t = -s_3)} \tag{4.18}$$

(4.17) can be rewritten as

$$L(u_t) = \log \frac{1 + e^{L(x_t)+L\left(x_{t-1}\right)}}{e^{L(x_t)} + e^{L\left(x_{t-1}\right)}} \tag{4.19}$$

(4.19) is a well known relation in the log-likelihood algebra. It describes the LLR of the mod 2 sum (XOR) of two binary inputs x_t and x_{t-1}, with the associated LLRs $L(x_t)$ and $L(x_{t-1})$ [HOP96][Bos98, Appendix B]. So, the result of the APP decoder without a priori information is equal to the differential decoding as described in Section 4.2.2. Often, the following approximation is used for (4.19)

$$\begin{aligned} L(u_t) &= L(x_t \oplus x_{t-1}) = L(x_t) \boxplus L(x_{t-1}) \\ &\approx \operatorname{sign}(L(x_t)) \operatorname{sign}(L(x_{t-1})) \min(|L(x_t)|, |L(x_{t-1})|) \end{aligned} \tag{4.20}$$

In conclusion, the reliability of the decoder's decision is determined by the symbol LLR of the less reliable symbol.

The simplifications concerning the forward and backward recursions can only be made as long as it is valid to assume equal a priori probabilities for both values of u_t. Otherwise, the full forward and backward recursions have to be calculated for the entire length of $\mathbf{y}$ sequence, to calculate the exact a posteriori LLRs.

The performance of the APP decoder for FM0 was verified in a simulation. The simulated system is set up as shown in Figure 4.9, with a convolutional

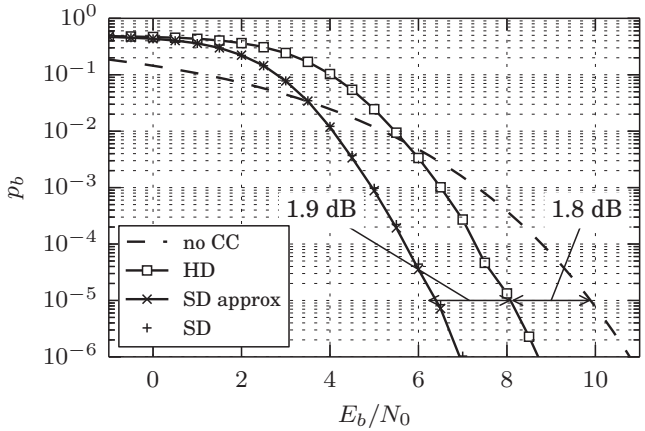


Figure 4.10: Comparison of the bit-error rates in a system with the FM0 baseband code without (no CC) and with an outer convolutional code $[23_8, 35_8]$, and different variants of the FM0 decoder: hard decoder (HD), soft decoder (SD) and soft decoder with approximated calculation of the LLRs (SD approx).

code as the outer code, and the FM0 baseband code as the inner code. An interleaver is placed between the channel coder and the baseband coder, and the corresponding deinterleaver is placed between the baseband and the channel decoder. As pointed out before, when a single symbol is received in error, the baseband decoder will produce two-bit burst errors at its output. The deinterleaver spreads the errors pseudo randomly, so they become distributed independently at the input of the channel decoder. The interleaver also fulfills an important function in the system with an iterative decoder, proposed later in Section 4.4.2. Figure 4.10 shows simulation results for the bit-error rate in a system with different FM0 baseband decoders. The channel code used in the simulations is a convolutional code with a memory of $\nu = 4$ and the generator polynomials $[23, 35]$ (in octal notation). This code is an Optimum Free Distance (OFD) code with a free distance of $d_f = 7$ [Bos98]. The Viterbi algorithm is used to decode the convolutional code.

If there is no outer decoder, the performance of the (soft) APP decoder is equal to a hard decoder, and the bit-error probability is given by (4.5) (dashed curve). When an outer code is used in the system, the variant with the FM0 soft decoder outperforms the hard decoding variant, as expected. When the inner decoder takes hard decisions, the coding gain compared to the system without the outer channel code is $G_{hard} \approx 1.8\,\mathrm{dB}$, at $p_b = 10^{-5}$. With an inner soft-decoder, the gain is $G_{soft} \approx 3.7\,\mathrm{dB}$. This is somewhat less than the coding gain that can be achieved with the same convolutional code, when coherent BPSK is used as opposed to the differential FM0 code. In that case the asymptotic coding gains are $\hat{G}_{hard} \approx 2.43\,\mathrm{dB}$, and $\hat{G}_{soft} \approx 5.44\,\mathrm{dB}$ for hard and soft decision decoding, respectively. The simulations also show that there is virtually no difference between the two soft-decoder variants from (4.19) (+), and the variant (4.20), which uses the LLR approximation (x).

4.4 Iterative decoder

4.4.1 Iterative and turbo decoder, and their applications

Decoding the FM0 baseband code in an EPCglobal system requires only a simple differential decoder, which can achieve a performance that is equal to the APP decoder, as long as no a priori information is available. For baseband Miller, only a symbol-by-symbol decoder, or the trellis based Viterbi, or APP decoders are available.

In a system with an outer channel code as depicted in Figure 4.8, the channel coder is followed by the baseband coder. Together, both encoders produce a serial concatenated code. Though concatenated codes are usually composed of two or more *channel* codes, the baseband code can also be regarded as the inner code of a concatenated system. Generally, concatenated codes are valuable because short codes can be combined to form longer codes, concatenated codes correct independent single-bit errors as well as burst errors, and the decoding complexity is low if the codes are decoded separately [Bos98].

These advantages of concatenated codes were already pointed out by [For65]. Later in the seventies and eighties, concatenated systems with Reed Solomon outer and convolutional inner codes were widely used by the NASA for deep space communication. Another important milestone was the discovery of the

turbo codes [BGT93]. Berrou et al. proposed a system with parallel concatenated convolutional coders. The most important contribution in their work is the iterative decoder that is proposed by the authors. The *turbo decoder* consists of individual SISO decoders for each of the concatenated codes. Both decoders are used to decode the convolutional codes separately, but the knowledge that each decoder gains during the decoding is passed along to the other decoder in the form of the so-called *extrinsic information*. Through several decoding iterations, each decoder uses the extrinsic information of the other decoder as a priori information on the information sequence. This exchange of information between the two stages drastically improves the decoding performance.

Since the discovery of turbo codes, the turbo principle, i.e. the iterative decoder structure, has been extended to serial concatenated codes [BM96] and has also been used beyond the scope of channel coding. E.g. [HL99] proposes to use a turbo approach for the incoherent demodulation of Differential Phase Shift Keying (DPSK). Hoher's work sparked the idea of using an iterative decoder to iteratively decode the baseband and channel code in the EPCc1g2 system. The corresponding results that are presented in the remainder of the chapter were originally published in [Sch13].

The following sections describe the architecture of the iterative decoder. It will be pointed out what comprises the extrinsic information that is passed between the decoder stages, and how it is calculated. Lastly, the decoder is analyzed using monte carlo simulations of the bit-error rate. In addition to that, the convergence behaviour of the decoder is analyzed, using the Extrinsic Information Transfer charts (EXIT charts) of the decoder [tB01, tB99].

4.4.2 Decoder architecture and extrinsic information

The decoder architecture presented in this section is universally applicable for systems with a serial concatenation of codes. Its basic structure originates from the decoding scheme proposed in [BM96, Fig. 2b], which was also used in many other works, e.g, for turbo equalization [KST04], interference cancellation [WP99], and demodulation [HL99].

Figure 4.11 shows the system model consisting of the concatenated coders and the corresponding iterative decoder. The information sequence consisting of the binary values v_j is encoded with a convolutional coder to obtain the

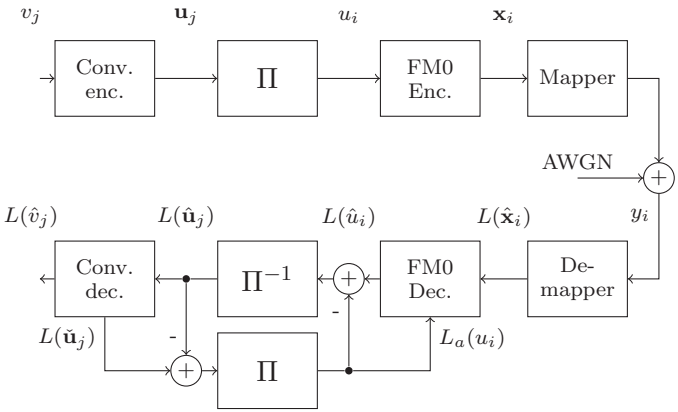


Figure 4.11: Coder and recursive decoder architecture, reprinted from [Sch13], © 2013 IEEE

code symbols $\mathbf{u}_j$, which consist of $\frac{1}{R_c}$ bits each, where R_c is the rate of the outer coder. The code symbols are serialized and passed through a bit-wise interleaver. The interleaved code bits u_i are then coded by the baseband coder (FM0 or baseband Miller), and the symbols $\mathbf{x}_i$ of the baseband code are mapped to waveforms (e.g. BPSK symbols) for the transmissions over the channel. Note that the indexes i and j differ beyond the permutation in the interleaver. The index j enumerates code *symbols*, i enumerates code *bits*. I.e., if there are J information bits overall, an $R_c = 1/n$ convolutional coder generates J channel code symbols $u_1 \dots u_J$, which consist of n bits each, $I = n \cdot J$ bits overall. The interleaver does bit-wise interleaving; u_i does refer to any of interleaved code *bits* $u_1 \dots u_I$. The index i marks the index of a trellis segment at the inner (de)coder, whereas j marks the index of the trellis segments in the outer (de)coder. Furthermore, $u_j^{(m)}$ shall denote the m^{th} bit of the j^{th} channel code symbol.

At the receiver, the demapper calculates the LLR $L(\hat{\mathbf{x}}_i)$ for the received waveforms y_i. Note that the FM0, or Miller symbol, consists of two LLRs; one for the first half, and the second half of the symbol. Both LLRs are passed as $L(\hat{\mathbf{x}}_i)$ to the baseband decoder, which carries out the APP algorithm using the trellis representation of the corresponding baseband code (Figures 4.5d &

4.6b). The baseband decoder calculates the a posteriori LLRs $L(\hat{u}_i)$ for the channel-code bits. The calculation consists in determining the branch metrics for every trellis transition, and carrying out the forward and backward recursions [BCJR74]

$$\begin{aligned} L(\hat{u}_i) &= \log \frac{\sum_{\mathcal{U}_i^{(+1)}} \prod_{j=k}^{n} P(u_k, y_k)}{\sum_{\mathcal{U}_i^{(-1)}} \prod_{j=k}^{n} P(u_k, y_k)} \\ &= \log \frac{P(u_i = +1) \sum_{\mathcal{U}_i^{(+1)}} P(y_i | u_i = +1) \prod_{\substack{k=1 \\ k \neq i}}^{n} P(u_k, y_k)}{P(u_i = -1) \sum_{\mathcal{U}_i^{(-1)}} P(y_i | u_i = -1) \prod_{\substack{k=1 \\ k \neq i}}^{n} P(u_k, y_k)} \\ &= L_a(u_i) + \underbrace{\log \frac{\sum_{\mathcal{U}_i^{(+1)}} P(y_i | u_i = +1) \prod_{\substack{k=1 \\ k \neq i}}^{n} P(u_k, y_k)}{\sum_{\mathcal{U}_i^{(-1)}} P(y_i | u_i = -1) \prod_{\substack{k=1 \\ k \neq i}}^{n} P(u_k, y_k)}}_{L_{in}(\hat{u}_i)} . \end{aligned} \tag{4.21}$$

So, the i^{th} a posteriori LLR depends on the conditional probability $P(y_i \mid u_i)$, where y_i is the channel observation, and $u_i \in \{-1, +1\}$ are to possible code-bit values. Furthermore, the LLR depends on the remaining sequence through the forward and backward recursion values $\alpha_{i-1}(\sigma_i)$ and $\beta_i(\sigma_i)$ (Not shown in (4.21)), and on the a priori LLR for the i^{th} channel-code bit $L_a(u_i) = \log \frac{P(u_i=+1)}{P(u_i=-1)}$. At the first decoding attempt no a priori knowledge is available, i.e., the value of $L_a(u_i)$ is zero.

The a posteriori LLRs from (4.21) are passed through the deinterleaver and to the outer decoder. Note that each $L(\hat{u}_i)$ depends on the channel observation $L(\hat{\mathbf{x}}_i)$, on the decoding result of the outer decoder through the a priori values $L_a(u_i)$ and, in an indirect way, on all other channel observations and a priori LLRs through the forward and backward recursions, i.e. the inherent dependency of code symbols that can be represented by the code trellis. The a priori information must not be fed to the outer decoder, since it does not comprise new, statistically independent information [Hag97]. It is removed by subtracting $L_a(u_i)$ from the output of the inner decoder. This leaves behind L_{in}, which is then deinterleaved, grouped and passed to the outer decoder as $L(\hat{\mathbf{u}}_j)$. The outer decoder is a modified APP decoder, as proposed by [BDMP97]. The decoder does not only calculate the a posteriori LLRs for the information bits $L(\hat{v}_j)$, but the a posteriori LLRs for the code symbols $L(\check{\mathbf{u}}_j)$ as well, which are

n values for the $1/n$ convolutional code. Let $L\left(\check{u}_j^{(m)}\right)$ be the a posteriori LLR of the m^{th} bit in the j^{th} code symbol. The a posteriori LLR can then be calculated from the forward and backward recursions, and the transition probabilities

$$L\left(\check{u}_j^{(m)}\right) = \log \frac{\sum\limits_{\sigma_j}\sum\limits_{\sigma_{j+1}} \gamma_0^{(m)}\left(u_j, \sigma_j, \sigma_{j+1}\right) \cdot \alpha_{j-1}\left(\sigma_{j-1}\right) \cdot \beta_j\left(\sigma_j\right)}{\sum\limits_{\sigma_j}\sum\limits_{\sigma_{j+1}} \gamma_1^{(m)}\left(u_j, \sigma_j, \sigma_{j+1}\right) \cdot \alpha_{j-1}\left(\sigma_{j-1}\right) \cdot \beta_j\left(\sigma_j\right)}$$

$$= \underbrace{\log \frac{p\left(u_j^{(m)}=0\right)}{p\left(u_j^{(m)}=1\right)}}_{L_{in}\left(\hat{u}_j^{(m)}\right)} + \underbrace{\log \frac{\sum\limits_{\sigma_j}\sum\limits_{\sigma_{j+1}} \cdot\alpha_{j-1}\left(\sigma_{j-1}\right) \cdot g\left(u_j^{(m)}=0, \sigma_j, \sigma_{j+1}\right) \cdot \beta_j\left(\sigma_j\right)}{\sum\limits_{\sigma_j}\sum\limits_{\sigma_{j+1}} \cdot\alpha_{j-1}\left(\sigma_{j-1}\right) \cdot g\left(u_j^{(m)}=1, \sigma_j, \sigma_{j+1}\right) \cdot \beta_j\left(\sigma_j\right)}}_{L_{ext,\,outer}} \tag{4.22}$$

(4.22) is very similar to the calculation of the a posteriori LLRs of the information bits, with the distinction that $\gamma_0^{(m)}\left(\mathbf{u}_j, \sigma_j, \sigma_{j+1}\right)$ and $\gamma_1^{(m)}\left(\mathbf{u}_j, \sigma_j, \sigma_{j+1}\right)$ mark the transition probability for a trellis branch that is related to a code symbol, in which the m^{th} *code* bit is zero or one, respectively. The gamma values depend on the output of the inner decoder and on the trellis only. No a priori information on v_j is available to the outer decoder. So the gamma values can be written as

$$\gamma_z^{(m)}\left(u_j, \sigma_j, \sigma_{j+1}\right) = p\left(u_j^{(m)}=z\right) g\left(u_j^{(m)}=z, \sigma_j, \sigma_{j+1}\right) \tag{4.23}$$

where $p\left(u_j^{(m)}=z\right)$ is the probability for the m^{th} bit in the j-th code symbol having the value z, which is exactly the APP from the inner decoder. The function $g\left(u_j^{(m)=z}, \sigma_j, \sigma_{j+1}\right)$ is either one, if there exists a branch in the trellis from state σ_j to σ_{j+1} that produces the value z at the m^{th} output bit of the coder, or zero otherwise. The probability $p\left(u_j^{(m)}=0\right)$ is a common factor in the numerator of (4.22), and $p\left(u_j^{(m)}=1\right)$ in the denominator. Together, they

make up $L_{in}\left(\hat{u}_j^{(m)}\right)$ on the left-hand side of (4.22). The right hand summand of (4.22) makes up the extrinsic information of the outer decoder that is fed back to the inner decoder.

It is noteworthy that the extrinsic information and decoder input are mutually independent, i.e., they can be factorized in the numerators and denominators, and written as a separate summands in (4.21) and (4.22). However, the independence assumption is only true for the first decoding iteration. The decoder output of the inner decoder is used in the calculation of the extrinsic information in the outer decoder, which is then fed back in the next iteration. Clearly, the value $L_{ext,outer}\left(u_j^{(m)}\right) = L\left(\check{u}_j^{(m)}\right)$, which becomes $L_a\left(u_i\right)$ after the interleaver in the feedback loop, depends of $\mathbf{u}_{j-1}$, $\mathbf{u}_{j+1}$, etc., through the forward and backward recursions. This dependence becomes negligible when two trellis segments are far apart. The interleaver serves to mitigate the effects of statistical dependence [tB01], by randomizing the positions of the dependent symbols before they are used in the next decoder stage, assuring a quasi independence.

4.4.3 Implementation and simulation results

The iterative decoder described previously, was implemented and tested using MATLAB and the *Coded Modulation Library (CML)*[2]. The CML provides MATLAB functions for the encoding and decoding of certain channel codes, including convolutional codes, functions for generating standard, and random interleavers and auxiliary function for calculating LLRs, etc. Performance-critical parts of the CML are implemented in C. Fortunately; the CML is published under the GNU Lesser General Public License, so the source code is available for adaptation. The library includes an implementation of the SISO APP decoder, which was used to realise the iterative decoder. The CLM's implementation includes the Log-MAP and Max-Log-MAP variants of the APP decoder. For the simulations, the Log-MAP variant [RVH95] was used, which gives equal results as the original APP algorithm from [BCJR74], but is numerically stable and better suited for implementation.

The implementation of the concatenated coder and the iterative decoder is straightforward, as shown in Figure 4.11. The outer code is a convolutional code. Figure 4.12 shows the used coder, which is an OFD code with the generator

[2] `http://www.iterativesolutions.com/Matlab.htm`, last checked: 2014-11-03

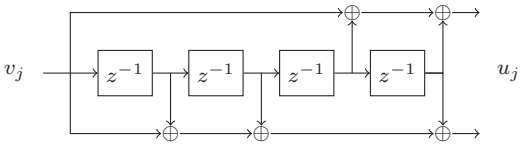


Figure 4.12: Convolutional coder used in the simulations

polynomials [23, 35]. The coder output is serialized and passed through a bit-wise interleaver. The interleaved values are then encoded according to the rules of the baseband code. Both, FM0 and baseband Miller, are considered in the simulations. The binary output of the baseband coder is then mapped to BPSK symbols. White Gaussian Noise (WGN) is added to the symbols, where the noise amplitudes are Gaussian distributed with a variance of $\sigma^2 = N_0/2$. Each baseband symbol is comprised of two samples, and E_s is normalized. The noise power is then chosen so that the bit energy remains constant for codes with different code rates R_c

$$\frac{E_s}{N_0} = R_c \frac{1}{2} \frac{E_b}{N_0} \tag{4.24}$$

The noise variance of the AWGN channel has to be known to the receiver in order to calculate the LLRs in the soft demapper [Bos98]

$$L\left(\hat{\mathbf{x}}_i\right) = \frac{2}{\sigma^2} \cdot y_i = L_{ch} \cdot y_i \tag{4.25}$$

where L_{ch}, the channel reliability, determines how channel information and a priori information are weighted in the inner decoder.

The channel LLR are then passed to the iterative decoder, where the decoding takes place in the following steps.

1. The channel information $L\left(\hat{\mathbf{x}}_i\right)$ is received from the demapper.

2. The inner decoder calculates $L\left(\hat{u}_i\right)$ from the channel observations and the a priori information (4.21). During the first iteration, there is no a priori information, and the a priori input is $L_a\left(u_i\right) = 0$.

3. The a priori values $L_a\left(u_i\right)$ are subtracted from the a posteriori LLRs $L\left(\hat{u}_i\right)$ to obtain the extrinsic information $L_{in}\left(u_i\right)$.

4. The values $L_{in}(u_i)$ are deinterleaved and passed to the outer decoder, $L(\mathbf{u}_j)$ being the LLRs for the j-th outer code symbol.

5. The outer decoder calculates the a posteriori LLRs for the information bits $L(\hat{v}_j)$, and the a posteriori LLRs for the code symbols $L(\check{\mathbf{u}}_j)$ (4.22).

6. The extrinsic information is extracted from $L(\check{\mathbf{u}}_j)$ by subtracting $L(\mathbf{u}_j)$.

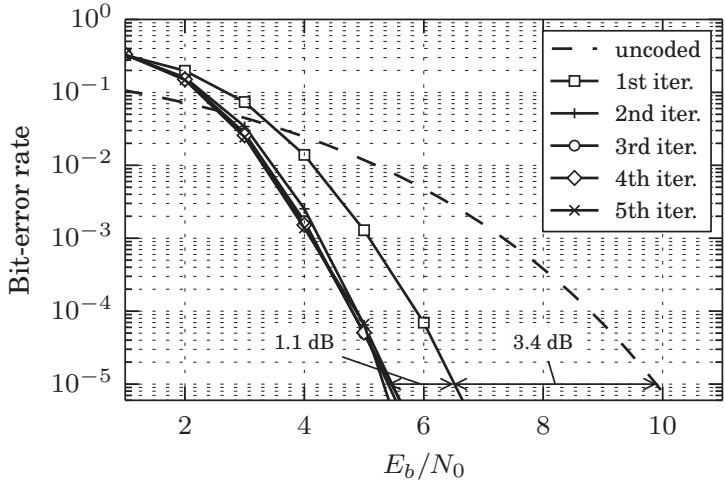
7. The extrinsic information of the outer decoder is serialized, passed through the interleaver, and feed to the inner decoder as a priori information $L_a(u_i)$.

8. The next decoding iteration starts at step 2 with $L(\hat{\mathbf{x}}_i)$ unchanged.

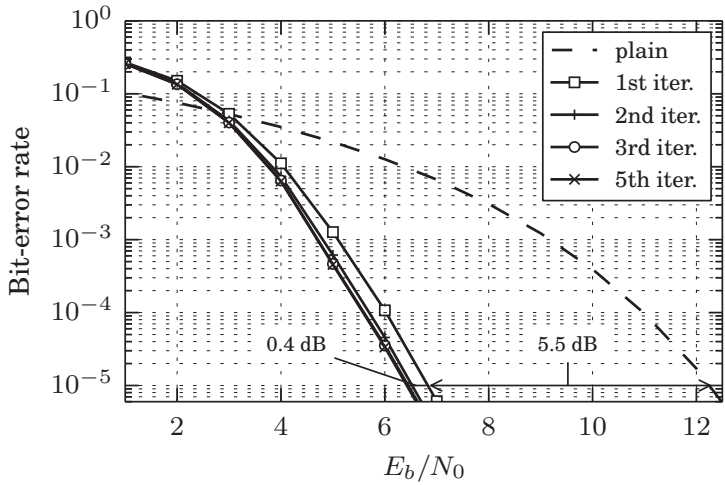
The iterative decoder may be stopped after a fixed number of iterations is reached, or if some other stopping criterion is fulfilled. In the given implementation, the decoding is stopped after a fixed number of iterations.

The results presented on the following pages were obtained using the outer code that is generated by the $R = 1/2$ convolutional coder with memory $\nu = 4$ and the generator polynomials $[23_8, 35_8]$ (see Figure 4.12). The simulation results presented in Figures 4.13a (originally [Sch13]) and 4.13b) are based on a message length of 100 bit with zero-tail termination. The zero-tail termination leads to a rate loss of ≈ 0.17 dB. On the decoder side, the number of decoding iterations is limited to five.

Figure 4.13a shows the simulation results for the variant with FM0 as the baseband (inner) code. The curves show the bit-error rate after the iterative decoder, plotted over E_b/N_0, ranging from one to ten decibel. Furthermore, the bit-error probability of the differential decoder (4.5) without an outer channel code is plotted as the reference. The bit-error rate after the first decoding iteration (□) (I.e., not information was fed back from the outer to the inner decoder yet) shows a coding gain of ≈ 3.4 dB at a BER of 10^{-5}. This is comparable to the results shown in Figure 4.10. Note that the results differ slightly, because different frame length and different interleavers were used. After the first feedback and the second decoding iteration (+), the BER improves at the output of the iterative decoder, resulting in approximately 1.1 dB additional coding gain. The successive decoding iterations 3-5 (∘,×,△) do not yield a significant improvement of the coding gain at a BER of 10^{-5}, though a minuscule improvement is visible at a higher BER of $\approx 10^{-3}$.



(a)

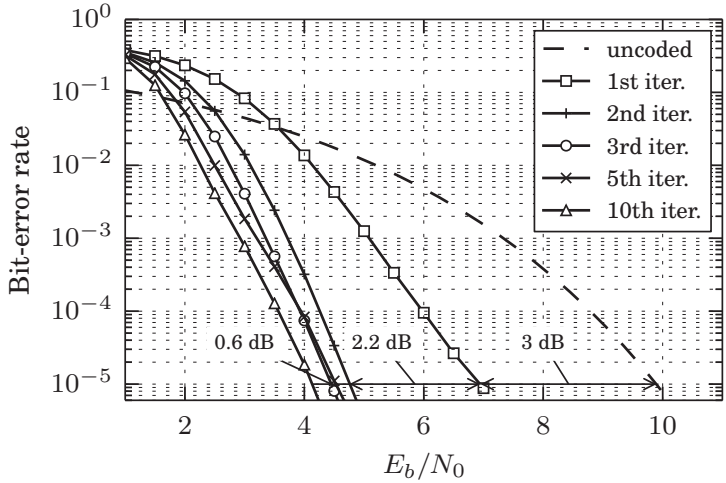


(b)

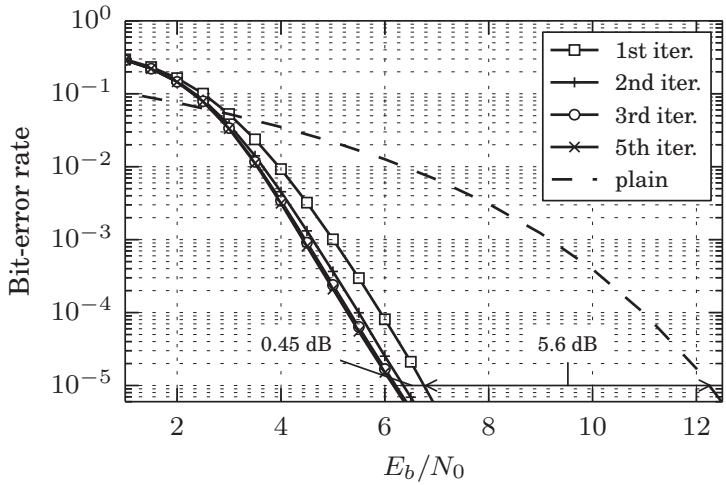
Figure 4.13: Simulation results for the bit-error rate after the iterative decoder for FM0 (a) and Miller coding (b), $K = 100$

The results of the simulation for the baseband Miller variant with equal parameters are shown in Figure 4.13b. Here, the dashed reference curve was obtained by simulation, and shows the BER of the trellis-based Miller decoder (see 4.2.3). For the variant with FEC, the coding gain at 10^{-5} is about 5.5 dB, compared to the reference curve. Note that, despite the greater coding gain, the Miller variant is still inferior to the FM0 variant with coding. In consecutive decoding iterations, the Miller variant does only improve slightly ≈ 0.4 dB. And while the baseband Miller variant has a greater coding gain of up to ≈ 5.9 dB at 10^{-5}, compared to ≈ 4.5 dB for FM0, it is still inferior to FM0, when comparing the BER.

For a message length of 100 information bits, the coding gain does not improved much after the second iteration, for either of the two baseband codes. Short message length, and thus short interleaver depth, are in general not desirable when turbo decoders are to be used. However, the short message length shown above, marks only one possible length of an *EPC* frame in a EPCc1g2 system. In general, the length of the *EPC* frames is potentially longer than that. An EPC length of up to 496 bit is permitted by the standard. The maximum case with an overall frame length of 528 bit is considered in the following simulation results. Figure 4.14a shows the results for the FM0 variant. At a BER of 10^{-5}, the coding gain after a maximum of ten iterations has now increased to 5.8 dB. After the first iteration, a gain of 3 dB is achieved; after the second iteration, the coding gain is increased by another 2.2 dB. With consecutive iterations, the BER is only slightly improved. The same can be observed of the baseband Miller variant (Fig. 4.14b), although to a smaller extend. With the longer block length, the coding gain with one iteration hardly changes, a second decoding iteration however improves the coding gain by 0.75 dB. The overall coding gain of 6.35 dB is only a slight improvement compared to the shorter message length.



(a)



(b)

Figure 4.14: Simulation results for the bit-error rate after the iterative decoder for FM0 (a) and baseband Miller (b), EPC496

4.4.4 EXIT chart analysis

The simulations demonstrate that the iterative decoder architecture provides a coding gain that is superior to the non-iterative decoding, when a concatenation of an outer convolutional and an inner baseband code (with memory) is used. To get a better understanding of the limitations of the proposed decoder architecture, the decoder was further analyzed based on its extrinsic information transfer characteristics. EXIT charts were originally proposed by [tB01, tB99], to analyze and visualize the convergence behaviour of iterative decoders. According to ten Brink, an iterative decoder can essentially be characterized by the information *transfer characteristics* of the individual constituent decoders. The transfer characteristics of a decoder expresses the relation of the information at the output of the decoder to the information at the input of the decoder. More specifically, the information is measured as the mutual information of X and U [tB01, eq. (12)]

$$I_{e,inner}(X;U) = \frac{1}{2} \sum_{x=\{-1,1\}} \int_{-\infty}^{\infty} p_U(\xi \mid X = x) \cdot \log_2 \frac{2p_U(\xi, X = x)}{p_U(\xi \mid X = 1) + p_U(\xi \mid X = -1)} d\xi \qquad (4.26)$$

where X is a binary random variable denoting the transmitted binary values of the baseband code x_i, and U is a continuous random variable denoting the extrinsic LLRs $L_{in}(\hat{u}_i)$ at the output of the inner decoder. (4.26) corresponds to the general definition of the mutual information [PS08]. Evaluating the equation is usually not possible, because the PDFs $p_U(\xi \mid X)$ and $p_U(\xi, X)$ are unknown and depend on the mutual information $I_{a,inner}$ of the a priori LLR, as well as the channel E_b/N_0. Although solving (4.26) in closed form is usually not feasible, it can be evaluated when the PDFs are obtained by monte carlo simulation.

For the inner decoder, the transfer characteristics are obtained by feeding the decoder with an output sequence of the AWGN channel at a fixed E_b/N_0. Further, the decoder is fed with a priori LLRs that share the mutual information $I_{a,inner}$ with the (known) transmission sequence. After the decoding has been carried out, $I_{e,inner}$, the mutual information of the transmit sequence and the extrinsic output, can be measured. By varying $I_{a,inner}$ between zero and one, and recording $I_{e,inner}$, the transfer characteristic can be obtained.

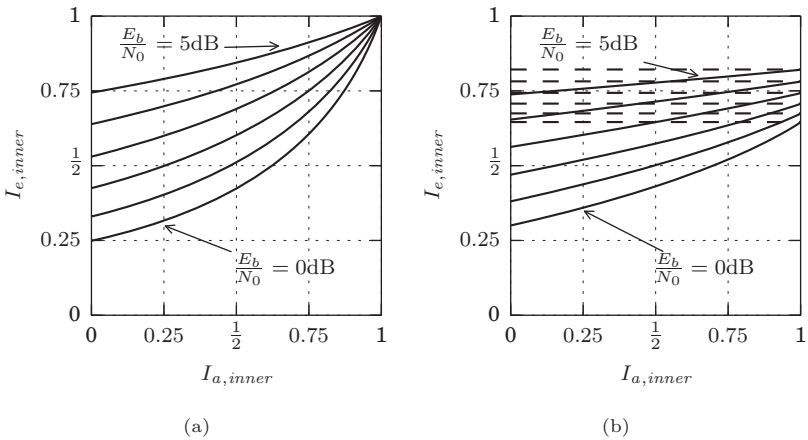


Figure 4.15: Transfer characteristics for the FM0 (a) and Miller (b) baseband decoder.

Figure 4.15 shows the transfer characteristics for both variants of inner decoders. The curves were obtained by simulation, using a set of scripts provided by [Mau08]. Each line in the plots shows the transfer characteristics of the decoder for a different channel SNR, between 0 dB to 5 dB. It can clearly be seen that for incomplete a priori information ($I_{a,inner} < 1$), the mutual information of the extrinsic output and the transmitted data, depends on the channel SNR. When the mutual information at the a priori input increases, so does the mutual information at the extrinsic output. For FM0 (Figure 4.15a), the mutual information I_e of the extrinsic information is one, when $I_a \to 1$, i.e., perfect a priori information is available. This means for FM0, that the decision at the i-th position is determined fully by the a priori information of the surrounding symbols. Not so for the Miller code variant (Figure 4.15b).

The apparent difference between FM0 and baseband Miller is that, for the Miller decoder, high mutual information of the a priori input $I_{a,innner} \to 1$ will not lead to $I_{e,innner} \to 1$ at the extrinsic output of the decoder. To understand this, consider the role that the a priori information plays in the APP decoder. Suppose the a priori information is perfect ($I_{a,innner} \to 1$), and that the decoder,

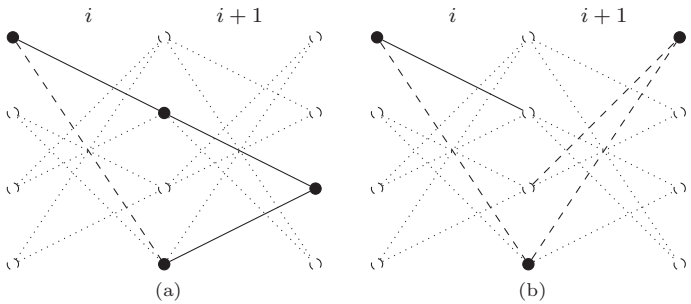


Figure 4.16: Trellis section of the Miller decoder.

therefore, knows that the coder was in state σ_{i-1} after the $i-1$-st symbol, and in state σ_{i+1} after the $i+1$-st symbol. The extrinsic decoder output at instant i then depends on all remaining paths (with non-zero probability) that start in state σ_{i-1} and end in σ_{i+1}. These paths are shown solid in the trellis in Figure 4.16. In the first example (a) σ_{i-1} is 0, and σ_{i+1} is 2. The backward recursion assigns equal probabilities to the states $\sigma_i = 1$ and $\sigma_i = 3$, because both branch metrics are equal[3] at $i+1$.

$$
\begin{aligned}
L\left(\hat{u}_i\right) &= \log \frac{\alpha_{i-1}(0) p(y_i|x_i = s_1) \cdot P(x_1 = -1)\beta_i(1)}{\alpha_{i-1}(0) p(y_i|x_i = s_3) \cdot P(x_1 = 1)\beta_i(3)} \\
&= \underbrace{\overbrace{\log \frac{\alpha_{i-1}(0)}{\alpha_{i-1}(0)}}^{0} + \log \frac{p(y_i|x_i = s_1)}{p(y_i|x_i = s_3)} + \overbrace{\log \frac{\beta_i(1)}{\beta_i(3)}}^{0}}_{L_{in}} + \log \frac{P(x_1 = -1)}{P(x_1 = 1)}. \quad (4.27)
\end{aligned}
$$

Since only two transitions (with non-zeros probability) are left, and both transitions originate in the same state and end in states with equal probabilities, the values of the forward and backward recursions cancel out for this case in (4.27). The extrinsic LLR at i then only depends on the channel observation.

In the second case (b), suppose only $\sigma_{i-1} = 0$ and $\sigma_{i+1} = 0$ have non-zero probabilities, i.e. $\alpha_{i-1}(0) = 1$ and $\beta_{i+1}(3) = 1$. The backward recursion assigns equal probabilities to the states $\sigma_i = 2$ and $\sigma_i = 3$. But in this case, there is

[3] Incoming transitions to the same state are associated to the same input bit and the same output symbol for baseband Miller

only one transition left that has a non-zero probability; the one originating in $\sigma_{i-1} = 0$, going into $\sigma_i = 3$. The extrinsic LLR at i is then

$$\begin{aligned} L(\hat{u}_i) &= \log \frac{\alpha_{i-1}(0) p(y_i|x_i = s_1) \cdot P(x_1 = -1)\beta_i(1)}{\alpha_{i-1}(0) p(y_i|x_i = s_3) \cdot P(x_1 = 1)\beta_i(3)} \\ &= \underbrace{\overbrace{\log \frac{\alpha_{i-1}(0)}{\alpha_{i-1}(0)}}^{0} + \log \frac{p(y_i|x_i = s_1)}{p(y_i|x_i = s_3)} + \overbrace{\log \frac{0}{\beta_i(3)}}^{-\infty}}_{L_{in} = -\infty} + \log \frac{P(x_1 = -1)}{P(x_1 = 1)}. \end{aligned} \quad (4.28)$$

In this case the extrinsic output is fully determined by the (perfect) a priori information.

In total, there are 16 different cases for the four different originating states $\sigma_{i-1} = 0 \ldots 3$ and the two following coder input bits. In eight of this, cases the extrinsic output LLR is perfectly given by the priori information; in the other eight cases, the extrinsic LLR depends on the channel observation. So, in half of the cases, the mutual information of the extrinsic output values and baseband Miller coder input values is one; in the other eight cases it is equal to the mutual information of the channel input (x_i) and the channel output (y_i), i.e. the channel capacity. When all sixteen cases occur with equal probability, then the average mutual information must be the arithmetic mean of *one* and the (AWGN) channel capacity

$$I_{e,innner}^{max} = \frac{1}{2}\left(1 + C_{\mathrm{BPSK}}\right). \quad (4.29)$$

C_{BPSK} is given e.g. in [PS08, (6.5-31)]. The result given in (4.29) is consistent with the simulation results, as shown in Figure 4.15. The maximum values of the mutual information $I_{a,inner}$ are shown dashed for each E_b/N_0 in the figure.

To obtain the full EXIT chart of the iterative decoder, the transfer characteristics of the outer decoder are required as well. The transfer characteristics are obtained in the same manner as for the inner decoder, with the distinction that there is no channel input for the outer decoder in a serial concatenated system. The mutual information at the extrinsic output of the outer decoder depends only on the (extrinsic) output of the inner decoder. Figure 4.17a shows the transfer characteristics that were obtained by simulation. In the plot, the roles of the abscissa and the ordinate have been swapped with respect to Figure 4.15.

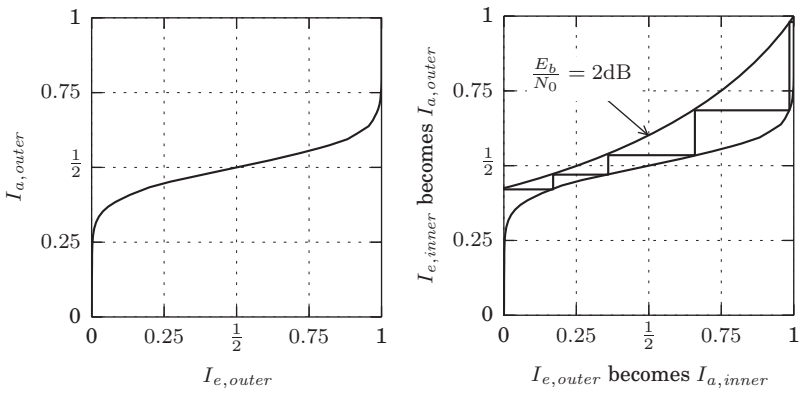


Figure 4.17: Inverse transfer characteristics for the outer APP decoder of the $R = \frac{1}{2}$ convolutional code with the generator polynomials $[23_8, 35_8]$ (a), and EXIT chart of the iterative decoder, with FM0 as the inner code (b).

Together, the transfer characteristics of the inner decoder and the inverted transfer characteristics of the outer decoder, form the EXIT chart as shown in Figure 4.17b. The example shows the predicted convergence behaviour of the decoder for $E_b/N_0 = 2\,\mathrm{dB}$. The figure can be interpreted in the following way: During the first decoding iteration, the inner decoder has no a priori information available, thus the average mutual information at the extrinsic output depends only on the channel SNR. For the given example it is about $I_{e,inner} \approx 0.42$. The extrinsic values of the inner decoder are then fed to the outer decoder. After decoding the convolutional code, the mutual information of the extrinsic values of the outer decoder is expected to be $I_{e,outer} \approx 0.17$. In the next decoding iteration, the inner decoder then carries out the decoding again. This time, with the extrinsic output values of the outer decoder as a priori input. Given the additional information, the decoding result is expected to improve (the BER decreases). In the example, the mutual information at the output of each of the decoders increases with every successive iteration, up to the point where

$I_{e,inner} = I_{e,outer} = 1$. So, it is expected that the data will be fully recovered (i.e., without errors, though the BER simulation shows that this point is not reached).

Figure 4.18 shows two cases for the convergence behaviour of the iterative decoder when baseband Miller is used as the inner (baseband) code. The first chart Fig. 4.18a shows the so-called pinch-off case. After the first iteration, the increase of mutual information is not high enough to significantly affect the decoding result in the following iteration. The decoder gets *stuck*, and the BER does not improve with successive iterations. Figure 4.18b shows a case with a slightly higher channel SNR. Here, there is an opening between the curves, and the trajectory shows an increase of the mutual information over the course of about four iterations. However, the mutual information of the extrinsic output of the inner decoder is not expected to exceed the limit as given in (4.29).

So, in how far can the EXIT chart analysis give an insight to understand the results of the BER simulations? While the EXIT charts of both iterative decoders, for baseband Miller and FM0, suggest that about four to five decoding iterations will be required to achieve optimum results, the BER simulations suggest that two iterations are sufficient. More than two iterations do not significantly improve the decoding result. This is however, not a contradiction to the results of the EXIT chart analysis. The EXIT chart also reveals that more than two iterations are only beneficial in a very limited SNR range. The Figures 4.18c and 4.18d show the EXIT charts for an $E_b/N_0 = 4\,\mathrm{dB}$, for FM0 and Miller, respectively. The decoding trajectories show that for high SNR, two decoding iteration are sufficient to get good results. Especially the simulation results shown in Figure 4.14a coincide with this prediction. In the range from $E_b/N_0 = 2\,\mathrm{dB}$ up to $4\,\mathrm{dB}$, the BER improves with each additional iteration. For E_b/N_0 values higher than that, there is hardly any gain for more than two decoding iterations. For FM0 with a short block length (Figure. 4.13a), an increase of the coding gain is barely detectable at all, when more than two decoding iterations are carried out. Most probably this is due to the fact that short interleaver length are generally not favourable for iterative decoding, because the assumption of statistical independence [Hag97] of the a priori values and the extrinsic information is not valid after a few iterations.

In the case of baseband Miller, the EXIT chart analysis brought the authors attention to the limited usefulness of a priori information during the iterative

decoding process. As explained above, inherent structure of the Miller code limits the influence of (even perfectly known) a priori values in the APP decoding of baseband Miller.

4.4.5 Evaluation of the iterative decoder

Observing the results of the previous two sections, it can be concluded that an iterative approach for decoding the inner baseband code, and the outer channel code (if present), contributes an additional coding gain. At the same time, the added complexity is fairly low. The decoders first decoding (without a priori information) is generally carried out, in any concatenated system. Only one additional decoding iteration is required to achieve an added coding gain from 0.4 dB for baseband Miller up to about 2.2 dB for FM0.

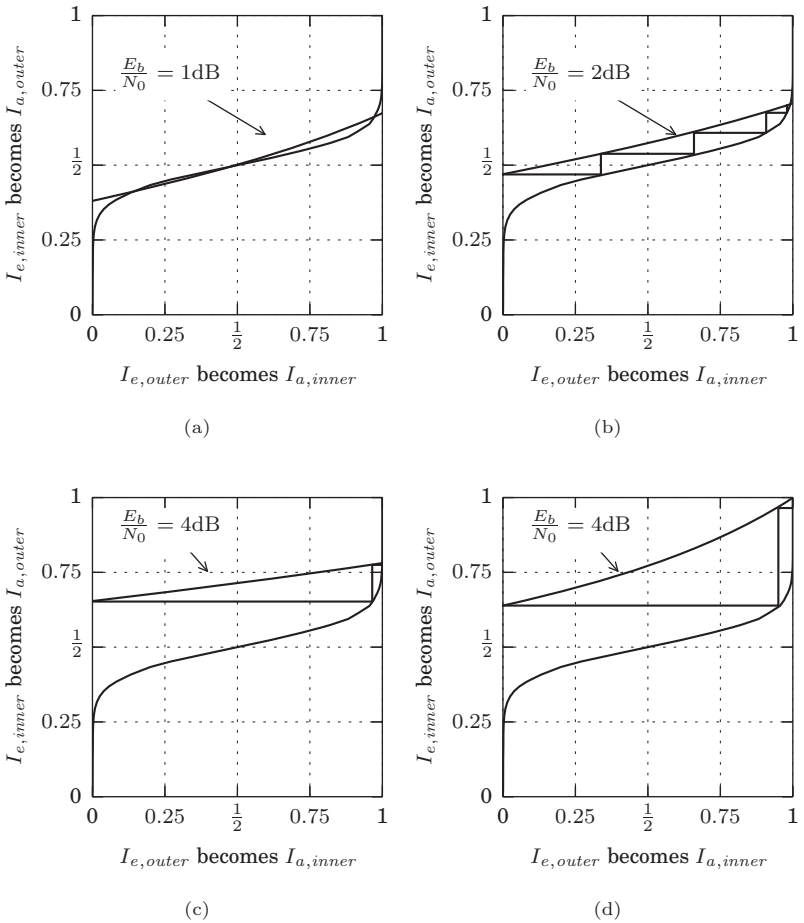


Figure 4.18: EXIT charts of the iterative decoder with baseband Miller as the inner code, $E_b/N_0 = 1\,\mathrm{dB}$ (a) and $E_b/N_0 = 2\,\mathrm{dB}$ (b), and EXIT charts for a higher SNR $E_b/N_0 = 4\,\mathrm{dB}$ for baseband Miller (c) and FM0 (d).

5 Retransmission schemes for EPCglobal systems

5.1 Retransmissions in the standard protocol

In the previous chapter it has been established that using FEC on the T-to-R link of a EPCc1g2 should be considered as an option to improve the duration of tag inventory process. Further, it has been shown how the decoding of the channel codes may be implemented in the reader, to yield good results when the EPCc1g2 standard baseband codes are used, which would be a necessity to retain backward compatibility with the current standard systems. This chapter deals with the integration of a coding scheme into the standard protocol, and with the error recovery and retransmission in general.

In Section 3.7.2 it has already been discussed that there are various ways in which transmission errors interfere with the identification of a tag during the inventory process. Errors in the early stages of the exchange, i.e., in the *RN16* response and *ACK* command, are not detectable and cannot be corrected. The *EPC* response of the tag, on the other hand, contains a CRC checksum that allows the reader to detect errors with a high probability. The EPCc1g2 standard leaves some liberties for the implementation of the protocol in the reader to recover from the second kind of errors. The reader may then abort the tag identification (Fig. 3.15a) or request a retransmission (Fig. 3.15b). In the first case, the tag identification is deferred to a subsequent inventory round; in the second case, the tag may be identified right away, if the retransmission is successful.

In the latter, the first variant is referred to as *deferred retransmission*, because the retransmission/identification is deferred to the next round(s); the second variant is referred to as *instant retransmission*, respectively. It might seem counterintuitive to distinguish between this two cases, because one might argue that the instant of the retransmission has no impact on the throughput. In the

case of the deferred retransmission however, a tag takes part in the subsequent inventory rounds, occupying a slot and potentially causing collisions. Also, for the deferred retransmission, the reader has no simple way of detecting whether a tag reply in a later inventory round is a retransmission of a previously failed tag reply or not. So, although both tag replies are the same, except for one or few errors, the reader cannot exploit that fact. For instant retransmissions, the reader can, however, exploit the fact that the retransmission is a copy of the first, erroneous tag reply. It may use a combining scheme to improve the chance of receiving the retransmission correctly.

5.2 Combining

When the reader prompts the tag for instant retransmissions, the tag sends an identical copy of the *EPC* frame every time (Fig. 3.15b). Instead of decoding the repeated copy independently, the reader can as well regard all transmissions as part of a time diversity scheme, and use a combining method to jointly decode multiple retransmissions. [Bre59] discusses the properties of different linear combining methods for diversity systems, which can be used to combine multiple received copies in the EPCc1g2 system as well. Maximum Ratio combining (MR combining) yields the best results in fading channels and is used in the latter, though it shall be noted that MR combining and equal gain combining yield the same results in the AWGN channel. It is also understood that diversity techniques are meant to lower the outage probability in fading channels, and no such effect exists in the AWGN channel model. However, combining n repetitions of a frame increases the SNR n-fold [Bre59]. When the received signal $s(t)$ has an average symbol energy E_s, and the amplitude of the data $E\left[|d(iT)|\right] = 1$ is unity, then

$$\frac{E\left[P_s\right]}{E\left[P_n\right]} = \frac{E\left[\left(\sum_n s(t)\right)^2\right]}{E\left[\left(\sum_n n(t)\right)^2\right]} = \frac{E\left\{\left[n\sqrt{E_s}d(iT)\right]^2\right\}}{n\sigma^2} = \frac{nE_s}{\sigma^2} = \frac{nE_s}{N_0} \tag{5.1}$$

where $n(t)$ is a complex noise process with real and imaginary parts that are normally distributed with variance $\sigma_r^2 = \sigma_i^2 = \frac{N_0}{2}$, after the matched filter and sampling. The combining can be regarded as repetition coding, so when normalizing with respect to the code rate $R_c = 1/n$, it is clear that there is no

coding gain. Nevertheless, combining multiple copies of the same frame is, of course, preferable to decoding them separately.

5.3 Hybrid ARQ

5.3.1 Introduction to Hybrid ARQ

Hybrid Automatic Repeat Request (HARQ) protocols are the combination of classical Automatic Repeat Request (ARQ) schemes and FEC, and are superior to systems that use regular ARQ, even when combining is used as proposed before. [CC84] proposed two basic variants of implementing HARQ. *Type I* systems combine an FEC code with error detection. The FEC code is designed with a fixed code rate, so that most of the transmission errors can be corrected, using the redundant information of the channel code. If the error correction capabilities of the code are exceeded, remaining errors can be detected with high probability, using a additional error-detecting code. The receiver can then request a retransmission of the coded data. *Type II* systems, unlike the previous variant, do not have a fixed code rate. The first transmission is sent encoded with a certain code rate. If the transmission contains errors, the frame is repeated, and the copies are decoded jointly using Chase Combining (CC) [Cha85], so this variant combines FEC with a rateless repetition coding. A second variant of type II HARQ uses retransmissions that are distinct and carry a different portion of the redundancy of an FEC code each time. Again, the receiver saves all transmissions, which when combined form a punctured codeword of the so-called mother code. After each retransmission, more parts of the codeword are available, i.e., the redundancy increases. Thus, this variant is called HARQ with Incremental Redundancy (IR). Finally, *type III* HARQ[Kal95] is similar to type II, with the additional requirement that each coded frame shall be invertible on its own, i.e., it shall be possible to decode any single transmission to recover the original data. The distinction between type II and type III is not always made. E.g., [CC84] already mentions the characteristic properties of type III in his description of type II HARQ.

5.3.2 RCPC Codes

Type I HARQ does not require a specific type of channel code to be used. Any FEC code with characteristics that matches the error characteristics of the channel can be used. Type II systems with CC do only differ from type I systems in the way in which the receiver side is implemented. HARQ with IR, however, requires codes that are designed to have good properties for every different code rate after puncturing. [Hag88] introduces such a class of codes called Rate Compatible Punctured Convolutional (RCPC) codes. A class of RCPC codes is created by puncturing a *mother* convolutional code of rate $R_c = \frac{k}{n}$, with a puncturing pattern of a certain length P, which is applied periodically. The punctured positions of every code in a class can be described by a set of $P-1$ matrices with $n \times P$ elements. Each line of the puncturing matrix contains the pattern for one of the n coder outputs, where a zero signifies a punctured code bit. A puncturing matrix contains between $P+1$ and $P \cdot n$ one entries that mark the non-punctured positions, leading to code rates from $\frac{P}{P+1}$ to $\frac{k}{n}$. A puncturing matrix of a low rate code must contain ones at all positions where the higher rate code of the same class does. From all patterns for each possible rate, the puncturing patterns are chosen that yield the best distance properties [Hag88, II.B]. Consider the example of a $R_c = \frac{1}{2}$ mother code with puncturing period $P = 4$ from [Hag88]:

$$\mathbf{a}(1) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}, \mathbf{a}(2) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$
$$\mathbf{a}(3) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \mathbf{a}(4) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \tag{5.2}$$

The four patterns produce codes with the rates: $R_{c1} = \frac{5}{8}$, $R_{c2} = \frac{3}{4}$, $R_{c3} = \frac{7}{8}$ and $R_{c4} = \frac{1}{2}$. An extended selection of RCPC codes with optimized distance profiles is given in [Hag88, Lee94].

The use of RCPC codes, in a type II HARQ system, has already been suggested as an application in Hagenauer's original publication. He describes an IR approach, where the first transmission contains the code sequence of the highest rate code of the RCPC family. Consecutive parity frames then contain the parity information that is required to complete a codeword of a lower rate

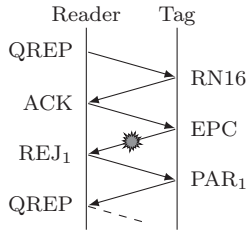


Figure 5.1: Proposed type II HARQ error recovery procedure

code from the same family. So with each transmission, the code rate reduces, and the redundancy increases, until the receiver has the complete code sequence of the mother code.

5.3.3 Extending the EPCglobal protocol by HARQ functionality

The EPCc1g2 standard protocol is not prepared for dealing with retransmissions during neither the identification nor the access phase. While retransmissions are possible, and the system behaves like a stop-and-wait ARQ system as shown in Section 3.7.2, there are no commands specified explicitly for this task (The *ACK* and *NAK* commands do not have the same meaning like in a regular ARQ system). In the HARQ scheme, the reader needs to be able to prompt the tag to transmit parity information, after the transmission of an *EPC* failed (see Figure 5.1). Reader and tag need to be in agreement about the set of parity information exchanged at each retransmission attempt. To keep the complexity of the tag low, i.e., keeping the number of state variables in the tag to a minimum, it is proposed that the reader shall explicitly request a particular parity set. For this purpose, the EPCc1g2 protocol should be extended by a REJ_n command, which the reader issues to a tag that has been previously singled out, i.e., that has successfully received the *ACK* command. The tag shall reply with REJ_n to this command by backscattering a parity frame PAR_n, when the tag is in the *Acknowleged*, *Open* or *Secured* state. The entire exchange of commands and responses is shown in Figure 5.1.

The parity information that the tag sends in the PAR_n reply shall be derived from the *EPC* data using an RCPC code, according to the scheme described in [Hag88]. Upon receiving the REJ_1 command, the tag shall encode the data,

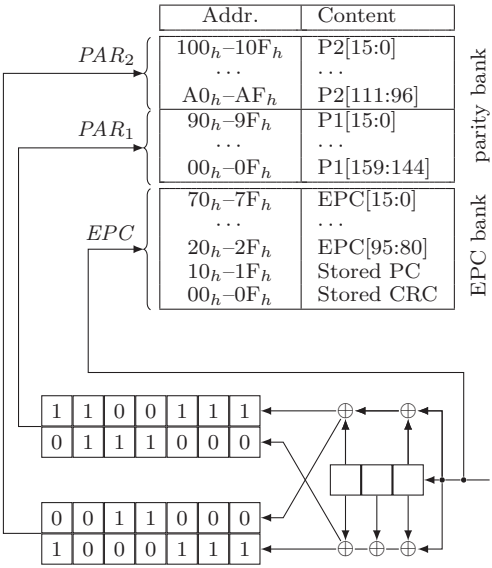


Figure 5.2: Memory scheme for storing precomputed parity information in the tag memory. From [SR13] © 2013 IEEE.

puncture the code sequence according to the puncturing scheme with the highest code rate R_{c1}, and send then non-punctured code bits in the PAR_1 reply. The reader can then decode the PAR_1 reply, using the previously received (erroneous) *EPC* reply as a priori information. If the CRC checksum still shows errors after decoding, the reader may request additional redundancy by issuing REJ_2. The tag shall reply by puncturing the code sequence with the second scheme from the RCPC code class, but it shall only include the parity bits in PAR_2 that were not previously sent in PAR_1. Together, PAR_1 and PAR_2, form a code sequence from the RCPC class with the rate R_{c2}. This scheme may continue until all parity from the RCPC code class has been sent. At this point, the reader has received the *EPC* reply and all REJ_n frames that make up a code sequence with the rate R_c, which is the rate of the non-punctured mother code of the RCPC class.

In order to keep the complexity of the tag low, neither the convolutional coding nor the puncturing need to actually be implemented in the tag chip. Instead,

this tasks may be offloaded to the reader. During the initial programming of the tag, the reader may pre-calculate the parity, i.e., do the convolutional coding, carry out the puncturing, serialize the data and store it in a dedicated memory bank in the tag. Upon receiving a PAR_n frame, the tag must then simply reply with the stored pre-calculated parity information. Figure 5.2 shows an example for storing the parity information. In the example, the parity information is generated using a $R_c = \frac{1}{2}$ convolutional coder, with the generator polynomials 15_8 and 17_8 as a mother code. The code sequence is then punctured with a puncturing period of $P = 7$ [Lee94]. The example does not use all codes out of the family of RCPC codes. The first parity reply, PAR_1, is a code sequence generated with the highest rate code $R_c = \frac{7}{8}$. In the second step, the remaining code bits of the mother code are transmitted in PAR_2.

5.4 System performance and simulations of the tag inventory

5.4.1 Implementation and analysis of the retransmission schemes

The different retransmission schemes, both standard compatible schemes without (Section 3.7.2) and with combining (Section 5.2), as well as the proposed HARQ extension, were evaluated by computer simulation [SR13]. The simulation program implements a major part of the EPCc1g2 protocol, with some simplifications, and models the entire identification process of a tag population of a given size by *one* reader. While running, the simulation fully records the type, number and length of all frames exchanges during the identification process, from which the overall identification time, as well as the average throughput per round, can be calculated.

The simulations include the FSA part of the tag inventory phase, that precedes the exchange of the *EPC* identifier. In this phase, it is assumed that the reader has complete knowledge of the tag population size, and adjusts the number of slots in the inventory round to the optimal values given in [BDVA11]. The capture effect is not considered, i.e., collisions of two or more tag replies in an inventory slot lead to the loss of all tag information, and the tags are not identified in the given round. For each singleton tag reply, the exchange of *RN16*, *ACK*, *EPC*, etc. is carried out according to the procedures described below. Four different retransmission strategies were compared: deferred retrans-

mission, instant retransmission, instant retransmission with packet combining, and the HARQ variant. In all variants, errors during the exchange of *RN16* and *ACK* cause the tag identification to fail. Otherwise, after a tag has successfully been acknowledged, it sends its *EPC* identifier. From there on, the four variants differ.

The flowcharts in the Figures 5.3 and 5.4 illustrate how exactly the retransmission schemes are implemented. Each flowchart shows the implemented behavior of the reader within one slot, starting with the transmission of the *ACK* command by the reader. With the deferred retransmission scheme (Figure 5.3a), the reader only sends one *ACK* to request a tag's *EPC* reply during one round. If the attempt fails, the tag remains unidentified in this particular inventory round. The instant retransmission scheme (Figure 5.3b) makes multiple $(a+1)$ attempts to read a tag that replies in a slot, treating each tag reply separately. The variant with shown in Figure 5.4a also makes up to $a+1$ attempts to read a tag, however, it uses combining, and exploits all the copies of the tag reply received up to the given instant. The HARQ (Figure 5.4b) variant is a little more complex than the previous schemes. If the first copy of the *EPC* is received with errors, the reader requests the parity information, which consists of the punctured code symbols of RCPC code. If all parity information has been requested and the decoding still does not succeed, the reader requests the *EPC* and the PAR_n information again, and uses diversity combining on the received copies. Note that in Figure 5.4b, some details, like the calculation of the buffer positions, are omitted. The decision whether there is *"parity left"* that can be requested, depends on the number of attempts and the puncturing scheme.

The four proposed schemes are ultimately compared by monte carlo simulation. However, a prediction of the expected number of transmissions that each scheme requires, can also be obtained analytically for the first tree variants. It is assumed here that the R-to-T channel is error-free, as shown in Figure 5.5, i.e., the error-probability is negligibly small compared to the T-to-R channel. The T-to-R channel is modeled as an AWGN channel. For the deferred, and instant retransmission schemes, the bit-error probability p_b does not depend on the number of attempts m, so the frame error probability for a frame with length l is

$$p_f = 1 - (1 - p_b)^l \tag{5.3}$$

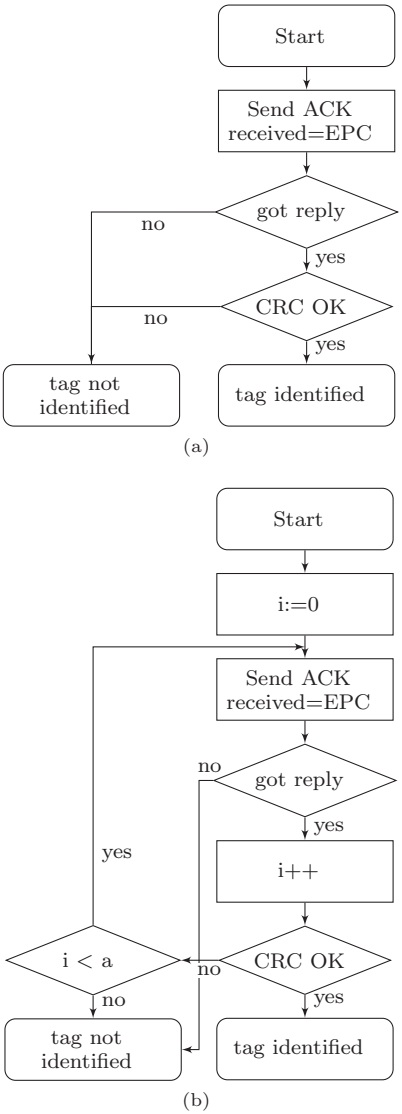
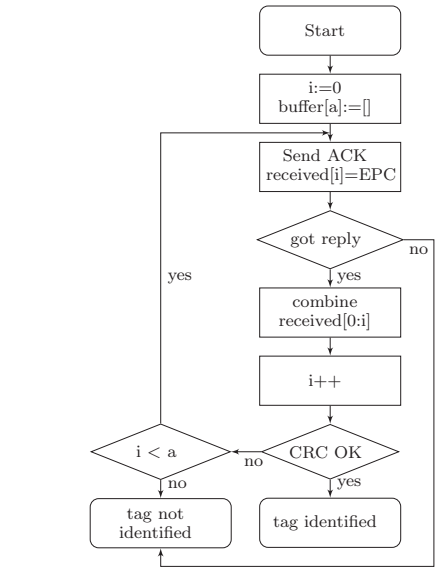
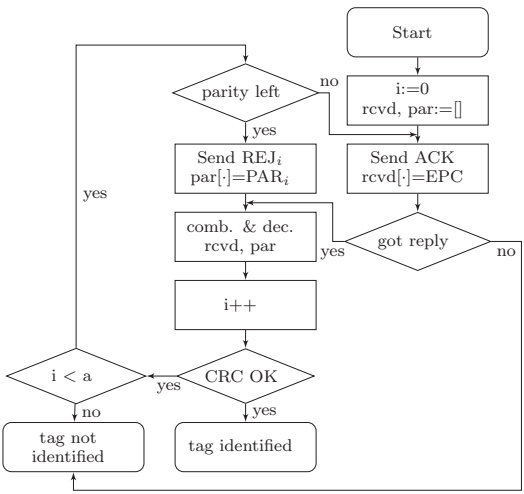


Figure 5.3: Flowcharts for the deferred (a) and instant (b) retransmission scheme as implemented for the simulation.



(a)



(b)

Figure 5.4: Flowcharts for the combining (a) and HARQ (b) retransmission scheme as implemented for the simulation.

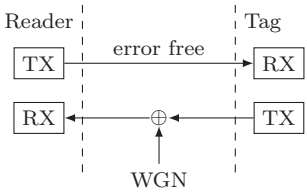


Figure 5.5: Simplified channel model used in the evaluation of the different retransmission schemes. From [SR13] © 2013 IEEE.

In general, the bit-error probability for the T-to-R channel is determined by the noise variance of the channel, and depends further on the number of transmissions, when combining or HARQ is used. For the combining and the HARQ variant, the frame-error probability decreases with every consecutive transmission. Let $p_{b,m}(m)$ be the bit-error probability after combining or jointly decoding m transmissions. Then the frame-error probability depends on m as well

$$p_{f,m}(m) = 1 - (1 - p_{b,m}(m))^l \,. \tag{5.4}$$

Now, let M be a random variable describing the number of transmissions that are required to receive the frame without errors. The PMF of M is then a geometric distribution for the deferred and instant retransmission

$$p_g(M = m) = (1 - p_f) \cdot p_f^{m-1}. \tag{5.5}$$

For the combining variant, the frame-error probability changes with m, so the probability of having $m - 1$ failed attempts can be written as a product of the $m - 1$ distinct frame-error probabilities

$$p_g(M = m) = (1 - p_{f,m}(m)) \prod_{i=1}^{m-1} p_{f,m}(m). \tag{5.6}$$

The expected value of M for the first three cases can be calculated from (5.5) and (5.6)

$$E\,[M] = \sum_{j=1}^{\infty} j \cdot p_g\,(M = j)\,. \tag{5.7}$$

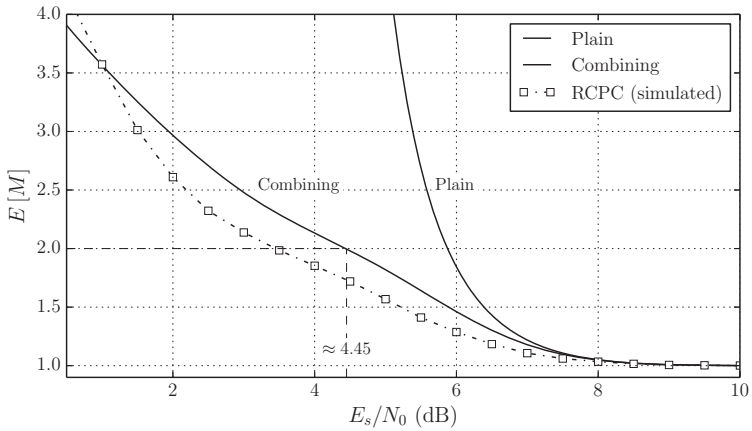


Figure 5.6: Expected number of transmissions until the tag reply is received error-free for the different retransmission strategies. From [SR13] © 2013 IEEE.

(5.7) can now be evaluated for the transmission over an AWGN channel. In that case, the bit-error probability for a certain SNR is known, and it is further known that the SNR is improved m-fold when combining m retransmission of the same frame (5.1).

Figure 5.6 shows the expected number of transmissions according to (5.7) for the different cases. The results are plotted against the symbol-energy-to-noise ratio in the T-to-R channel. Further, the assumption is made that the R-to-T channel is error-free, and that the noise in the T-to-R channel can be described by the AWGN model (Figure 5.5). Note that, the expected number of transmissions is identical for the deferred and the instant retransmission scheme, however, the identification time differs, as will be shown in the simulations. As expected, the combining variant requires less retransmission compared to the deferred and instant retransmission schemes. The number of retransmissions for the RCPC type II HARQ variant cannot be calculated directly, since the frame-error probability is not available in close-form. Though there are some error-bounds given in [Hag88], the results for the HARQ variant, shown dashed in Figure 5.6, were obtained by simulation. It is clear from this results, that

the HARQ variant outperforms the combining variant in terms of the required transmission attempts. However, the length of the frames differs for the combining and HARQ, which should be taken into account as well.

A more detailed analysis was obtained by simulating the identification time for an entire tag population. The simulation starts with a tag population of a given size N, and terminates when the reader has identified all the tags, over the course of several rounds. A limit a, for the maximum number of retransmission, is introduced in the simulation. If a tag replies in a singleton slot, but cannot be identified after $a + 1$ attempts, the reader proceeds to the next slot, and the tag remains unidentified. The tag then participates in the next inventory round, and is identified later. For the deferred retransmission scheme, the reader only makes one attempt to request the tag identifier. For the instant retransmission scheme, the reader requests up to one ($a = 1$) retransmission. In the same manner, there are two variants of the combining scheme: For $a = 1$, the reader attempts one retransmission; for $a = 2$, the reader attempts up to two retransmission, thus a maximum of two and three frames are available for combining. For the HARQ variant, the number of transmissions is also limited to three. In the first attempt, the reader requests the plain *EPC* frame (systematic part); in consecutive trials, the reader requests up to two distinct sets of parity information, PAR_1 and PAR_2. The parity information is generated using a subset of codes from an RCPC code family. The same codes are used, which were previously depicted in Figure 5.2. Essentially, the maximum number of code bits exchanged with the HARQ variant is equal to the combining variant with $a = 2$.

The simulation results of the identification time for the different variants are shown in Figure 5.7. At high SNR, above $E_s/N_0 \approx 9\,\mathrm{dB}$, all variants perform equally well. At this channel conditions, the frame-error probability is very low, so that the first, uncoded transmission of the *EPC* is almost certainly received error free. The identification time then only depends on the number of tags in the system, the number of slots in the FSA frame (see model in Section 3.7.3, [Vog02b, BDVA11]), and the transmission rate. In the simulation, the tag-population size is $N = 128$ tags, and the reader chooses the optimal frame size. The transmission rates are $40\,\mathrm{kbit\,s^{-1}}$ for the T-to-R, and $30\,\mathrm{kbit\,s^{-1}}$ for the R-to-T transmissions. When the channel degrades, i.e., at lower SNRs, the expected number of transmissions increases, and consequently so does the

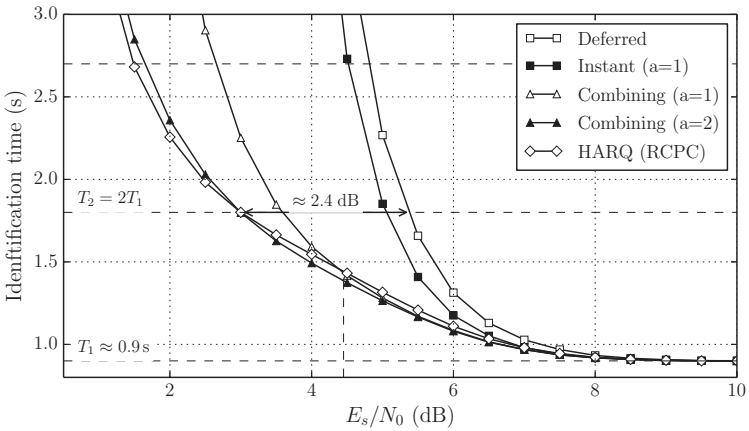


Figure 5.7: Duration of the tag inventory process for a tag-population size of $N = 128$, 96 bit EPCs, and transmission rates of $40\,\mathrm{kbit\,s^{-1}}$ and $30\,\mathrm{kbit\,s^{-1}}$, for the T-to-R and R-to-T links. From [SR13] © 2013 IEEE.

identification time. The variants deferred retransmission (plain), and instant retransmission without combining (instant $a = 1$) perform the worst. At an SNR of $\approx 5.4\,\mathrm{dB}$, the identification time has doubled for the plain variant, compared to the noise-free case, and it continues to grow exponentially for lower SNR values. The instant retransmission scheme performs just slightly better. The improvement can be attributed to the fact that the instant retransmission scheme saves some transmissions compared to the deferred variant. If the first tag reply has errors, the instant retransmission prevents the tag from entering into the FSA process again, avoiding that the tag causes further collisions that would contribute to the degradation of the throughput.

Similar observations can be made for the other variants. Combining with one retransmission performs equally well as the HARQ variant, down to about 4.5 dB. At this SNR, the expected number of transmission exceeds two (shown dashed in Figure 5.7) for the combining variant. So, since the $a = 1$ combining variant only makes two attempts per round, the majority of attempts fail for SNR lower than 4.5 dB. The combining variant with $a = 2$ retransmissions

even performs slightly better than the HARQ. While both variants have similar chances of succeeding after the second retransmission, the HARQ variant occupies the channel for a slightly longer amount of time, because the PAR_1 frame used by the HARQ variant is longer than the retransmission of the EPC. Also note that, while the HARQ variant eventually outperforms the combining variant at $E_s/N_0 = 1\,\mathrm{dB}$ and below, it does not show a substantial improvement, at least not in the configuration and the SNRs considered in the simulation. Nevertheless, the HARQ variant has a coding gain and can improve the system throughput, as will be shown below for a different configuration and in [BR13]. From the simulation results, it is clear however, that the proposed combining and HARQ scheme both can improve the performance of a EPCc1g2 system.

5.4.2 Throughput simulations

After the publication of [BR13], it seemed desirable to compare the proposed scheme to Boyer's results. Boyer suggests a slightly different approach for the implementation the HARQ scheme. He suggests that the tag shall send its data continuously after having been acknowledged, instead of having the reader send ACK and REJ_n commands, as proposed here. For instance: To implement the combining scheme, he proposes that the tag shall continuously repeat the data of the EPC frame. The reader shall then attempt to decode (do the parity check) after receiving a full copy of the information. If this fails, the reader continues to receive the tag's transmission until receiving the next copy of the data is complete, and then combines both copies using MR combining. When the decoding eventually succeeds, the reader powers down, i.e. lower the carrier power, for a short instant, to signal the tag to stop transmitting. Compared to the methods proposed in this work, Boyer's scheme saves the time that is spent by the reader transmitting REJ_n commands, and also avoids sending the frame preamble several times.

Regardless of the practical implementation and the absolute performance of Boyer's protocol and the one proposed here, it is also interesting to compare the throughput of the IR HARQ schemes proposed in both works. Boyer uses a mother code with a rate of $R_c = \frac{1}{4}$ and memory order of $m = 4$, while the mother code in this work is a $R_c = \frac{1}{2}$ and memory order of $m = 3$ that, combined with the first, uncoded transmission of the EPC frame, has an overall

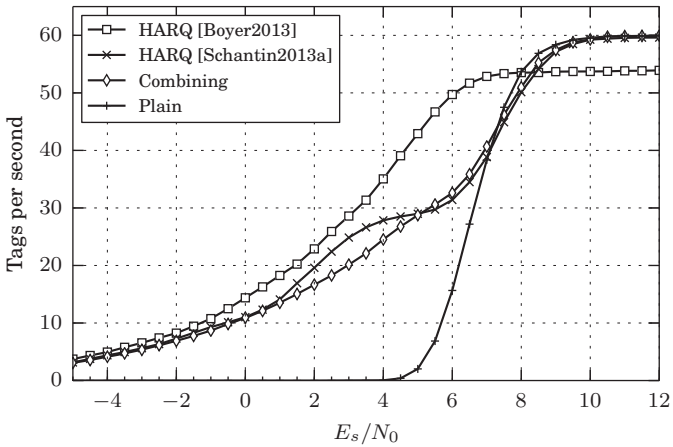


Figure 5.8: Throughput in identified tags per second for different retransmission schemes, comparing the RCPC variants proposed in [BR13] and [SR13].

rate of $R_c = \frac{1}{3}$. Clearly, the code used in this work, and in [SR13], is inferior to the code used by Boyer, in terms of coding gain, because it has a lower free distance $d_f = 8$ (including the systematic part), as opposed to $d_f = 15$ in [BR13]. Nevertheless, both variants were compared in a simulation.

Figure 5.8 shows simulation results that were created using the simulation framework from [SR13], comparing Boyer's code and puncturing schemes with the one described in Section 5.3. The plot shows the system throughput in terms of tags identified per second. The transmission rates were chosen equal to the ones presented in the previous section. The length of the *EPC* frame was chosen according to [BR13], who considered a length of 528 bit.

It can clearly be seen that in the low SNR, Boyer's choice of code outperforms the combination proposed in this work, which is not surprising given the lower code rate and higher memory order of Boyer's chosen code. Also note that the [SR13] HARQ variant is only superior to the combining variant in a limited SNR region between $E_s/N_0 \approx 1\,\text{dB}$ to $\approx 5\,\text{dB}$. At a higher SNR about 8 dB, Boyer's variant is inferior to any systematic code, as well as to the combining

variant. This is due to the fact that the first transmission of Boyer's scheme is coded with rate $R_c = \frac{8}{9}$ and is thus 12.5% longer than the plain tag data, regardless of how good the channel SNR is.

The shortcomings of both approaches, Boyer's and the one proposed previously in this work, can be overcome by using a different code. The code should be systematic, so the first transmission only contains the systematic data (*EPC*). In that way, the information can be transmitted with no redundancy when the channel is good, and thus the maximum throughput can be achieved. The scheme presented in Section 5.3.3 uses a convolutional code, which is systematic and can be described by the generator polynomials $[10_8 15_8 17_8]$. This code however, only has free distance of $d_f = 8$, which is less than optimal. OFD codes with rate $R_c = \frac{1}{3}$ have a higher free distance, e.g. $d_f = 10$ for $[13_8 15_8 17_8]$ for the same constraint length [Bos98]. In general, non-recursive systematic codes do not achieve OFD. Instead, Recursive Systematic Convolutional (RSC) codes may be used, which are systematic while providing the same distance properties like a non-systematic, non-recursive code [BGT93, Bos98].

The following example compares the code and puncturing scheme from [BR13], with a RSC variant with equal rate $R_c = \frac{1}{4}$. The generator polynomials are

$$\mathbf{G}_1 = \begin{pmatrix} 23_8 & 35_8 & 37_8 & 33_8 \end{pmatrix} \qquad \mathbf{G}_2 = \begin{pmatrix} 1_8 & \frac{25_8}{37_8} & \frac{27_8}{37_8} & \frac{33_8}{37_8}. \end{pmatrix} \tag{5.8}$$

where $\mathbf{G}_1$ is from [BR13] and $\mathbf{G}_2$ is the RSC variant of an OFD code taken from [Bos98]. The puncturing scheme for the systematic variant was optimized by a few trial-and-error experiments and is not guaranteed to be optimal. However, as shown in the simulation results in Figure 5.9, the systematic variant performs nearly equal to Boyer's variant at low SNR, while achieving the maximum throughput at high SNR. So in conclusion, an RSC is preferable over a non-recursive coder, because it performs well for both high and low SNR.

It has been shown here, that an HARQ scheme can be readily integrated into the EPCc1g2 standard, and that using HARQ can improve the throughput in the system over a wide range of SNRs. If a modification of the EPCc1g2 standard is not possible, then the system performance can be improved by implementing combining, which requires a modification in the reader only.

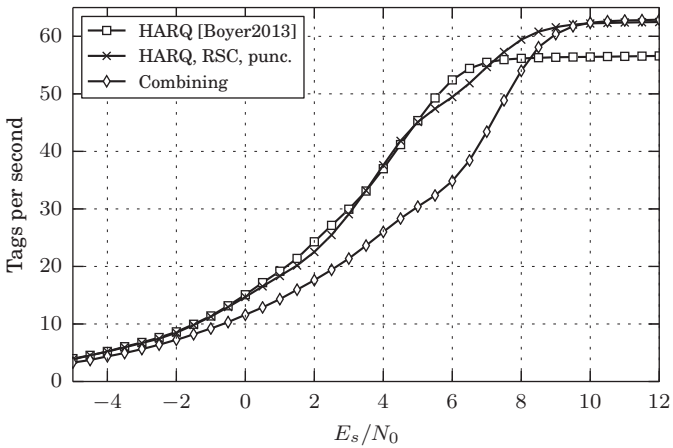


Figure 5.9: Throughput in identified tags per second for different retransmission schemes, comparing the RCPC variants proposed in [BR13] and the RSC variant.

6 The influence of FEC on captures

6.1 A bit-error based capture model

In the previous chapters, FEC was considered as an option to improve the throughput in EPCc1g2 systems, by reducing frame errors on the T-to-R link. The analysis was focused on the exchange of the *EPC* frames, and the capture effect was neglected. Collisions during the inventory phase, however, limit the throughput in a EPCc1g2 system, and while approaches have been proposed to actively resolve tag collisions [ALR10], most readers do not yet implement these. Nevertheless, not all collision slots are lost, due to the capture effect. FEC coding may be an option to improve the frequency of captures, without having to actively separate colliding tag replies. The capture model, which is proposed in this chapter, is more detailed than the classical power-ratio based capture models [Rob75, HKS89]. The capture model is applied to determine in how far FEC coding in the early stages of the tag inventory, specifically coded *RN16* frames, can improve the system throughput.

Suppose an RFID reader is interrogating a tag population, and suppose further that the distance between tags is small compared to the distance of each tag to the reader. E.g. many tags are affixed to items on a palette and the distance to the reader is approximately the same for each tag. Then, the path loss that every tag experiences is equal, except for small scale fading. In the case of a collision of R tags, the received signal at the reader then is the superposition of all tag R signals

$$y_{BP}(t) = \Re\left\{\sum_{r=1}^{R} |h_r| x_r(t) \exp\left(j(2\pi f_c t + \Phi_r)\right) + n(t)\right\} \tag{6.1}$$

where f_c is the carrier frequency, $|h_r|$ is the real, positive channel coefficient, and Φ_r is the random phase term of each tag. The noise at the reader is modelled by a complex white Gaussian noise process $n(t)$. In backscatter systems, f_c is

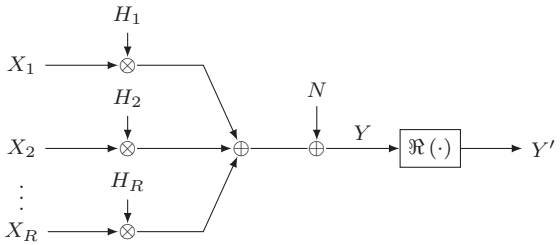


Figure 6.1: Channel model for the collision of R tags (equivalent baseband).

truly equal for all tags, since the carrier signal is generated by the reader, and thus is available for the demodulation at the receiver side. So, the equivalent signal in the complex baseband can be written as

$$y(t) = \sum_{r=1}^{R} |h_r| x_r(t) \exp(j\Phi_r) + n(t) \tag{6.2}$$

The real channel coefficient $|h_r|$ and the phase term $\exp(j\Phi_r)$ can be combined and expressed by a complex channel coefficient h_r. The realisations of the channel coefficient are modeled by a suitable random processes to represent the small scale fading in the system. Figure 6.1 shows the symbol-rate channel model for the R-tag-to-reader collision channel. All values X_r, H_r, N, as well as the received signal Y and its real part Y', can be described by random variables.

The definition of a capture in this model shall be as follows. In the case of a tag collision, i.e., the occupation number of a slot is $R \geq 2$, a capture occurs when the reader is able to receive and decode one of the R tag transmissions without any uncorrectable error. If a t-error correcting code is used, it is then straightforward to define the capture probability as the probability of having t or less errors in one of the tag transmissions. Normally, when no FEC code is used, no errors can be tolerated, so $t = 0$.

Suppose that the reader in only prepared to receive a single tag reply, that coherent detection is used, and that the reader can lock to the signal of one of the colliding tags. Without loss of generality, the index $r = 1$ can be assigned to this tag. The channel coefficient H_1 of the first tag is then a real value, while all other channel coefficients $H_2 \ldots H_R$ are complex values. Furthermore, it is

assumed that the channel coefficients are slowly changing, with respect to the frame duration. This implies that the fading in the system can be described as block fading, which was suggested by [LGSV09, ALR10]. The T-to-R channel is then a discrete memoryless channel, where bit-errors are independent for all bits in the frame. For BPSK transmissions, the reader only evaluates the sign of the real part of the received signal, so a bit error occurs when the sign of the received signal Y' is different from the sign of the transmitted signal X_1. So, the bit-error probability is

$$\begin{aligned} p_b = &\Pr\left(X_1 = 1\right) \Pr\left(Y' < 0 \mid X_1 = 1\right) + \\ &\Pr\left(X_1 = -1\right) \Pr\left(Y' > 0 \mid X_1 = -1\right) \end{aligned} \tag{6.3}$$

where $Y' = \Re\{Y\}$. The value of the input bits of the R tags is described by the binary random variables $X_r \in \{-1, 1\}$, whereas the channel output is a continuous random variable Y. If all transmissions are equally likely, then $Pr\left(X_1 = 1\right) = Pr\left(X_1 = -1\right) = \frac{1}{2}$ and (6.3) can be rewritten using the conditional PDF of Y'.

$$p_b = \frac{1}{2} \int_{-\infty}^{0} p\left(y' \mid x_1 = 1\right) dy + \frac{1}{2} \int_{0}^{\infty} p\left(y' \mid x_1 = -1\right) dy \tag{6.4}$$

The channel output does not only depend on X_1 however, but on all tag transmissions $\mathbf{X} = (X_1, \ldots, X_R)$, the channel coefficients and the added noise. While y is the complex signal received by the reader, only the real part $Y' = \Re\{Y\}$ is taken into account in the demodulation. Let $\mathbf{H}$ be the random vector of the R complex channel coefficients with $\mathbf{H} = \mathbf{A} + j\mathbf{B}$. Furthermore, let N be a real zero-mean Gaussian random variable with the variance $\sigma^2 = \frac{N_0}{2}$, describing the real part of the AWGN. The real part of the channel output is then

$$Y' = \mathbf{A}\mathbf{X} + N \tag{6.5}$$

and its conditional PDF is

$$p\left(y' \mid \mathbf{x}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{\left(y' - \mathbf{A}'\mathbf{x}\right)^2}{\sigma^2}\right) \tag{6.6}$$

To distinguish the desired signal from the interfering tags, x_1 can be written separately, and the interference caused by the colliding tags can be written as $\mathbf{A}_2'\mathbf{x}_2$ in (6.6), where $\mathbf{A}_2 = (A_2, \cdots A_R)$ and $\mathbf{x}_2 = (x_2, \cdots x_R)^T$. Summing over all possible transmit vectors $\mathcal{X}_2$ of the interferers, finally yields the conditional PDF for y' given x_1

$$p\left(y' \mid x_1\right) = \frac{1}{2^{R-1}} \sum_{\mathbf{x}_2 \in \mathcal{X}_2} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{\left[y' - (A_1 x_1 + \mathbf{A}_2'\mathbf{x}_2)\right]^2}{\sigma^2}\right) \tag{6.7}$$

Observing that

$$\Pr\left(Y' < 0 \mid X_1 = 1\right) = \Pr\left(Y' > 0 \mid X_1 = -1\right)$$

and with (6.7) in (6.4), the bit-error probability is

$$\begin{aligned} p_b = & \frac{1}{2^{R-1}} \sum_{\mathbf{x}_2 \in \mathcal{X}_2} \frac{1}{\sqrt{2\pi\sigma^2}} \times \\ & \int_{-\infty}^{0} \exp\left(-\frac{1}{2}\frac{\left(y - (A_1 + \mathbf{A}_2\mathbf{x}_2)\right)^2}{\sigma^2}\right) dy \end{aligned} \tag{6.8}$$

which can be further simplified using the complementary error function $\operatorname{erfc} u = \frac{2}{\sqrt{\pi}} \int_u^\infty e^{-t^2}$ and the substitutions

$$u = \frac{y - (A_1 + \mathbf{A}_2\mathbf{x}_2)}{\sqrt{2}\sigma} \qquad \frac{dy}{du} = \sqrt{2}\sigma \tag{6.9}$$

The bit-error probability is then a sum of 2^R complementary error functions.

$$p_b = \frac{1}{2^R} \sum_{\mathbf{x}_2 \in \mathcal{X}_2} \operatorname{erfc}\left(\frac{A_1 + \mathbf{A}_2\mathbf{x}_2}{\sqrt{2}\sigma}\right) \tag{6.10}$$

The capture probability, as defined before, is the probability of having less than t errors in a frame, where t is the error correction capability of the FEC code, if any is used, or zero otherwise. So,

$$p_{c0}\left(\mathbf{A}\right) = \sum_{i=0}^{t} \binom{n}{i} p_b\left(\mathbf{A}\right)^i \left(1 - p_b\left(\mathbf{A}\right)\right)^{n-i} \tag{6.11}$$

is the capture probability for *one* realisation of the channel coefficients, where $\mathbf{A}$ is a random vector whose values are changing for every transmitted frame. The average capture probability is the expected value of (6.11)

$$p_c(R) = \int_{\mathcal{A}} f(\mathcal{A}) p_{c0}(\mathcal{A}) d\mathcal{A} \tag{6.12}$$

where p_c then only depends on the occupation number of the slot. If the channel coefficients in $\mathbf{A}$ are independent random variables, which is a valid assumption if the colliding tags are separated spatially, the PDF $f(\mathcal{A})$ is merely the product of the R individual PDFs

$$f(\mathcal{A}) = \prod_{r=1}^{R} f(a_r). \tag{6.13}$$

Note that $f(a_1)$, the PDF of the first tag differs from the remaining ones, since only the first tag is received coherently.

The PDF for the first tag is known for common channel models (Rayleigh distribution etc.). The PDFs of the channel coefficients of the interferers are derived below for some relevant channel models.

6.1.1 AWGN channel

In a system where there is no small-scale fading altogether, and the tags are all placed at an approximately equal distance to the reader, the T-to-R channels can be described with a simple AWGN model. In this case, the magnitude of the channel coefficients is not random at all, i.e., $|h_r| = 1$ for all tags. In the case of the first, coherently received tag, the reader also recovers the phase, so that $h_1 = 1$. For all other tags however, the coefficients h_r include a random phase term. The phase of the r^{th} ($r > 1$) tag is thus a random variable Φ_r with a uniform PDF

$$f_{\Phi_r}(\varphi_r) = \frac{1}{2\pi} \qquad -\pi < \varphi_r \leq \pi \tag{6.14}$$

The impact of the interference created by these tags, though, depends on the real part of A_r only. So, for the evaluation (6.12) $f(a_r)$, the PDF of $A_r = \Re\{H_r\}$ is needed. It can be obtained by first deriving the Cumulative Distribution Function (CDF) for A_r. As shown in Figure 6.2a, there are two intervals of φ_r

that correspond to $A_r \leq a_r$, $\cos\Phi_r \leq a_r$ for $\varphi_r < 0$ and $\varphi_r \geq 0$, so in the range of $-1 \leq a_r \leq 1$

$$F(A_r) = \Pr(-1 \leq A_r \leq a_r) \tag{6.15}$$

$$= \Pr(-1 \leq \cos\Phi_r \leq a_r)\Big|_{0\leq\varphi_r<\pi} + \Pr(-1 \leq \cos\Phi_r \leq a_r)\Big|_{-\pi\leq\varphi_r<0} \tag{6.16}$$

The inverse of the cosine function is applied to both arguments on the right-hand side. For positive phase angles, $0 \leq \Phi_r < \pi$, the arccos is used, while $-\arccos$ is used to obtain the inverse of the cosine that will give the angles $-\pi \leq \Phi_r < 0$.

$$\Pr(A_r \leq a_r) = \Pr(\pi > \Phi_r \geq \arccos a_r) + \Pr(-\pi \leq \Phi_r \leq -\arccos a_r) \tag{6.17}$$

$$= \int_{\arccos a_r}^{\pi} \frac{1}{2\pi} d\varphi_r + \int_{-\pi}^{-\arccos a_r} \frac{1}{2\pi} d\varphi_r \tag{6.18}$$

$$= \frac{1}{2\pi}\left[\pi - \arccos a_r - \arccos a_r + \pi\right] \tag{6.19}$$

$$= 1 - \frac{1}{\pi}\arccos a_r \tag{6.20}$$

(6.20) defines the CDF on the interval $-1 \leq a_r \leq 1$. For all other values, it can be defined piecewise

$$\Pr(A_r \leq a_r) = \begin{cases} 0 & a_r < -1 \\ 1 - \frac{1}{\pi}\arccos a_r & -1 < a_r < 1 \\ 1 & a_r > 1 \end{cases} \tag{6.21}$$

so the CDF becomes a continuous function. Deriving (6.21) gives the PDF of A_r as shown in Figure 6.2b.

$$f_{A_r}(a_r) = \frac{d}{da_r}\Pr(A_r \leq a_r) = \begin{cases} 1 - \frac{1}{\pi}\frac{d}{da_r}\arccos a_r & -1 \leq a_r \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.22}$$

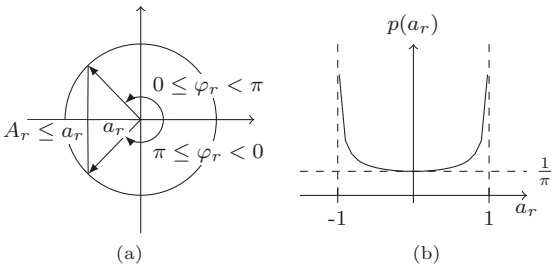


Figure 6.2: Dependency of the real part of a complex random variable with uniform phase distribution (a), and the corresponding density function (b)

$$= \begin{cases} \frac{1}{\pi} \frac{1}{\sqrt{1-a_r^2}} & -1 \leq a_r \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{6.23}$$

Substituting (6.23) in (6.12), the capture probability for the AWGN case can be calculated. It should be noted that only the coefficients $A_2 \ldots A_R$ are random variables in the AWGN case, so the integration from (6.12) is reduced by one dimension.

6.1.2 Rayleigh channel

Modeling the T-to-R channel as AWGN channel might not be appropriate for all RFID scenarios. As discussed in 3.5, small-scale fading may be present on the T-to-R link. Rayleigh fading can be incorporated easily into the proposed capture model. In the Rayleigh channel model, there are multiple propagation paths between a tag and the reader, each associated with a random path-loss and a random phase term. The resulting channel coefficients $H_r = A_r + jB_r$ are random variables, where the real component A_r (in-phase) and the imaginary components B_r (quadrature-phase), both follow a zero-mean normal distribution. Both, real and imaginary part, have a variance of $\frac{\Omega}{2}$, where $\Omega = E\left[|H_r|^2\right]$ is the second order moment of the Rayleigh random variable $|H_r|$ (average received power). The complex PDF of H_R is invariant with respect to phase-

rotation, so regardless of any additional phase terms, the real part $A_r = \Re\{H_r\}$ follows a normal distribution.

$$f(a_r) = \frac{1}{\sqrt{\pi\Omega}} \exp\left(-\frac{a_r^2}{\Omega}\right) \tag{6.24}$$

The coefficient $H_1 = A_1$ of the first tag is Rayleigh distributed

$$f(a_1) = \begin{cases} \frac{2a_1}{\Omega} \exp\left(-\frac{a_1^2}{\Omega}\right) & a_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.25}$$

6.1.3 Dyadic backscatter channel

The Dyadic Backscatter Channel (DBC) (see 3.5.4) was specifically defined to model the fading in a backscatter system. In a backsctter channel both, forward and backward link, are subjected to fading. The channel coefficient of the DBC is thus the product of two complex random variables, H_f and H_b. In a monostatic reader setup, the forward and the backward channel are identical, due to the reciprocity theorem of electromagnetism. So, the random variables H_f and H_b are fully correlated. In general, i.e., for a bistatic reader setup, there may be some correlation between H_f and H_b, or they may be independent, if the transmit and receive antennas of the reader are sufficiently separated in space.

When the forward and backward channels are modeled as Rayleigh channels, then the corresponding channel coefficients are $H_f = X_f + jY_f$, and $H_b = X_b + jY_b$. The real random variables X_f, X_b, Y_f, Y_b are zero-mean Gaussian distributed, with the variances $\frac{\Omega_f}{2}$, $\frac{\Omega_b}{2}$. In the *fully-correlated* case, the forward and backward channel coefficients are identical ($H_f = H_b$)

$$H_r = H_f H_b = H_f^2 = \underbrace{X_f^2 - Y_f^2}_{A_r} + 2jX_fY_f \tag{6.26}$$

So, the interference in the capture model is a random variable, which is the difference of two χ_1^2 random variables. The PDF of the difference $A_r = X_f^2 - Y_f^2$

is the convolution of χ^2-distribution and a χ^2-distribution that is reversed in $x_r = -y_r$. The characteristic function of the χ_1^2 distribution is

$$\phi_{\chi_1^2}(t) = (1 - 2jt)^{-\frac{1}{2}} \tag{6.27}$$

and the characteristic function of the reversed PDF is reversed in t (time reversal property of the Fourier transform). So, the characteristic function of the difference is

$$\phi_{\chi_1^2 - \chi_1^2}(t) = (1 - 2jt)^{-\frac{1}{2}}(1 + 2jt)^{-\frac{1}{2}} \tag{6.28}$$

$$= \frac{1}{\sqrt{(1 - 2jt) \cdot (1 + 2jt)}} \tag{6.29}$$

$$= \frac{1}{\sqrt{1 + 4t^2}} = \frac{\frac{1}{2}}{\sqrt{\left(\frac{1}{2}\right)^2 + t^2}} \tag{6.30}$$

$$\tag{6.31}$$

The PDF corresponding (6.31) can be found in the literature [Obe73, Eq. 400]

$$f_{\chi_1^2 - \chi_1^2}(x) = \frac{K_0\left(\frac{1}{2}|x|\right)}{2\pi} \tag{6.32}$$

where K_α is a modified bessel function of the second kind with order α. (6.32) is also known as the product normal distribution [Wei14]. Note that this also shows that real and imaginary part in (6.26) both have the same PDF. (6.32) holds only if X_f, Y_f are standard normal distributed. For X_f, $Y_f \sim \mathcal{N}\left(0, \frac{\Omega}{2}\right)$

$$f_{A_r}(a_r) = \frac{K_0\left(\frac{|a_r|}{\Omega}\right)}{\pi\Omega} \tag{6.33}$$

The PDF for $H_1 = A_1$ is given in [GD08, Table II] for the fully correlated DBC.

$$f_{A_1}(a_1) = \begin{cases} \frac{1}{\Omega}\exp\left(\frac{-a_1}{\Omega}\right) & a_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.34}$$

where Ω is the parameter of the Rayleigh distribution of the forward H_f and backward H_b channels. For a normalized receive power, i.e., $E[A_1] = 1$, the parameter has to be chosen as $\Omega = \frac{1}{\sqrt{2}}$, in the fully correlated DBC.

In the case of *independent* forward and backward channel coefficients, H_f and H_b are independent random variables.

$$H_r = H_f H_b = (X_f + jY_f)(X_b + jY_b) = \underbrace{X_f X_b - Y_f Y_b}_{A_r} + j(X_f Y_b + Y_f X_b) \tag{6.35}$$

Here, A_r is the difference – or sum, since the PDF of the normal product distribution is an even function – of two product normal random variables. The PDF of A_r is, again, most easily derived using the characteristic function (6.32). For X_f, Y_f, X_b and $Y_b \sim \mathcal{N}\left(0, \frac{\Omega}{2}\right)$, the PDFs and the corresponding characteristic function are

$$f_{X_f X_b}(x) = \frac{1}{\pi \frac{\Omega}{2}} K_0\left(\frac{x}{\frac{\Omega}{2}}\right) \circ\!\!-\!\!\bullet\, \phi_{X_f X_b}(t) = \frac{\frac{2}{\Omega}}{\sqrt{\left(\frac{2}{\Omega}\right)^2 + t^2}} \tag{6.36}$$

$$f_{A_r}(a_r) = \frac{4}{\Omega}\sqrt{\frac{|a_r|}{\Omega\pi}} K_{\frac{1}{2}}\left(\frac{2|a_r|}{\Omega}\right) \circ\!\!-\!\!\bullet\, \phi_{A_r}(t) = \frac{\left(\frac{2}{\Omega}\right)^2}{\left(\frac{2}{\Omega}\right)^2 + t^2} \tag{6.37}$$

The inverse transformation of (6.37) [Obe73, Eq. 406] of the square of the characteristic function of a normal product distribution yields the PDF of A_r.

The channel coefficient $H_1 = A_1$ for independent forward and backward channel in the DBC is given in [GD08, Table I]

$$f_{A_1}(a_1) = \begin{cases} \frac{4a_1}{\Omega^2} K_0\left(\frac{2a_1}{\Omega}\right) & a_1 \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.38}$$

and $\Omega = 1$ has to be chosen for a normalized receive power.

6.2 Capture probabilities for selected cases

The PDFs given above for the distribution of the interference amplitudes in the different channel models can now be used together with (6.10) to (6.12), to calculate the capture probabilities for the different channels and varying SNRs.

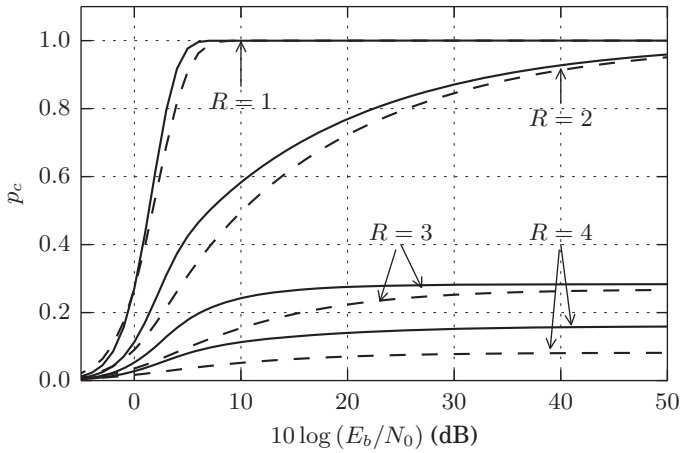
The Figures 6.3 and 6.4 show plots of the capture probability for various cases that were obtained by numerical evaluation of (6.12). The capture probabilities were calculated for the transmission of frames with 16 data bits (*RN16*).

The analysis compares two variants for each channel model and different occupation numbers. The first variant is *plain* transmission. In this case, the tag(s) transmit 16 data bits without redundancy, and a capture only happens if all $n = 16$ bits of one tag are received error-free. This case corresponds to the regular *RN16* frame in an EPCc1g2 system. In the second case, an FEC code is used to add redundancy to the transmitted data. In the system as described above, bit-errors are independent and the frame length is short. Thus, a Bose-Chaudhuri-Hocquenghem (BCH) code with the parameters $(31, 16, 7)$ was chosen from the literature [Bos98]. This code exactly accommodates $k = 16$ information bits, and can correct $t = \left\lfloor \frac{7-1}{2} \right\rfloor = 3$ bit errors. So, three errors can be tolerated in the transmission; however, the frame length is increased from 16 to $n = 31$ bit. A capture happens only if three, or less of the 31 symbols, are received with errors (see (6.11)). When comparing the two variants, the SNR is adjusted so that the average bit-energy E_b remains the same in the plain and the FEC case. This corresponds to an $1/R_c$-fold increase of the noise variance σ^2 in (6.11), where R_c is the code rate of the BCH code (Here, $R = \frac{16}{31}$).

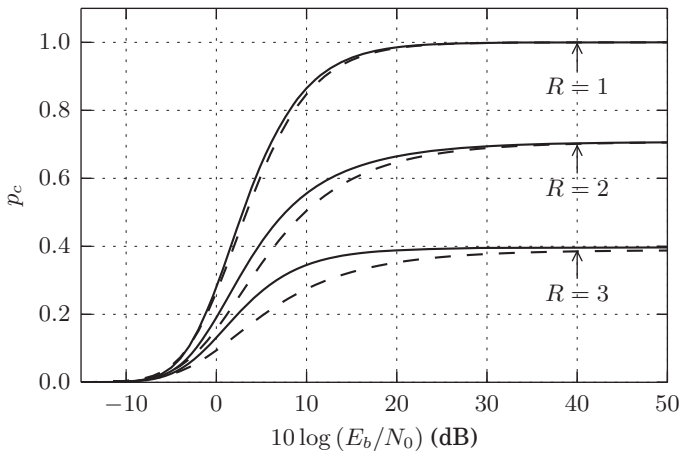
All plots in the Figures 6.3 and 6.4 show the capture probability plotted over a range of E_b/N_0. In the plots, the plain variant is shown as a dashed line, and the variant with the BCH code is shown as a solid line. The capture probability $p_c(R)$ is shown for $R = 1 \ldots 3$, and $R = 4$ in the AWGN case. Note that $R = 1$ marks a singleton slot; so, for this case the p_c is the complement of the frame-error probability. Strictly speaking, $R = 1$ does thus not constitute a capture, because no collision occurs in the first place.

The results of the AWGN case exhibit one distinctive difference compared to the capture probabilities in the fading channels. Figure 6.3a shows that for $R = 2$, the capture probability asymptotically becomes equal to one for $E_b/N_0 \to \infty$, with and without coding. This means that, in the absence of noise, an arbitrarily small phase difference between the coefficients H_1 and H_2 is sufficient to recover the first tag reply. For higher occupation numbers, collisions are not likely to be resolved in the AWGN channel. For $R = 3$ the asymptotic capture probability drops to 0.28, and to 0.16 for $R = 4$.

The asymptotic capture probabilities in the fading channels differ only slightly among each other. In the Rayleigh channel $p_c \to 0.7$ for $K = 2$ and $E_b/N_0 \to \infty$; for the DBC, $p_c \to 0.66$ and $p_c \to 0.63$, in the independent and in the fully

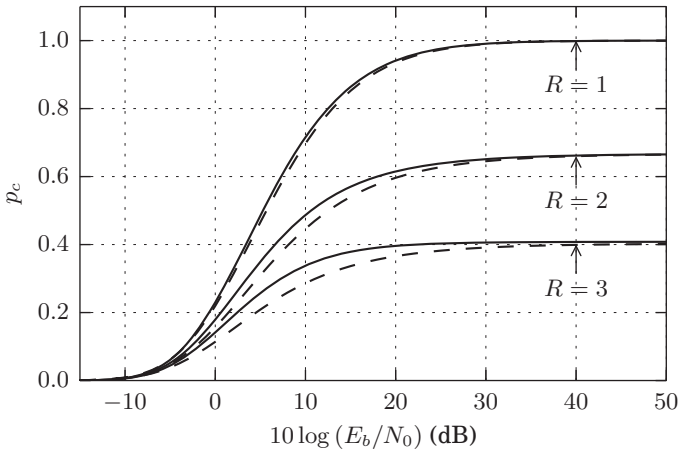


(a) AWGN

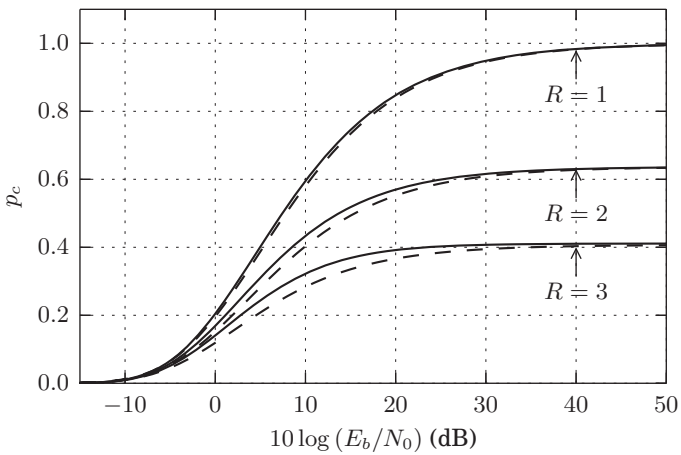


(b) Rayleigh

Figure 6.3: Capture probability for an R-tag collision slot in the AWGN and Rayleigh channel models. Comparison of plain transmission (dashed) and coded transmission (solid), using a $(31, 16, 7)$ BCH code for FEC (previously published in [SR14] © 2014 IEEE).



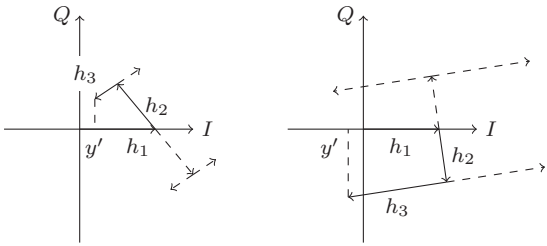
(a) Independent



(b) Fully correlated

Figure 6.4: Capture probability for an R-tag collision slot in the DBC model, for the case of independent, and fully correlated forward and backward links. Comparison of plain transmission (dashed) and coded transmission (solid), using a $(31, 16, 7)$ BCH code for FEC (previously published in [SR14] © 2014 IEEE).

correlated case, respectively. For $R = 3$, the capture probability drops to 0.40 for Rayleigh, and to 0.41 for both cases of the DBC.



(a) $\Re\left(h_2 + h_3\right) < h_1$ (b) $\Re\left(h_2 + h_3\right) > h_1$

Figure 6.5: Examples for constellations of the noise-free signal in an $R = 3$ collision slot. (a) constellation with reduced noise margin (error-free in the absence of noise). (b) constellation with an inversion of the sign of y' for $c = 2$ constellation points.

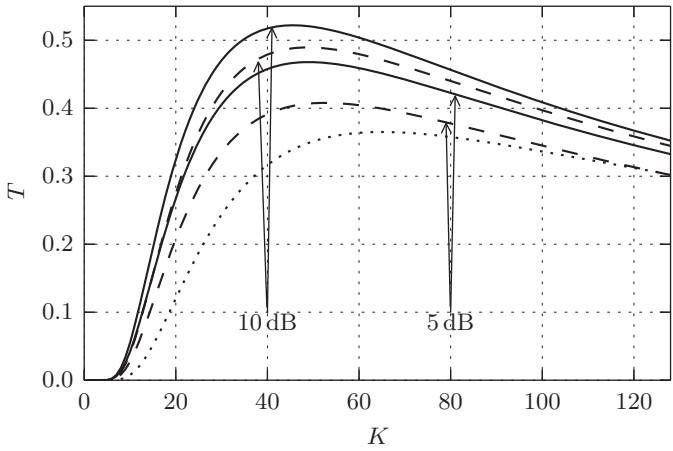
The asymptotic capture probability at high E_b/N_0 is determined only by the distribution of the channel coefficients, for all the cases. In the absence of noise, bit-errors do not occur when $|\Re(\mathbf{X}_2\mathbf{H}_2)| < |H_1|$ for all possible $X_1, \mathbf{X}_2$. Figure 6.5) illustrates this for two examples. When all the constellation points at the channel output lie in the same half-plane as $X_1 H_1$ (Fig. 6.5a), then X_1 is always received correctly in the absence of noise. Otherwise, if one or more constellation points lie in the opposite half-plane (Fig. 6.5b), then bit-errors happen with a probability $\frac{c}{2^{R-1}}$, where c is the number of constellation points, which cause errors.

Figures 6.3 and 6.4 also show the influence of the BCH code on the capture probability. It can clearly be seen that the two variants do not differ for high SNRs, for any of the channel models and the chosen code. This is because either all bits are received correctly for the case (a) illustrated above, or the error probability is very high and the correction capabilities of the BCH code are always exceeded, for case shown in (b). So, for high SNRs, the capture probability is not increased for the BCH variant. At low SNRs, $E_b/N_0 < 0\,\mathrm{dB}$ for AWGN and $E_b/N_0 < 5\,\mathrm{dB}$ for the fading channels, the variant using the BCH code performs even slightly worse than the plain variant, though this in hardly

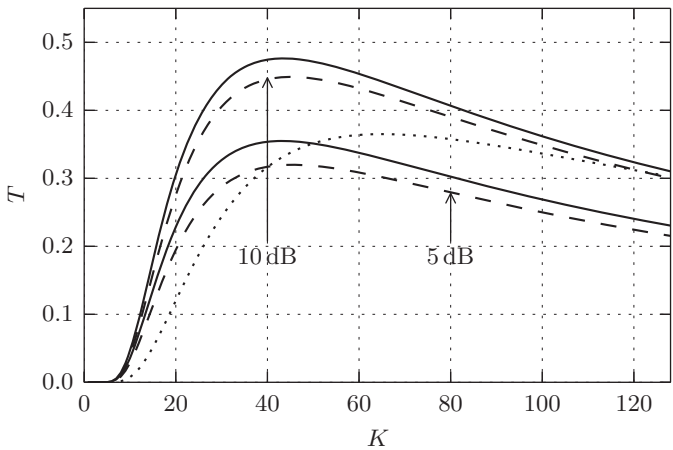
visible in the plots. In this low SNR regime, the channel code does not have a coding gain, i.e., the code cannot make up for the energy spent on transmitting the redundancy bits. An improvement of the capture probability can only be found for E_b/N_0 in the range of approximately 0 dB to 25 dB, depending on the channel model. But even in this range it is complicated to give a conclusive quantification in the form of a coding gain, since the gain strongly depends on the number of colliding tags. E.g. in the AWGN channel, there is a coding gain of ≈ 1.5 dB for $p_c = 0.9$ in a singleton slot. For $R = 2$, the coding gain is bigger, approximately ≈ 2.5 dB, though it is hard to make out in the plot. For $R > 2$, no coding gain can be defined at $p_c = 0.9$, since the capture probability is always below this values.

The improvements can be analyzed better when the throughput is considered. Recall T_c, the frame-local throughput, that is defined as the ratio of useful slots to the number of available slots in the FSA frame in Section 3.6.3. Useful slots, in this case, are all slots that will contribute to the tag identification, i.e., which contain an error-free identifier, regardless of whether a collision happens or not. With the proposed capture model, for every channel model and a given SNR, the throughput can be calculated by using (6.12) in (3.18). The results of the frame-local throughput for a tag population size of $N = 64$ are shown in the Figures 6.6 and 6.7. Each plot shows the frame-local throughput over a variable number of slots for a certain channel model. The plots show the variant without coding (dashed) and the variant using the BCH code for SNRs of 5 dB and 10 dB. In addition, each plot shows the frame-local throughput of FSA according to (3.14) for $M = 1$, i.e., the FSA throughput disregarding the capture effect and bit- or frame-errors altogether.

Figure 6.6a illustrates the predicted throughput in a system, where the T-to-R communication can be modeled by an AWGN channel model. The results obtained with the proposed capture model differ in various way from the classical model without captures and frame errors. The predicted maximum throughput is 0.41 at 5 dB and 0.49 at 10 dB, as opposed to the well known value of 0.37 for the classical FSA model, due to the non-zero capture probability. So, the predicted throughput is higher, although the new model considers frame-errors, and the frame error-probability is approximately $p_f \approx 10^{-1}$ at 5 dB. Also, the maximum of the throughput is achieved for a lower number of slots. While the classical analysis of FSA states that the maximum throughput is reached for an

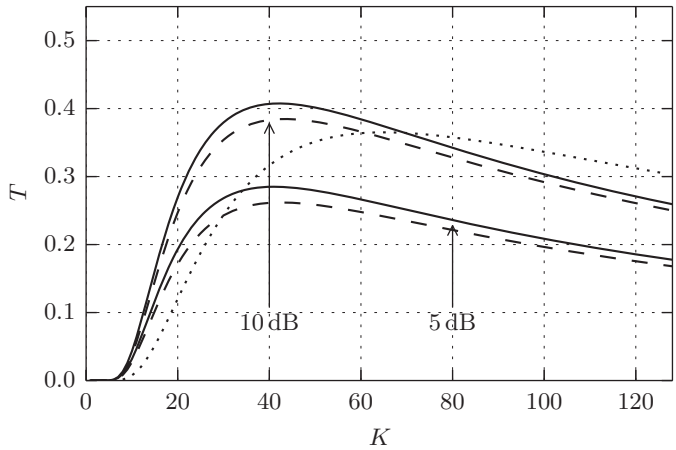


(a) AWGN

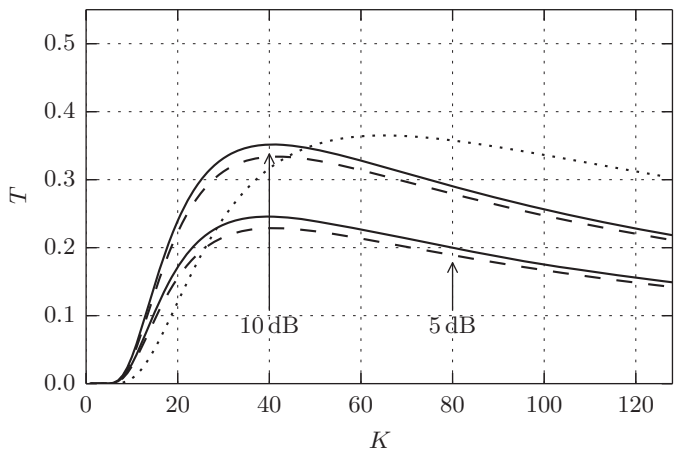


(b) Rayleigh

Figure 6.6: Frame-local throughput in the AWGN channel (a) and in the Rayleigh channel (b), for a tag population of $N = 64$ and $E_b/N_0 = 5\,\mathrm{dB}$, and $E_b/N_0 = 10\,\mathrm{dB}$. Plain transmission (dashed) and FEC variant (solid) (previously published in [SR14] © 2014 IEEE).



(a) Independent



(b) Fully correlated

Figure 6.7: Frame-local throughput in the DBC, with independent (a) and fully correlated (b) forward and backward links. Shown for a tag population of $N = 64$ and $E_b/N_0 = 5\,\mathrm{dB}$, and $E_b/N_0 = 10\,\mathrm{dB}$. Plain transmission (dashed) and FEC variant (solid) (previously published in [SR14] © 2014 IEEE).

equal number of slots and tags ($N = K$), the consideration of captures reveals that the optimal number of slots is actually lower. Sure enough, this fact alone has been pointed out previously by [Rob75, BDVA11] and many others. But while [BDVA11] prediction relies on empirical values from [BLW09] to define the capture probability, the proposed model gives a similar result, based just on the channel characteristics.

Using the capture probabilities calculated with the proposed capture model, the optimal number of slots and the maximum frame-local throughput can now be determined, and the improvements when using the BCH code can be quantified. For the AWGN channel, the maximum throughput $T = 0.41$ is achieved at a slot-count of $K = 53$, for an E_B/N_0 of 5 dB. For 10 dB, the predicted throughput is $T = 0.49$ at the optimal slot-count of $K = 49$. The throughput increases slightly when the BCH code is used. The improved maximums are $T = 47$ for $K = 49$ and $T = 0.52$ at $K = 46$ slots, for $E_b/N_o = 5$ dB and 10 dB, respectively. Similar observations can be made for the fading-channel models. These results show as well that the optimal throughput is only reached when choosing a smaller number of slots $K < N$, compared to the value of $K = N$ that is often suggested for FSA when the capture effect is neglected. Also for this channels, using a BCH code on the T-to-R link slightly increases the throughput, because the frequency of captures is increased. Table 6.1 summarizes the results for the selected combinations of channel models, E_b/N_0 with and without coding. For the fading channels, the improvements achieved by the introduction of the BCH code are modest. The improvements of the throughput are approximately two to three percent points, while the number of necessary slots is reduced by one to two. The reduction of the optimal slot number is fairly irrelevant though, considering that in a practical system K cannot be changed arbitrarily in an EPCc1g2 system, but has to be a power of two ($K = 2^Q$).

6.3 Limitations of the capture model

The capture model presented in this chapter permits to evaluate the capture probabilities and the associated frame-local throughput in a backscatter RFID system. The model requires no parameters, except for the choice of the channel model, and yields results that are comparable to other capture models [BDVA11] that rely on empirical values to characterize the capture probability. Unlike

Table 6.1: Summery of the throughput analysis

Channel model	E_b/N_0	FEC	$T_{\max}$	K_{opt}
AWGN	5 dB	none	0.41	53
		$(31, 16, 7)$	0.47	49
	10 dB	none	0.49	49
		$(31, 16, 7)$	0.52	46
Rayleigh	5 dB	none	0.32	46
		$(31, 16, 7)$	0.35	43
	10 dB	none	0.45	45
		$(31, 16, 7)$	0.48	43
DBC (indep.)	5 dB	none	0.26	43
		$(31, 16, 7)$	0.29	41
	10 dB	none	0.38	43
		$(31, 16, 7)$	0.41	42
DBC (fully corr.)	5 dB	none	0.23	41
		$(31, 16, 7)$	0.25	40
	10 dB	none	0.33	42
		$(31, 16, 7)$	0.35	41

simpler models, the proposed model calculates the (frame) capture probability, based on the bit-error probability and the frame length of the data exchanged in an FSA slot. It thus can be used to evaluate the impact of noise, changes in frame length and, most importantly, the influence of FEC codes on the capture probability and the throughput, by considering the error-correction capability of the code.

The proposed model has, however, a few drawbacks, some of which can be overcome more easily than others. Most of the existing capture models (only) take the spatial distribution of the contending stations into account [Rob75]; the model proposed above does not do so at all. It is assumed that power variations among different tags can be modeled by the fading process. If this is not sufficient, large-scale fading could be easily included in the model, by treating the average E_b/N_0 as a random variable and providing a suitable PDF to describe the spatial distribution of the tags, e.g. a uniform distribution over a range of distances from the reader. (6.12) would then also have to calculate the expected value with respect to the spatial PDF. Another drawback

is the assumption of coherent detection at the reader, though it is also possible and desirable to use incoherent detection, when the differential FM0 baseband code is used. Clearly, for incoherent detection, the calculation of the bit-error probability becomes much more difficult. Finally, it is in general, fairly difficult to solve (6.12), as R increases.

In spite of this shortcomings, the model can be a useful tool to quickly analyze the impact of certain FEC coding schemes on the capture probability. Where the model is not sufficient, monte carlo simulations have to be used as an alternative to determine the capture probability, throughput, etc., in a system.

7 Summary

7.1 Results

The results presented in this thesis show that classical error-control methods can significantly improve the performance of long-range RFID systems, which was specifically shown for the EPCc1g2 standard systems. The improvements have been shown by way of simulation, using different analytical models and by using combinations of both methods.

The results given in Section 4.1 show an improvement of the duration of the tag inventory process, when FEC is used for a system where errors are dominant in the T-to-R link. The results are based on the DTMC model presented in Section 3.6.2. The model was extended to consider the number of tag replies that fail in an inventory round, due to frame-errors in the tag replies. The required frame-error rates can either be calculated, for a simple channel model and simple codes, or obtained by simulation, when more complex codes are considered. Using the proposed model, it was shown in Section 4.1 that the total number of required inventory rounds for the FSA based tag inventory process increases drastically when tags are far away from the reader. Using FEC on the T-to-R link can reduce the number of required rounds, and significantly speed up the tag inventory process in this case. The obtained results support the findings from [JC09], while additionally providing an estimation of the range improvements. The expected extension of the read range is about 4 m ($\approx 25\%$), when free-space propagation is present in the system.

If FEC codes are used on the T-to-R link of a EPCc1g2 system, they will be used in combination with the standard baseband (line) codes: FM0 and Miller coding. In the remainder of Chapter 4, the bit-error performance of the standard baseband codes was analyzed, and SISO decoders were proposed for the baseband codes in the Section 4.2. In a concatenated system, SISO decoders are required for the inner decoder, here the baseband decoder, to achieve good

decoding results in the outer decoder. The decoding of the concatenation of baseband and channel code was discussed in Section 4.3, and an improved turbo-like iterative decoder was presented in Section 4.4. It was shown by BER and EXIT chart analysis that the iterative decoder achieves an added coding gain between 0.4 dB to 2.8 dB compared to a separate, non-iterative decoding of the baseband and the FEC code.

In Chapter 5, several retransmission schemes for a standard EPCc1g2 system were analyzed. It was demonstrated how combining can be used in a standard system to reduce the identification times and improve the throughput for bad T-to-R channel conditions. This can be achieved by extending *only* the reader, while retaining full standard compatibility. Furthermore, an extension for the EPCc1g2 standard that is based on HARQ and RCPC codes was proposed. The presented simulation results show that both HARQ and the standard compatible combining methods can improve the throughput and overall identification time of a tag population, which was independently shown by [BR13]. Subsequently, it was shown in Section 5.4.2 that the results for HARQ can be improved further, when RSC convolutional codes are used. The system can then achieve maximum throughput at high SNR, while remaining close to the capacity limit over a wide range when the SNR is decreased.

Lastly, in Chapter 6, a capture model for collision slots was proposed. This model allows the calculation of the capture probability in a collision slot for multiple tag signals and added Gaussian noise. The interference of the colliding tags is described by a set of random channel coefficients in this model. The required PDFs were presented for the AWGN, the Rayleigh and the dyadic backscatter channel. The proposed model is suitable for evaluating the impact of FEC coding on the capture probability. Corresponding results for the capture probabilities and the frame-local throughput were presented for a modified tag-inventory process, that uses a BCH coded identifier instead of the regular *RN16* during the collision phase. The results show an improvement of the capture probability for medium channel conditions and a minor increase in the throughput. Since the improvements are only marginal, it has to be concluded that a modification of the EPCc1g2 tag inventory will not significantly improve the read rates of the system.

The results from this thesis, and the existing works that were reviewed, strongly suggest that adding FEC to the T-to-R link of a long-range RFID

will bring significant improvements concerning the throughput in the system, the read rates and the overall identification times. Adding an HARQ scheme on the T-to-R link of the system can be done without adding too much complexity to the tag, and should be considered as an extension for the EPCc1g2 standard. The system will mainly benefit from the proposed changes, if the T-to-R link is the limiting factor. However, this will be increasingly the case while the sensitivity of passive tags keeps improving, and especially with the advent of the new BAP tags, specified in the latest EPCglobal and ISO standards ([EPC13],[ISO13, Clause 7.5]).

7.2 Future work

Certain simplifications were made throughout this work, when analyzing the existing EPCc1g2 standard and the proposed schemes. Often this meant that the AWGN model was used, and fading effects were neglected, when simulating transmission. Considering small-scale fading would probably give more realistic results, e.g. for the range estimations in Chapter 4. Since the coding gain is usually greater in fading channels, the absolute improvements of the throughput and the identification times would likely be even larger.

The proposed HARQ scheme from Chapter 5 could potentially be improved as well. It was already noted that the puncturing scheme that was used with the RSC variant is sub-optimal. Though the results that were achieved are quite good, optimizing the puncturing patterns according to the criterion given by [Hag88] could improve them further. In his work, [BR13] hinted that punctured turbo codes can be used with HARQ, but does not include them in his work either. So in general, different classes of codes might be considered.

In the capture model presented in Chapter 6, it is assumed that the reader uses coherent detection of the tag signals. Especially FM0 though, is well suited for incoherent detection. An adaptation of the capture model for incoherent detection would thus be desirable. If this proves to be impractical, different FEC schemes, and their influence on the capture probability, can always be evaluated in simulations. Also, it is assumed in the model that the reader is completely unaware of the collisions and does not actively attempt to resolve then. Since various [ALR10, KBDK11, BKDK12] proposals have been made to

actively resolve tag collisions, it would also be desirable to evaluate in how far FEC can be beneficial in these systems.

In the end, the proposed schemes should be deployed in an actual RFID system, so the results presented in this work can be backed up with measurements. If the identification times of a tag population in a bulk-reading scenario can be improved or if the modified system shows a faster read range, or an improved reliability under bad conditions, the proposed schemes can be deemed a complete success.

Bibliography

[Abr70] N. Abramson, "The ALOHA system: Another alternative for computer communications," in *Proceedings of the November 17-19, 1970, Fall Joint Computer Conference*, ser. AFIPS '70 (Fall). New York, NY, USA: ACM, 1970, pp. 281–285.

[ALR10] C. Angerer, R. Langwieser, and M. Rupp, "RFID Reader Receivers for Physical Layer Collision Recovery," *Communications, IEEE Transactions on*, vol. 58, no. 12, p. 3526–3537, Dec. 2010.

[BCJR74] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate (corresp.)," *Information Theory, IEEE Transactions on*, vol. 20, no. 2, p. 284–287, Mar. 1974.

[BDMP97] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "A soft-input soft-output APP module for iterative decoding of concatenated codes," *Communications Letters, IEEE*, vol. 1, no. 1, pp. 22–24, Jan. 1997.

[BDVA11] M. Bueno-Delgado and J. Vales-Alonso, "On the optimal frame-length configuration on real passive RFID systems," *Journal of Network and Computer Applications*, vol. 34, no. 3, p. 864–876, 2011, RFID Technology, Systems, and Applications.

[BG96] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *Communications, IEEE Transactions on*, vol. 44, no. 10, p. 1261–1271, Oct. 1996.

[BGT93] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *Communications, 1993. ICC 93. Geneva. Technical Program, Conference*

Record, IEEE International Conference on, vol. 2, May 1993, p. 1064–1070 vol.2.

[BKDK12] A. Bletsas, J. Kimionis, A. Dimitriou, and G. Karystinos, "Single-antenna coherent detection of collided FM0 RFID signals," *Communications, IEEE Transactions on*, vol. 60, no. 3, p. 756–766, 2012.

[BLW09] Y. Y. Bo Li and J. Wang, "Anti-collision issue analysis in Gen2 protocol – anti-collision issue analysis considering capture effect," Auto-ID Labs, Tech. Rep., Mar. 2009. [Online]. Available: http://www.autoidlabs.org/uploads/media/AUTOIDLABS-WP-HARDWARE-047.pdf

[BM96] S. Benedetto and G. Montorsi, "Iterative decoding of serially concatenated convolutional codes," *Electronics Letters*, vol. 32, no. 13, pp. 1186–1188, Jun. 1996.

[Bos98] M. Bossert, *Kanalcodierung : mit 36 Tabellen und 211 Beispielen*, 2nd ed., ser. Informationstechnik. Stuttgart: Teubner, 1998.

[Boy14] C. S. Boyer, "RFID: A communication system perspective," Ph.D. dissertation, University of Washington, Apr. 2014. [Online]. Available: http://hdl.handle.net/1773/25480

[BR12] C. Boyer and S. Roy, "Coded QAM backscatter modulation for RFID," *Communications, IEEE Transactions on*, vol. 60, no. 7, p. 1925–1934, 2012.

[BR13] ——, "Hybrid ARQ improvements for RFID systems," in *RFID (RFID), 2013 IEEE International Conference on*, 2013, p. 223–230.

[Bre59] D. G. Brennan, "Linear diversity combining techniques," *Proceedings of the IRE*, vol. 47, no. 6, p. 1075–1102, Jun. 1959.

[Bun10] Bundesnetzagentur, "Allgemeinzuteilung von Frequenzen in den Frequenzbereichen 865–868 MHz und 2446–2454 MHz für Funkanwendungen für Identifizierungszwecke; („Radio Frequency Identification Applications", RFID)," online, Mar. 2010. [Online]. Available: http://www.bundesnetzagentur.de/SharedDocs/Downloads/

DE/Sachgebiete/Telekommunikation/Unternehmen_Institutionen/Frequenzen/Allgemeinzuteilungen/RFID_Vfg62010pdf

[BW08] M. Buettner and D. Wetherall, "An empirical study of UHF RFID performance," in *Proceedings of the 14th ACM international conference on Mobile computing and networking*, ser. MobiCom '08. New York, NY, USA: ACM, 2008, pp. 223—-234.

[CC84] R. Comroe and J. Costello, D.J., "ARQ Schemes for Data Transmission in Mobile Radio Systems," *Selected Areas in Communications, IEEE Journal on*, vol. 2, no. 4, p. 472–481, 1984.

[CFRU01] S.-Y. Chung, J. Forney, G.D., T. Richardson, and R. Urbanke, "On the design of low-density parity-check codes within 0.0045 db of the shannon limit," *Communications Letters, IEEE*, vol. 5, no. 2, pp. 58–60, Feb 2001.

[CFV00] D. Chizhik, G. Foschini, and R. Valenzuela, "Capacities of multi-element transmit and receive antennas: Correlations and keyholes," *Electronics Letters*, vol. 36, no. 13, pp. 1099–1100, 2000.

[Cha85] D. Chase, "Code combining–a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *Communications, IEEE Transactions on*, vol. 33, no. 5, p. 385–393, 1985.

[CP73] M. W. Cardullo and W. Parks, "Transponder apparatus and system," US Patent US3 713 148 A, Jan. 23, 1973. [Online]. Available: https://www.google.com/patents/US3713148

[CYX+09] W. Che, Y. Yang, C. Xu, N. Yan, X. Tan, Q. Li, H. Min, and J. Tan, "Analysis, design and implementation of semi-passive Gen2 tag," in *RFID, 2009 IEEE International Conference on*, Apr. 2009, p. 15–19.

[DHK+11] H. Dong, J. He, A. Kang, C. Zhang, and T. Han, "Implementation of error-correcting encoded SAW RFID tags," in *Ultrasonics Symposium (IUS), 2011 IEEE International*, 2011, p. 581–583.

[Dob08] D. M. Dobkin, *The RF in RFID : passive UHF RFID in practice*, ser. Communications engineering series. Amsterdam: Elsevier Newnes, 2008.

[EPC07] *Tag Class Definitions*, EPCglobal Std., Nov. 2007. [Online]. Available: http://www.gs1.org/docs/epcglobal/TagClassDefinitions_1_0-whitepaper-20071101.pdf

[EPC08] *EPC^{TM} Radio-Frequency Identity Protocols, Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz–960 MHz*, EPCglobal Std., Oct. 2008. [Online]. Available: http://www.gs1.org/gsmp/kc/epcglobal/uhfc1g2/uhfc1g2_1_2_0-standard-20080511.pdf

[EPC13] *EPC^{TM} Radio-Frequency Identity Protocols Generation-2 UHF RFID Specification for RFID Air Interface Protocol for Communications at 860 MHz – 960 MHz Version 2.0.0 Ratified*, EPCglobal Std., Nov. 2013. [Online]. Available: http://www.gs1.org/sites/default/files/docs/uhfc1g2/uhfc1g2_2_0_0_standard_20131101.pdf

[ETS07] "Electromagnetic compatibility and radio spectrum matters (ERM); technical characteristics of RFID in the UHF band; system reference document for radio frequency identification (RFID) equipment; part 1: RFID equipment operating in the range from 865 MHz to 868 MHz," ETSI, Tech. Rep. TR 102 649-1, Apr. 2007, 2007. [Online]. Available: http://www.etsi.org/deliver/etsi_tr/102600_102699/10264901/01.01.01_60/tr_10264901v010101p.pdf

[ETS10] *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short range devices; Radio equipment to be used in the 1 GHz to 40 GHz frequency range; Part 2: Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive*, ETSI Std. ETSI EN 300 440-2, Rev. V1.4.1, Aug. 2010. [Online]. Available: http://www.etsi.org/deliver/etsi_en%5C300400_300499%5C30044002%5C01.04.01_60%5Cen_30044002v010401p.pdf

[ETS11] *Electromagnetic compatibility and Radio spectrum Matters (ERM); Radio Frequency Identification Equipment operating in the band*

865 MHz to 868 MHz with power levels up to 2 W; Part 2: Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive, ETSI Std. ETSI EN 302 208-2, Rev. V1.4.1, Nov. 2011. [Online]. Available: http://www.etsi.org/deliver/etsi_en%5C302200_302299%5C30220802%5C01.04.01_60%5Cen_30220802v010401p.pdf

[ETS12] *Electromagnetic compatibility and Radio spectrum Matters (ERM); Short Range Devices (SRD); Radio equipment to be used in the 25 MHz to 1 000 MHz frequency range with power levels ranging up to 500 mW; Part 2: Harmonized EN covering essential requirements under article 3.2 of the R&TTE Directive*, ETSI Std. EN 300 220-2, Rev. V2.4.1, May 2012. [Online]. Available: http://www.etsi.org/deliver/etsi_en%5C300200_300299%5C30022002%5C02.04.01_60%5Cen_30022002v020401p.pdf

[ETS14] *Electromagnetic compatibility and Radio spectrum Matters (ERM); Radio Frequency Identification Equipment operating in the band 865 MHz to 868 MHz with power levels up to 2 W and in the band 915 MHz to 921 MHz with power levels up to 4 W; Part 2: Harmonized EN covering the essential requirements of article 3.2 of the R&TTE Directive*, ETSI Std. Draft ETSI EN 302 208-2, Rev. V2.1.0, Jun. 2014. [Online]. Available: http://www.etsi.org/deliver/etsi_en%5C302200_302299%5C30220802%5C02.01.00_20%5Cen_30220802v020100a.pdf

[Fel50] W. Feller, *An introduction to probability theory and its applications*, 2nd ed., ser. 1950. New York, NY: Wiley, 1950. [Online]. Available: http://archive.org/details/AnIntroductionToProbabilityTheoryAndItsApplicationsVolume1

[Fin08] K. Finkenzeller, *RFID-Handbuch : Grundlagen und praktische Anwendungen von Transpondern, kontaktlosen Chipkarten und NFC / Klaus Finkenzeller*, 5th ed. Munich: Hanser Verlag, 2008.

[Flo07] C. Floerkemeier, "Bayesian transmission strategy for framed ALOHA based RFID protocols," in *RFID, 2007. IEEE International Conference on*, Mar. 2007, p. 228–235.

[For65] G. D. Forney, "Concatenated codes," Ph.D. dissertation, Massachusetts Institute of Technology. Dept. of Electrical Engineering, 1965. [Online]. Available: http://hdl.handle.net/1721.1/13449

[Fri96] B. Friedrichs, *Kanalcodierung : Grundlagen und Anwendungen in modernen Kommunikationssystemen*, ser. Information und Kommunikation. Berlin: Springer, 1996.

[FS09] C. Floerkemeier and S. Sarma, "RFIDSim – a physical and logical layer simulation engine for passive RFID," *Automation Science and Engineering, IEEE Transactions on*, vol. 6, no. 1, p. 33–43, 2009.

[GD07] J. Griffin and G. Durgin, "Link envelope correlation in the backscatter channel," *Communications Letters, IEEE*, vol. 11, no. 9, p. 735–737, Sep. 2007.

[GD08] ——, "Gains for RF tags using multiple antennas," *Antennas and Propagation, IEEE Transactions on*, vol. 56, no. 2, p. 563–570, Feb. 2008.

[GD09] ——, "Complete link budgets for backscatter-radio and RFID systems," *Antennas and Propagation Magazine, IEEE*, vol. 51, no. 2, p. 11–25, Apr. 2009.

[GS97] C. M. Grinstead and J. L. Snell, *Introduction to Probability.* American Mathmatical Society, 1997, vol. 2. [Online]. Available: http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf

[Hag88] J. Hagenauer, "Rate-compatible punctured convolutional codes (rcpc codes) and their applications," *Communications, IEEE Transactions on*, vol. 36, no. 4, p. 389–400, 1988.

[Hag97] ——, "The Turbo Principle: Tutorial Introduction and State of the Art," in *International Symposium on Turbo-Codes, Proceedings*

of, Sep. 1997. [Online]. Available: http://www.lnt.e-technik.tu-muenchen.de/veroeffentlichungen/tut.ps.gz

[Has93] H. Hashemi, "The indoor radio propagation channel," *Proceedings of the IEEE*, vol. 81, no. 7, p. 943–968, Jul. 1993.

[HKS89] I. M. I. Habbab, M. Kavehrad, and C. E. W. Sundberg, "ALOHA with capture over slow and fast fading radio channels with coding and diversity," *Selected Areas in Communications, IEEE Journal on*, vol. 7, no. 1, p. 79–88, 1989.

[HL99] P. Hoeher and J. Lodge, ""turbo DPSK" iterative differential PSK demodulation and channel decoding," *Communications, IEEE Transactions on*, vol. 47, no. 6, p. 837–843, Jun. 1999.

[HOP96] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *Information Theory, IEEE Transactions on*, vol. 42, no. 2, p. 429–445, Mar. 1996.

[Imp12a] Impinj, "Datasheet : Indy® R1000 Reader Chip (IPJ-R1000)," 2012, last checked: 2013-12-02. [Online]. Available: http://www.impinj.com/Documents/Reader_Chips/Indy_R1000_Datasheet_v2_2/

[Imp12b] ——, "Datasheet : Indy® R2000 Reader Chip (IPJ-R2000)," 2012, last checked: 2013-12-02. [Online]. Available: http://www.impinj.com/Documents/Reader_Chips/Indy_R2000_Datasheet/

[Imp12c] ——, "Datasheet : Indy® R500 Reader Chip (IPJ-R500)," 2012, last checked: 2013-12-02. [Online]. Available: http://www.impinj.com/Documents/Reader_Chips/Indy_R500_Datasheet/

[Imp12d] ——, "Monza® 4 Tag Chip Datasheet," 2012, last checked: 2013-12-02. [Online]. Available: https://support.impinj.com/hc/en-us/articles/202756908-Monza-4-RFID-Tag-Chip-Datasheet

[Imp14a] ——, "Monza® 5 Tag Chip Datasheet," Mar. 2014, last checked: 2014-10-15. [Online]. Available: https://support.impinj.com/hc/en-us/article_attachments/200775538/Monza_5_Tag_Chip_Datasheet_3-24-14.pdf

[Imp14b] ——, "Monza® X-2K Dura Datasheet," Mar. 2014, last checked: 2014-10-15. [Online]. Available: https://support.impinj.com/hc/en-us/article_attachments/200775438/Monza_X-2K_Dura_Datasheet_3-24-14.pdf

[Int10] Intelleflex, "The intelleflex XC3 technology™ platform," 2010, last checked: 2013-10-15. [Online]. Available: http://www.intelleflex.com/downloads/white-papers/Intelleflex-XC3-ISO-C3-White-Paper.pdf

[ISO13] *Information technology — Radio frequency identification for item management — Part 63: Parameters for air interface communications at 860 MHz to 960 MHz Type C*, ISO/IEC Std. 18 000-63, 2013.

[JC09] K. Y. Jeon and S. H. Cho, "A RFID EPC C1 Gen2 System with Channel Coding Capability in AWGN Noise Environments," *IEICE Transactions on Communications*, vol. E92.B, no. 2, p. 608–611, Feb. 2009.

[JKC08] K. Y. Jeon, M. S. Kang, and S. H. Cho, "A performance enhanced Miller modulated RFID EPC C1 Gen2 system," in *Advanced Technologies for Communications, 2008. ATC 2008. International Conference on*, Oct. 2008, p. 102–105.

[Joe06] I. Joe, "A low-power hybrid ARQ scheme for the RFID system," in *High Performance Computing and Communications*, ser. Lecture Notes in Computer Science, M. Gerndt and D. Kranzlmüller, Eds. Springer Berlin Heidelberg, 2006, vol. 4208, p. 535–541.

[Kal95] S. Kallel, "Complementary punctured convolutional (CPC) codes and their applications," *Communications, IEEE Transactions on*, vol. 43, no. 6, pp. 2005–2009, Jun. 1995.

[KBDK11] J. Kimionis, A. Bletsas, A. Dimitriou, and G. Karystinos, "Inventory time reduction in Gen2 with single-antenna separation of FM0 RFID signals," in *RFID-Technologies and Applications (RFID-TA), 2011 IEEE International Conference on*, 2011, p. 494–501.

[KIS03] D. Kim, M. Ingram, and J. Smith, W.W., "Measurements of small-scale fading and path loss for long range RF tags," *Antennas and Propagation, IEEE Transactions on*, vol. 51, no. 8, pp. 1740 – 1749, Aug. 2003.

[KST04] R. Koetter, A. Singer, and M. Tüchler, "Turbo equalization," *Signal Processing Magazine, IEEE*, vol. 21, no. 1, pp. 67–80, Jan. 2004.

[Lan05] J. Landt, "The history of RFID," *Potentials, IEEE*, vol. 24, no. 4, pp. 8 – 11, Oct.-Nov. 2005.

[Lee94] L. H. C. Lee, "New rate-compatible punctured convolutional codes for viterbi decoding," *Communications, IEEE Transactions on*, vol. 42, no. 12, p. 3073–3079, 1994.

[LGS09] A. Lazaro, D. Girbau, and D. Salinas, "Radio link budgets for UHF RFID on multipath environments," *Antennas and Propagation, IEEE Transactions on*, vol. 57, no. 4, p. 1241–1251, Apr. 2009.

[LGSV09] A. Lazaro, D. Girbau, D. Salinas, and R. Villarino, "Channel modelling for UHF RFID," in *Microwave Conference, 2009. EuMC 2009. European*, Oct. 2009, p. 1105–1108.

[LL90] C. Lau and C. Leung, "Capture models for mobile packet radio networks," in *Communications, 1990. ICC '90, Including Supercomm Technical Sessions. SUPERCOMM/ICC '90. Conference Record., IEEE International Conference on*, 1990, p. 1226–1230 vol.30.

[Mau08] R. Maunder, "Matlab exit charts," online, 2008, last checked: 2014-11-10. [Online]. Available: http://users.ecs.soton.ac.uk/rm/resources/matlabexit/

[MIT02] *860MHz–930MHz Class I Radio Frequency Identification Tag Radio Frequency & Logical Communication Interface Specification Candidate Recommendation*, MIT Auto-ID Center Std., Nov. 2002. [Online]. Available: http://www.gs1.org/docs/epcglobal/standards/specs/860MHz_930_MHz_Class_1_RFID_Tag_Radio_Frequency_Logical_Communication_Interface_Specification.pdf

[MIT03] *Draft protocol specification for a 900 MHz Class 0 Radio Frequency Identification Tag*, MIT Auto-ID Center Std., Feb. 2003. [Online]. Available: http://www.gs1.org/docs/epcglobal/standards/specs/900_MHz_Class_0_RFIDTag_Specification.pdf

[MSG14] D. Merget, G. Smietanka, and J. Goetze, "UHF RFID transmission with soft-input BCH decoding," in *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, March 2014, pp. 199–202.

[MZ11] R. Morelos-Zaragoza, "On error performance improvements of passive UHF RFID systems via syndrome decoding," in *Internet of Things (iThings/CPSCom), 2011 International Conference on and 4th International Conference on Cyber, Physical and Social Computing*, 2011, p. 127–130.

[MZ12] ——, "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*, 2012, p. 36–40.

[NR06] P. Nikitin and K. Rao, "Performance limitations of passive UHF RFID systems," in *Antennas and Propagation Society International Symposium 2006, IEEE*, Jul. 2006, p. 1011–1014.

[NRM07] P. Nikitin, K. V. S. Rao, and R. Martinez, "Differential RCS of RFID tag," *Electronics Letters*, vol. 43, no. 8, pp. 431–432, 2007.

[NXP13] NXP, "Product short data sheet UCODE 7," May 2013.

[Obe73] F. Oberhettinger, *Fourier transforms of distributions and their inverses : a collection of tables*, ser. Probability and mathematical statistics ; 16. New York: Academic Press, 1973.

[PS08] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. Boston: McGraw-Hill, 2008.

[RD06] K. M. Ramakrishnan and D. D. Deavours, "Performance benchmarks for passive UHF RFID tags," in *Proceedings 13th GI/ITG Conference on Measuring, Modelling and Evaluation of Computer*

and Communication Systems, MMB 2006, March 27-29, 2006, Nürnberg, Germany, R. German and A. Heindl, Eds. VDE Verlag, 2006.

[Rob75] L. G. Roberts, "ALOHA packet system with and without slots and capture," *SIGCOMM Comput. Commun. Rev.*, vol. 5, no. 2, pp. 28—-42, Apr. 1975.

[Rob09] M. Roberti, "Mojix takes passive UHF RFID to a new level," Apr. 2009, Accessed: 2013-07-27. [Online]. Available: http://www.rfidjournal.com/articles/view?4019

[RVH95] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *Communications, 1995. ICC '95 Seattle, 'Gateway to Globalization', 1995 IEEE International Conference on*, vol. 2, Jun. 1995, p. 1009–1013.

[Sch83] F. Schoute, "Dynamic frame length ALOHA," *Communications, IEEE Transactions on*, vol. 31, no. 4, p. 565–568, 1983.

[Sch12] A. Schantin, "Forward Error Correction in Long-Range RFID Systems," in *Smart Objects, Systems and Technologies (SmartSysTech), Proceedings of 2012 European Conference on*. Munich, Germany: VDE Verlag, Jun. 2012, p. 1–6.

[Sch13] ——, "Iterative decoding of baseband and channel codes in a long-range RFID system," in *Industrial Technology (ICIT), 2013 IEEE International Conference on*. Cape Town, South Africa: IEEE, Feb. 2013, p. 1671–1676.

[SD06] M. Simon and D. Divsalar, "Some interesting observations for certain line codes with application to RFID," *Communications, IEEE Transactions on*, vol. 54, no. 4, pp. 583 – 586, Apr. 2006.

[SG12] G. Smietanka and J. Goetze, "Performance of BCH codes and transmit diversity for MISO UHF RFID communication," in *Smart SysTech 2012; European Conference on Smart Objects, Systems and*

Technologies; Proceedings of. Munich, Germany: VDE-Verlag, Jun. 2012.

[SGG13] 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*, 2013, p. 1666–1670.

[SR13] A. Schantin and C. Ruland, "Retransmission strategies for RFID systems using the EPC protocol," in *RFID Technologies and Applications, 2013 IEEE International Conference on*, 2013.

[SR14] ——, "A Bit-Error based capture model for EPCglobal RFID systems," in *Communication, Control, and Computing (Allerton), 2014 52nd Annual Allerton Conference on*, 2014.

[Sto48] H. Stockman, "Communication by means of reflected power," *Proceedings of the IRE*, vol. 36, no. 10, p. 1196–1204, Oct. 1948.

[tB99] S. ten Brink, "Convergence of iterative decoding," *Electronics Letters*, vol. 35, no. 10, p. 806–808, May 1999.

[tB01] ——, "Convergence behavior of iteratively decoded parallel concatenated codes," *Communications, IEEE Transactions on*, vol. 49, no. 10, p. 1727–1737, Oct. 2001.

[Vog02a] H. Vogt, "Multiple object identification with passive RFID tags," *IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2002.

[Vog02b] ——, "Efficient Object Identification with Passive RFID tags," in *Pervasive Computing*, ser. Lecture Notes in Computer Science, F. Mattern and M. Naghshineh, Eds. Springer Berlin / Heidelberg, 2002, vol. 2414, p. 98–113, 10.1007/3-540-45866-2_9.

[Wal73] C. Walton, "Electronic identification & recognition system," US Patent US3 752 960 A, Aug. 14, 1973. [Online]. Available: https://www.google.com/patents/US3752960

[Wal74] ——, "Electronic recognition and identification system," US Patent US3 816 708 A, Jun. 11, 1974. [Online]. Available: https://www.google.com/patents/US3816708

[Wei14] E. W. Weisstein, "Normal product distribution. From MathWorld – A Wolfram Web Resource," Mar. 2014. [Online]. Available: http://mathworld.wolfram.com/NormalProductDistribution.html

[WP99] X. Wang and H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *Communications, IEEE Transactions on*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.

[WW99] J.-H. Wen and J.-W. Wang, "A recursive solution to an occupancy problem resulting from tdm radio communication application," *Applied Mathematics and Computation*, vol. 101, no. 1, p. 1–3, 1999.

[ZY11] L. Zhu and T.-S. P. Yum, "A critical survey and analysis of RFID anti-collision mechanisms," *Communications Magazine, IEEE*, vol. 49, no. 5, p. 214–221, May 2011.