



# LUND UNIVERSITY

## Hardware-Conscious Wireless Communication System Design

Sarajlic, Muris

2019

*Document Version:*

Publisher's PDF, also known as Version of record

[Link to publication](#)

*Citation for published version (APA):*

Sarajlic, M. (2019). *Hardware-Conscious Wireless Communication System Design*. [Doctoral Thesis (compilation), Communications Engineering]. Lund University.

*Total number of authors:*

1

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# **Hardware-Conscious Wireless Communication System Design**

Muris Sarajlić

Lund 2019

Department of Electrical and Information Technology  
Lund University  
Box 118, SE-221 00 LUND  
SWEDEN

This thesis is set in Computer Modern 10pt  
with the L<sup>A</sup>T<sub>E</sub>X Documentation System

Series of licentiate and doctoral theses  
No. 118  
ISSN 1654-790X  
ISBN 978-91-7753-914-8 (print)  
ISBN 978-91-7753-915-5 (digital)

© Muris Sarajlić 2019  
Printed in Sweden by *Tryckeriet i E-huset*, Lund.  
January 2019.

*One must still have chaos in oneself to be able to give birth to a dancing star.*

Friedrich Nietzsche, Thus Spoke Zarathustra



# Popular Science

All of us have at a certain point felt the excruciating pain of our mobile phone shutting down in the middle of something important because the battery “died”. Batteries inevitably do need to get recharged, but everyone would love it if they could at least last longer. This is definitely doable, and there are many things that can be improved to extend battery life. In addition to improving the battery itself, we can also consider doing something about the things draining the battery. One of those things is a little piece of electronics, we can call it a modem, attached to the phone antenna. The modem will amplify the very weak signal that is captured by the antenna, transform it into numbers, and then perform a lot of calculations on those numbers to extract the information that is contained in the signal. Energy from the battery is needed for performing all those actions. In case of bad reception, the modem has to work extra hard to extract the information and ends up draining a lot of energy from the battery in the process. When the reception is good, the modem doesn’t really need to struggle or spend that much energy. However, because we never know when bad reception will hit, it’s easiest to make the modem work hard all the time just to be on the safe side; this is how modems are usually designed. The problem with this strategy is that bad reception does not really hit *that* often. This means that for most of the time, the reception is good but the modem is nevertheless working hard, spending a lot of energy for nothing. If it were possible to sense the quality of the reception and make the modem work harder or less hard based on the signal quality, a lot of energy could be saved and the battery would last longer! One part of the topics analyzed in this thesis is about how different parts of the modem - the amplifier, the converter to numbers, the calculator extracting the information from the numbers - adapt their actions to the quality of the signal so that they go without burning more energy than necessary.

Other parts of this thesis are also about efficient design, but the focus is on the design of the base station - a piece of wireless communication equipment that usually sits on a building somewhere and makes sure all the mobile phones

around it are served with the information they want. Recently, it was suggested that the base station should be equipped with an extreme number of antennas because this way it can serve more users than usual, or provide the existing users with a better user experience, or consume less energy which is good news for both the environment and the owner of the base station who pays the electric utility bill. As one of many good things claimed about base stations with lots of antennas, it was claimed that the converter from a signal to numbers - the same one as in the phone modem - can be made simpler because of the benefits provided by a large number of antennas and in this way save energy in the base station. One paper in the thesis investigates this claim in depth.

Everything said so far about putting many antennas on a base station sounds so good, so is there any downside to doing this? Of course there is. With a lot of antennas, the design of the base station becomes tricky. The connections between the antennas and the central computer handling all the signals carry a lot of data, sometimes too much for the central computer to handle at once. One paper in the thesis proposes a radical idea - removing the central computer altogether, connecting the antennas in a chain and letting them do all the processing of signals between themselves. This strategy proves to perform almost as good as if it was done on a central computer, with an important difference that now the data flows can be handled easily!

The last paper in the thesis is about pairs of wireless devices wanting to communicate directly and other devices in their vicinity helping those messages get across by relaying them. If these helpers, or relays as they are technically called, team up and relay the messages together, the quality of those messages improves. The quality becomes very good if there are many relays. Unfortunately, there is a caveat here too. The “teaming up” needs a lot of messages to be exchanged internally between the relays, which will consume resources. So, in the last paper it is proposed that relays only team up in smaller groups and do the relaying independently of other relay groups. In this way, the amount of resources spent on internal message exchange is reduced. The analysis shows that the quality of relayed messages is almost the same as when all relays are in one big group.

# Abstract

The work at hand is a selection of topics in efficient wireless communication system design, with topics logically divided into two groups.

One group can be described as *hardware designs conscious of their possibilities and limitations*. In other words, it is about hardware that chooses its configuration and properties depending on the performance that needs to be delivered and the influence of external factors, with the goal of keeping the energy consumption as low as possible. Design parameters that trade off power with complexity are identified for analog, mixed signal and digital circuits, and implications of these tradeoffs are analyzed in detail. An analog front end and an LDPC channel decoder that adapt their parameters to the environment (e.g. fluctuating power level due to fading) are proposed, and it is analyzed how much power/energy these environment-adaptive structures save compared to non-adaptive designs made for the worst-case scenario. Additionally, the impact of ADC bit resolution on the energy efficiency of a massive MIMO system is examined in detail, with the goal of finding bit resolutions that maximize the energy efficiency under various system setups.

In another group of themes, one can recognize systems where the system architect was *conscious of fundamental limitations stemming from hardware*. Put in another way, in these designs there is no attempt of tweaking or tuning the hardware. On the contrary, system design is performed so as to work around an existing and unchangeable hardware limitation. As a workaround for the problematic centralized topology, a massive MIMO base station based on the daisy chain topology is proposed and a method for signal processing tailored to the daisy chain setup is designed. In another example, a large group of cooperating relays is split into several smaller groups, each cooperatively performing relaying independently of the others. As cooperation consumes resources (such as bandwidth), splitting the system into smaller, independent cooperative parts helps save resources and is again an example of a workaround for an inherent limitation.

From the analyses performed in this thesis, promising observations about



hardware consciousness can be made. Adapting the structure of a hardware block to the environment can bring massive savings in energy, and simple workarounds prove to perform almost as good as the inherently limited designs, but with the limitation being successfully bypassed. As a general observation, it can be concluded that **hardware consciousness pays off**.

# Preface

Taking a look back at the timeline of my education, it is not easy to precisely pinpoint *when* or *how* I developed a fascination with wireless communications. It could have begun when I took the introductory courses in signal and systems and digital communications and saw those abstract mathematical concepts from the first couple of years of engineering studies being used to both create and deconstruct signals that convey information. Regardless of when and how the whole thing started, I am quite sure *why* I found wireless so appealing: there was something deeply mystical about the possibility of successfully communicating, sometimes over vast distances, without any need for a physical connection.

The initial fascination was followed, naturally, by a desire to know more. And I didn't want to focus on a very specific topic; I wanted to know something about everything, from adaptive equalization to how and why error control codes work. The quest for knowledge led me from my native Bosnia and Herzegovina to Sweden and Lund. Here, at the Department of Electrical and Information Technology at LTH, I got the opportunity to do a PhD, which, thanks to many projects that the department was involved in, presented me with a broad choice of topics to look into. It was exactly how I wanted it to be. Very importantly, PhD studies were not just a learning experience in an academic sense. They were also a process of learning about myself; a period of life-changing personal development.

The results of research activities performed during my PhD studies are collected in this thesis. The text itself comprises two parts. In the first part, a general overview of the covered topics is given, with the emphasis on the underlying leitmotif: hardware-consciousness. Here, hardware-consciousness relates to two points of view in system design: **design of hardware-conscious wireless systems** (those that dynamically tune the parameters of their constituent hardware parts to meet a satisfactory balance between performance and complexity/power consumption) and **hardware-conscious design of wireless systems** (keeping in mind hardware-related limitations when designing the

system with the goal of circumventing those limitations in an efficient way).

The second part of the thesis consists of a selection of original scientific publications written during my PhD studies, each related to one of the topics covered:

- [1] Muris Sarajlić, Liang Liu, Henrik Sjöland and Ove Edfors, “Low Power Receiver Front Ends: Scaling Laws and Applications,” submitted to *IEEE Transactions on Wireless Communications*, Jan. 2019.
- [2] Muris Sarajlić, Liang Liu and Ove Edfors, “When are Low Resolution ADCs Energy Efficient in Massive MIMO?,” in *IEEE Access*, vol. 5, pp. 14837-14853, July 2017.
- [3] Muris Sarajlić, Liang Liu and Ove Edfors, “Modified Forced Convergence Decoding of LDPC Codes with Optimized Decoder Parameters,” *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, PRC, Sept. 2015.
- [4] Muris Sarajlić, Fredrik Rusek, Jesús Rodríguez Sánchez, Liang Liu and Ove Edfors, “Fully Decentralized Approximate Zero-Forcing Precoding for Massive MIMO Systems,” in *IEEE Wireless Communication Letters*, doi: 10.1109/LWC.2019.2892044, Jan. 2019.
- [5] Muris Sarajlić, Liang Liu, Fredrik Rusek, Farhana Sheikh and Ove Edfors, “Impact of Relay Cooperation on the Performance of Large-scale Multipair Two-way Relay Networks,” *IEEE 37th Global Communications Conference (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018.

During the course of my PhD studies, I have contributed (either as principal author or as a secondary collaborator) to the following publications, which have not been included in this thesis:

- [6] Muris Sarajlić, Liang Liu and Ove Edfors, “Reducing the Complexity of LDPC Decoding Algorithms: An Optimization-Oriented Approach,” *IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, Washington DC, USA, Sept. 2014.
- [7] Muris Sarajlić, Liang Liu and Ove Edfors, “An Energy Efficiency Perspective on Massive MIMO Quantization,” *50th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, Nov. 2016.
- [8] João Vieira, Erik Leitinger, Muris Sarajlić, Xuhong Li and Fredrik Tufveson, “Deep Convolutional Neural Networks for Massive MIMO Fingerprint-Based Positioning,” *IEEE 28th Annual International Symposium on Per-*

*sonal, Indoor, and Mobile Radio Communications (PIMRC)*, Montreal, QC, Canada, Oct. 2017.

- [9] Jesús Rodríguez Sánchez, Fredrik Rusek, Muris Sarajlić, Ove Edfors and Liang Liu, “Fully Decentralized Massive MIMO Detection Based on Recursive Methods”, *IEEE Workshop on Signal Processing Systems (SiPS)*, Cape Town, South Africa, Oct. 2018.



# Acknowledgements

There are many people without whose support, patience, trust and love this thesis might have never been written.

First and foremost, I would like to thank my supervisor, Professor Ove Edfors, for believing in me and giving me unconditional support throughout my PhD studies. Thank you, Ove, for providing me with this unique learning opportunity, and for all the advice you have given me along the way.

To my second supervisor, Associate Professor Liang Liu, thank you for always being willing to help, to discuss whatever issues that would come up, and to give useful feedback and suggestions. Knowing that I could always rely on you meant a lot.

To the great man, teacher and friend, Associate Professor Fredrik Rusek, a big thank you for all the time and energy you have generously invested in our collaboration and for the support you gave me when I needed it the most. I am probably only yet to realize how much I learned by working with you.

To my course mate, office mate and great friend, João Vieira, thank you for being a big support and a true inspiration.

To all the colleagues and friends from the department, thank you for all the fikas, lunches, and all the moments that made this experience enriching in so many different ways. And to all the friends in Sweden, Bosnia and everywhere around the world, thank you for all your love.

Finally, to my parents: thank you for everything. Without your love and encouragement I would never have been in the place where I am now.

Muris Sarajlić



# List of Acronyms and Abbreviations

<b>ABE</b>	Analog Back End
<b>ADC</b>	Analog-to-Digital Converter
<b>AFE</b>	Analog Front End
<b>AP</b>	Access Point
<b>AQN</b>	Additive Quantization Noise
<b>AU</b>	Antenna Unit
<b>BB</b>	Baseband
<b>BS</b>	Base Station
<b>DL</b>	Downlink
<b>EE</b>	Energy Efficiency
<b>FC</b>	Forced Convergence
<b>FoM</b>	Figure-of-Merit
<b>IIP3</b>	Input-Referred Third-Order Intercept
<b>LDPC</b>	Low Density Parity Check
<b>LLR</b>	Log Likelihood Ratio
<b>LNA</b>	Low Noise Amplifier
<b>LTE</b>	Long-Term Evolution



<b>M2M</b>	Machine-to-Machine
<b>MaMI</b>	Massive MIMO
<b>MIMO</b>	Multiple Input, Multiple Output
<b>MSA</b>	Min-Sum Algorithm
<b>MSE</b>	Mean Square Error
<b>NGMN</b>	Next Generation Mobile Networks
<b>NR</b>	New Radio
<b>OOB</b>	Out-of Band
<b>OPEX</b>	Operational Expenditure
<b>PAPR</b>	Peak-to Average Power Ratio
<b>PQN</b>	Pseudoquantization Noise
<b>PSD</b>	Power Spectral Density
<b>RAN</b>	Radio Access Network
<b>RX</b>	Receiver
<b>SDMA</b>	Space Division Multiple Access
<b>SA</b>	Stochastic Approximation
<b>SE</b>	Spectral Efficiency
<b>SNR</b>	Signal-to-Noise Ratio
<b>SPA</b>	Sum-Product Algorithm
<b>TDD</b>	Time Division Duplex
<b>TX</b>	Transmitter
<b>UE</b>	User Equipment
<b>UL</b>	Uplink

# Contents

<b>Popular Science</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Preface</b>	<b>ix</b>
<b>Acknowledgements</b>	<b>xiii</b>
<b>List of Acronyms and Abbreviations</b>	<b>xv</b>
<b>Contents</b>	<b>xvii</b>
<b>I Introduction</b>	<b>1</b>
<b>1 Motivation and Outline</b>	<b>3</b>
1.1 The Energy Efficiency Issue . . . . .	4
1.2 Hardware-Consciousness . . . . .	7
1.3 Thesis Structure . . . . .	8
<b>2 Power-Performance Tradeoffs in Receivers</b>	<b>11</b>
2.1 Analog front end . . . . .	14
2.2 Analog-digital conversion . . . . .	18
2.3 Digital baseband . . . . .	28
<b>3 System Densification Under Hardware Constraints</b>	<b>35</b>
3.1 Massive MIMO . . . . .	35
3.2 Large-scale multipair two-way relay networks . . . . .	43
<b>4 Paper Summary and Discussion</b>	<b>47</b>

4.1	Research contributions . . . . .	47
4.2	Discussion and future work . . . . .	51
<b>II</b>	<b>Included Papers</b>	<b>61</b>
	<b>Low Power Receiver Front Ends: Scaling Laws and Applications</b>	<b>65</b>
1	Introduction . . . . .	67
2	Optimal power consumption of analog front ends . . . . .	69
3	Scaling laws of AFE power consumption . . . . .	72
4	Ramifications of the scaling laws . . . . .	78
5	Conclusion . . . . .	92
	<b>When are Low Resolution ADCs Energy Efficient in Massive MIMO?</b>	<b>101</b>
1	Introduction . . . . .	103
2	Preliminaries: ADC and AGC . . . . .	105
3	System model . . . . .	109
4	System sumrate . . . . .	112
5	Energy efficiency analysis . . . . .	116
6	Conclusion . . . . .	134
	<b>Modified Forced Convergence Decoding of LDPC Codes with Optimized Decoder Parameters</b>	<b>143</b>
1	Introduction . . . . .	145
2	Background . . . . .	145
3	Modified offset min-sum algorithm with layered scheduling and forced convergence . . . . .	146
4	Optimizing the LDPC decoder parameters . . . . .	147
5	Simulation and results . . . . .	154
6	Conclusion . . . . .	158
	Acknowledgment . . . . .	158
	<b>Fully Decentralized Approximate Zero-Forcing Precoding for Massive MIMO Systems</b>	<b>163</b>
1	Introduction . . . . .	165

---

2	Problem setup and system model . . . . .	166
3	Fully decentralized approximate zero-forcing . . . . .	167
4	Performance of the proposed algorithm. Implementation considerations . . . . .	173
5	Conclusion . . . . .	175
<b>Impact of Relay Cooperation on the Performance of Large-scale Multipair Two-way Relay Networks</b>		<b>181</b>
1	Introduction . . . . .	183
2	System setup . . . . .	184
3	System model . . . . .	186
4	Ergodic system sumrate calculation with per-group zero-forcing . . . . .	189
5	Analysis and discussion . . . . .	192
6	Conclusion . . . . .	195
	Acknowledgment . . . . .	196



## Part I

# Introduction



# Chapter 1

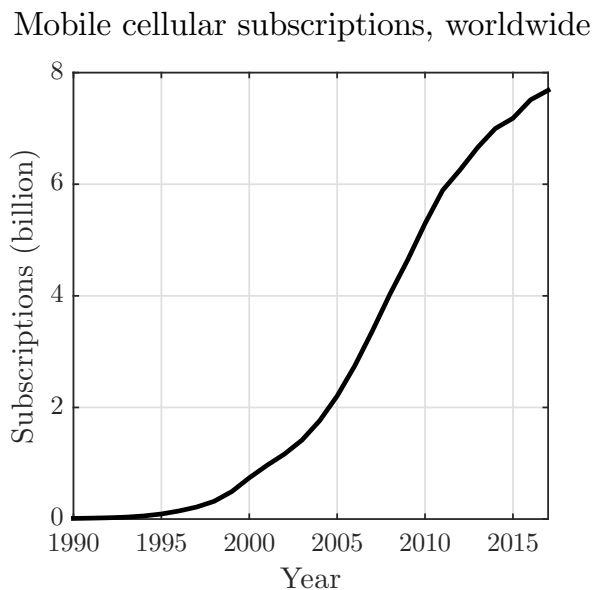
## Motivation and Outline

The explosive development of semiconductor technology, characterized by ever increasing computing power, ever reducing production costs and an exponential increase of the number of integrated circuit elements per unit area, as predicted by Moore [1], has been a key enabler of technological development in recent decades. More than that, it has been a stepping stone for ushering a new era in the history of human civilization, the one in which we find ourselves right now: the *Information Age*. The ways in which information is transferred, stored, accessed and - most importantly - used have a defining influence on the way we live our lives in the Information Age.

One of the most significant, as well as most rapid technological developments underlying the Information Age was the evolution of digital wireless communications. In the remainder of this text, the term “wireless communications” will be used to denote digital wireless communications, where wireless transmission techniques are employed in the context of transfer of digital information, in contrast with more traditional wireless technologies as FM radio. Motivation for (as well as consequence of) this brisk development can be found in the inherent traits of wireless communications: freedom of movement, expanded reach, and ease of use. A look at the evolution of the number of mobile cellular subscriptions, provided by World Bank [2] and shown in Fig. 1.1 reveals that in year 2017, the number of subscriptions was approximately equal to world population (7.68 versus 7.53 billions, respectively). Wireless communications are not only limited to personal use. Due to their ubiquitousness and flexibility, they are also being increasingly adopted to support communication between machines. In a white paper by Cisco [3] it is reported that in 2016, there were already 0.8 billion wireless machine-to-machine (M2M) connections, and it is also predicted that by 2021 this number will grow to 3.3 billions (cf.



Fig. 1.2).

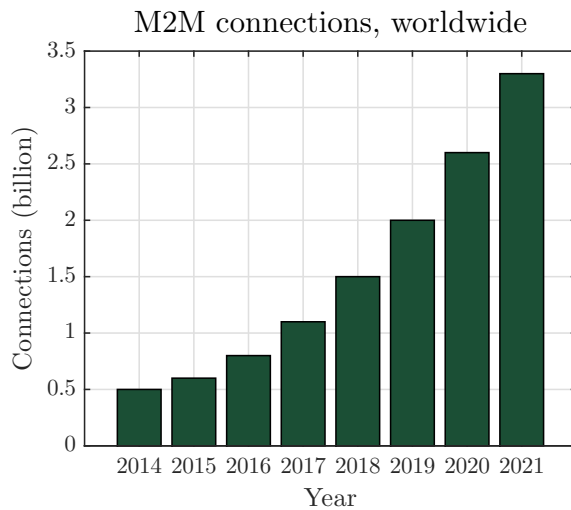


**Figure 1.1:** Evolution of the number of mobile cellular subscriptions worldwide [2]

The explosive growth of the number of wireless connections is also accompanied by an exponential increase of the amount of information transmitted wirelessly, with global mobile data traffic growing from 7 Exabytes ( $7 \times 10^{18}$  bytes) per month in 2016 to projected 49 Exabytes per month in 2021 [3]. The spatial data rate density, expressed in bits/s/m<sup>2</sup>, is projected to increase by a factor of 1000 between the fourth and fifth generation of broadband cellular technology (4G/5G) [6]. This development asks for a radical rethinking of ways wireless systems are designed, ranging from the design of individual units to how the entire system is organized and operating.

## 1.1 The Energy Efficiency Issue

A particularly important issue in the design of future wireless systems is the improvement of *energy efficiency*. There are numerous reasons why specifically energy efficiency is of primary concern. These can roughly be classified as being



**Figure 1.2:** Evolution of the number of wireless machine-to-machine connections [3] (values for 2017. - 2021. are predictions)

either environmental, economic or practical, with numerous interdependencies between the three groups.

- **Environmental:** A study from 2011 [4] predicts that the total CO<sub>2</sub> equivalent emissions related to mobile networks (manufacturing + operation), expressed in megatonnes of CO<sub>2</sub> equivalent (CO<sub>2</sub>e) will have increased from 86 Mt CO<sub>2</sub>e in 2007 to 235 Mt CO<sub>2</sub>e in 2020, which is equivalent to a third of total emissions of United Kingdom in 2011 [4]. A significant portion (30 %) of these emissions is due to radio access network (RAN) operation. Although small in comparison with other sources of greenhouse gas emissions, mobile industry emissions still need to be kept under control, having in mind the underlying exponential growth in the number of devices. Improved energy efficiency of mobile systems, on all levels, is the key to achieving this goal.
- **Economic:** Tightly connected with greenhouse emissions, the global RAN energy consumption is forecast to have grown from 49 TWh in 2007 to 109 TWh in 2020 [4], which is approximately equal to the total electrical energy consumption of Netherlands in 2016 [8]. This energy consumption can be directly related to operational expenditure (OPEX), and network operators will be interested in maintaining, or even reducing these costs with the proliferation of the number of mobile connections. As

with emissions, improved energy efficiency will be instrumental in keeping these costs low.

- **Practical:** A white paper by the Next Generation Mobile Networks (NGMN) alliance predicts that in 5G, battery life will be increased to “at least 3 days for a smartphone, and up to 15 years for a low-cost MTC [machine-type-communication, i.e. M2M] device.” [5]. Extended battery life will offer greater mobility and greater ease of use for mobile users, and simplified maintenance of M2M networks. Along with advances in battery technology, improving the energy efficiency of device hardware related to signal processing is key to achieving a longer battery life for battery-powered devices.

In short, energy efficiency needs to be improved in all segments of a wireless communication system. Improving the energy efficiency of RAN is important in the context of environment or cost. Currently, energy consumption of user equipment, normalized by the number of mobile subscribers, is much smaller than the energy consumption of RAN [4], [7]. Therefore, improving the energy efficiency of devices is of secondary importance in the context of environment or expenditure. However, when it comes to quality of user experience or convenience of maintaining a network of connected devices, energy efficiency is the prime design requirement. Energy efficiency will be the central design principle of the next-generation (5G) mobile systems, where the aforementioned 1000x increase in data traffic should be accompanied by reducing the overall power consumption of wireless systems to half of what it is today [5] or at least by maintaining it at the same level [6]. This corresponds to a staggering 1000x to 2000x improvement of overall energy efficiency!

Achieving this ambitious goal is envisioned to be the cumulative result of applying a plethora of system design techniques, which can be grouped as [9]:

- **Resource allocation:** techniques of allocating resources in a wireless system with the goal of maximizing energy efficiency, in contrast to maximizing throughput, which is the traditional approach;
- **Energy harvesting/transfer:** exploiting renewable energy sources and the electromagnetic energy of radio signals to provide energy for operation of wireless systems;
- **Network planning:** in general terms, these techniques are concerned with the spatial distribution and number of infrastructure nodes that support the operation of a wireless network. In the context of next-generation networks, we can distinguish two subgroups of energy-efficient techniques [6], [9]:

- *Densification of infrastructure nodes*: a (large) number of nodes with a smaller coverage is deployed to assist a base station (BS) with a wider coverage [6]. These “assisting nodes” can be other BSs or relays.
- *Densification of antennas at BS - massive multiple input, multiple output (MIMO)*: the number of antennas at the BS is (typically) set to the order of a hundred, which is one or two orders of magnitude larger than in traditional systems. This enables diverse possibilities of improving the energy and cost efficiency of the system.
- **Hardware**: the design of analog and digital hardware used to implement devices and network nodes is tailored specifically to improve energy efficiency.

## 1.2 Hardware-Consciousness

The topics analyzed and presented in this thesis are concentrated, for the most part, on the hardware-related aspects of energy efficiency (and related metrics such as cost) in the context of wireless systems design. Some of the investigated topics also fall in the category of network design, but with a strong underlying relation to hardware constraints. Hardware constraints lie in the heart of the entire work, and thus the qualifier *hardware-conscious* is used in the descriptions of all the topics considered. The online version of the Oxford English Dictionary [10] offers several definitions of “conscious”, out of which two selected ones are:

1. Aware of and responding to one’s surroundings.
2. Having knowledge or awareness.

In light of these two definitions, the topics covered in this thesis fall into one of the following categories:

- *Design of wireless communication systems that are hardware-conscious*: such systems consist of blocks whose hardware subblocks are tunable. This tuning allows trading performance with power consumption. More specifically, the value of the tuning parameter is chosen such that the power consumption of the block is minimized while the block delivers some predetermined level of performance, under fluctuating environment/system parameters. This is in contrast with designing the system such that the value of the parameter in question is determined according to some worst-case scenario of environment/system parameters. If the

tuning is performed according to the random fluctuations of the environment, it is done in real time. It can also be set according to a specific constellation of system parameters and then it is usually fixed. In any case, reduced power consumption of the block due to tuning improves the overall energy efficiency of the system. The block is “conscious of its hardware”, hence the hardware-conscious qualification. One example of a real-time tunable block is a receiver analog front end (AFE) with tunable linearity that adapts to the fluctuating level of an out-of-band interferer.

- *Hardware-conscious design of wireless communication systems*: the design of the system takes into account (“is conscious of”) a hardware-related constraint that is fundamental and cannot be changed but necessitates a workaround. This usually entails completely redesigning the system. One example of such a workaround is changing the massive MIMO BS topology from star to daisy chain in order to do away with connections from the periphery to the central processing unit. Moving from star to daisy chain topology asks for a complete redesign of signal processing algorithms, with a controlled (ideally minor) degradation of performance.

### 1.3 Thesis Structure

The physical layer of a wireless communication system can be broken down into constituent parts corresponding to different layers of abstraction. A bottom-up list of these layers relevant to the work at hand is

- **Hardware block** (e.g. low noise amplifier - LNA, mixer, channel decoder). One or more hardware blocks constitute a(n)
- **Analog front end/analog back end (AFE/ABE), mixed-signal section, digital baseband section (BB)**. Put together, these form a
- **Transmitter/receiver (transceiver)**. A transceiver together with an antenna constitutes a
- **Single-antenna unit (AU)**. Several AUs, together with optional additional hardware can be organized to form a
- **Base station/access point (BS/AP)**, or a **group of cooperating single-antenna user equipments (UEs) or relays**.

The first part of the thesis contains a detailed overview of selected parts of this stratification, where special focus is put on the impact of hardware on both power consumption/energy efficiency/cost efficiency and performance. This overview is divided into two sections corresponding to different chapters of the thesis:

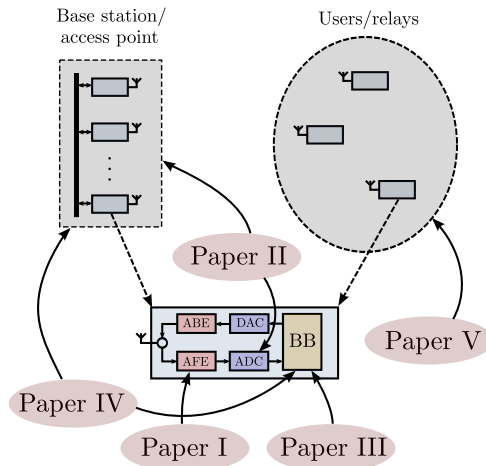
- **Chapter 2** analyzes the tradeoffs between power consumption/complexity and performance on the level of a receiver, with separate sections on the AFE, mixed-signal section (analog-to-digital conversion - ADC) and digital baseband where the low density parity check (LDPC) channel decoder is in focus.
- **Chapter 3** gives a brief overview of massive MIMO (MaMI) systems, with focus on the impact of hardware impairments and limitations on performance and energy/cost efficiency. Additionally, multipair two-way relay systems with a large number of relays are described, and some inherent hardware/cost limitations connected with the implementation of such systems are identified.

Finally, **Chapter 4** gives a detailed summary of the main findings and scientific contributions in each of the research papers.

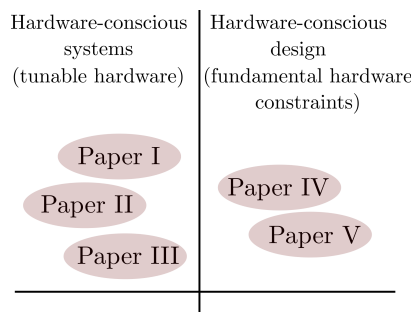
The second part of the thesis contains the five original research papers. Topics covered by each of the papers can be mapped to different layers of the aforementioned stratification, which is illustrated in Fig. 1.3. Additionally, Fig. 1.4 shows how the papers are classified based on whether they thematize hardware-conscious systems or hardware-conscious design. More specifically,

- **Paper I** looks into how power consumption of an AFE scales with linearity and thermal noise, which can be directly mapped to performance, expressed by the signal-to-noise-and-distortion ratio (*SNDR*);
- **Paper II** examines how energy efficiency of a massive MIMO system depends on the bit resolution of the ADCs at the BS;
- **Paper III** presents an LDPC decoder with a flexible tuning parameter that can be used to trade computational complexity with error rate performance;
- **Paper IV** proposes that the BS array in a massive MIMO system is arranged in form of daisy chain, thus obviating the central processing unit (CU) and corresponding links from AUs to CU. These links are known to form a bottleneck in terms of throughput (which could also be related to energy consumption or other cost metrics). Signal processing algorithms tailored to the daisy chain topology are also presented;

- **Paper V** examines a two-way relay system with a large number of relays and how the level of cooperation between relays impacts the performance. The level of cooperation can be directly connected with various cost metrics; the paper itself analyzes bandwidth utilization.



**Figure 1.3:** Illustration of a generalized wireless communication system and its constituent parts with mapping of research papers to each part/layer of abstraction. ABE = analog back end, AFE = analog front end, DAC = digital-to-analog converter, ADC=analog-to-digital converter, BB = digital baseband



**Figure 1.4:** Classification of research papers according to the type of hardware constraints analyzed

## Chapter 2

# Power-Performance Tradeoffs in Receivers

As it was pointed out in Chapter 1, optimizing hardware power consumption is one of the means of improving energy efficiency in wireless systems. A large portion of communication-theoretical works that implicitly or explicitly concern hardware in the context of wireless systems focuses on the *transmitter side*. This can include optimizing transmit power/energy, e.g. finding energy-efficient modulation and coding schemes, or improving the energy efficiency of the transmit power amplifier, e.g. by use of schemes for digital predistortion and peak-to-average-power ratio (PAPR) reduction.

When it comes to analyzing the energy efficiency of the entire communication system, with all transmitters (TX) and receivers (RX) taken into account, focusing solely on transmit power issues makes sense when distances between TX and RX are large. In this scenario, TX power is much larger than the power consumed by the circuits performing analog and digital signal processing in the receiver [11]. On the other hand, when distances in the system are small, *RX power consumption* also needs to be taken into account, as it can be on the order of, or even larger than the TX power. Finally, if the analysis focuses only on the energy efficiency of a receiver (e.g. in the case of battery-constrained RXs), optimizing RX power consumption becomes an issue of primary importance.

A standard approach in works analyzing the energy-efficiency on the level of the entire communication system is to assume that the RX circuitry consumes a constant amount of power regardless of system and environment parameters, as in e.g. [11]. Here the term “system parameters” subsumes system design variables that are under the control of a designer, such as



- Number of users in a multiuser system;
- Number of antennas in the base station/access point;
- Coverage area;
- Transmit power;
- Bandwidth.

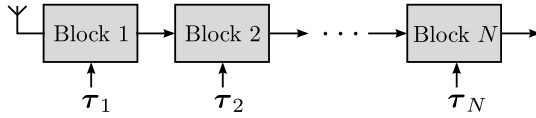
On the other hand, “environment parameters” refers to variables outside of the control of the designer that are usually stochastic in nature, such as

- Received signal power, which experiences fluctuations due to fading;
- In- and out-of-band interference.

The assumption of constant RX power consumption is well grounded in common receiver design practice. Namely, a very common approach in building receivers is to have one design that will deliver the prescribed performance in a wide range of system and environment parameters. Such a design is made assuming the most adverse combination of environment parameters, usually specified by the wireless standard.

However, this conservative approach in receiver design is markedly suboptimal. Namely, for the vast majority of time the environment conditions are better than the worst-case and a fixed, conservatively designed receiver is delivering a performance which exceeds the prescribed minimum level. Since the performance can always be traded for power consumption, this implies that a conservatively designed receiver will waste energy [12–14]. On the other hand, if one or more of the building blocks in the receiver are made tunable, it would be possible to change their properties as the environment changes. The tuning would be performed such that the performance constraint is always met, but the power consumption is not higher than what is necessary. Such a tunable, adaptive receiver would have a smaller energy consumption compared to the conservative worst-case design [12, 13].

The energy-efficient receiver design problem can now be generalized and formalized. Namely, under the condition that the receiver meets the performance requirement, its power consumption should be minimal *at any time instant of its operation, taking all system and environment parameters into account*. If the values of system and/or environment parameters change over time, the receiver necessarily needs to be flexible, i.e. its constituent blocks need to have a dynamic structure which will adapt to the changes. In line with the nomenclature introduced in Chapter 1, such a receiver can also be referred to as being *hardware-conscious*.



**Figure 2.1:** A general receiver structure

In order to give the described concept a mathematical formulation, we assume that the receiver is structured as a chain of  $N$  blocks, shown in Fig. 2.1. A set of design parameters, whose values are collected in a variable-length vector  $\boldsymbol{\tau}_i$  of dimensions  $D_i \times 1$  determines both the power consumption  $P_i$  and the performance/quality measure  $Q_i$  of the  $i$ th block in the chain. The power consumption of the entire receiver chain is simply a sum of individual power consumptions,  $P_{\text{tot}} = \sum_{i=1}^N P_i(\boldsymbol{\tau}_i)$ . On the other hand, the overall performance  $Q$ , measured at the output of the chain, is an involved, usually highly nonlinear function of individual quality measures  $Q_i$  and system variables. All design parameters can be represented by vector  $\boldsymbol{\tau} = [\boldsymbol{\tau}_1^T \boldsymbol{\tau}_2^T \dots \boldsymbol{\tau}_N^T]^T$  of dimensions  $D \times 1$ ,  $D = \sum_{i=1}^N D_i$ . Additionally, system parameters are represented by a  $S \times 1$  vector  $\boldsymbol{\sigma}$  and all environment parameters are collected in a  $E \times 1$  vector  $\boldsymbol{\nu}$ . The energy efficient receiver design can now be formulated as an optimization problem

$$\begin{aligned} & \underset{\boldsymbol{\tau}}{\text{minimize}} && P_{\text{tot}} = \sum_{i=1}^N P_i(\boldsymbol{\tau}_i) \\ & \text{subject to} && Q(\boldsymbol{\tau}, \boldsymbol{\sigma}, \boldsymbol{\nu}) \geq Q_{\min}, \end{aligned} \quad (2.1)$$

where  $Q_{\min}$  is some predetermined minimal performance/quality level.



**Figure 2.2:** A high-level structure of a wireless receiver

In the follow-up of Chapter 2, the specifics of problem (2.1) are analyzed in detail. The high-level structure of the receiver represented in Fig. 2.2 is assumed, where receiver blocks are organized into functional units depending on what type of signals are being processed (only analog, analog and digital, or only digital). For each of the three functional units, specific design parameters  $\boldsymbol{\tau}$  are identified and their connection with performance and power consumption/complexity is examined in detail. This will set the ground for the analyses described in Papers I - III, where each paper thematizes one of the functional

units using the approach summarized by (2.1).

## 2.1 Analog front end

In this work, analog front end (AFE) comprises all RF and baseband receiver blocks that perform *analog signal processing tasks*: amplification, frequency synthesis, downconversion and filtering. One example of an AFE is the direct conversion receiver, an analog signal processor that lends itself to an integrated, on-chip implementation and has thus become a virtual standard for the analog part of contemporary receiver implementations [16, Ch. 4]. However, the descriptions given here are general and valid for any type of analog receiver. In the analysis that follows, we focus on blocks performing amplification, downconversion and filtering.

In general, each block in the AFE chain can be described as having a nonlinear transfer function and adding thermal noise. More specifically, if memoryless nonlinearity is assumed for simplicity, the output signal for the block  $i$  with input signal  $x_{\text{in},i}$  can be written as

$$x_{\text{out},i}(t) = \alpha_{1,i}x_{\text{in},i}(t) + \alpha_{2,i}x_{\text{in},i}^2(t) + \alpha_{3,i}x_{\text{in},i}^3(t) + \dots + w_i(t), \quad (2.2)$$

where  $w_i(t)$  is the thermal noise signal.

Thermal noise and the effects of nonlinearities will corrupt the wanted (information-bearing) signal and thus have a decisive impact on the performance of the  $i$ th block, as well as the entire chain. The quantification of their effects will now be analyzed in detail.

### 2.1.1 Thermal noise

There exist two important measures of the impact of thermal noise in the  $i$ th block. One is the noise power spectral density (PSD)  $\bar{V}_{\text{n},i}^2$  [ $\text{V}^2/\text{Hz}$ ]. The other one is the noise factor, which is defined as

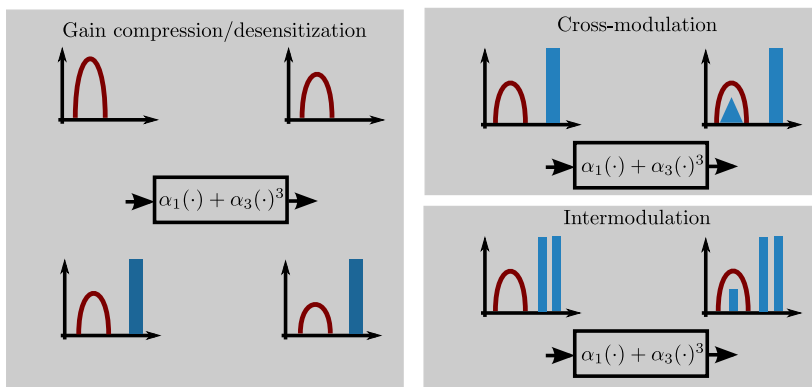
$$F_i = \frac{SNR_{\text{out},i}}{SNR_{\text{in},i}}, \quad (2.3)$$

where  $SNR_{\text{out},i}$  and  $SNR_{\text{in},i}$  are the signal-to-noise ratios at the output and input of the  $i$ th block, respectively.

### 2.1.2 Nonlinearities

In contrast to noise, which simply adds to the wanted signal, the effects of nonlinearities are more intricate in their genesis as well as their analysis. Nonlinearity effects that have the biggest impact on receiver performance are shortly

described here and illustrated in Fig. 2.3. A detailed description can be found in [16, Ch. 2.2].



**Figure 2.3:** A summary of most important nonlinearity effects

- **Gain compression:** decrease of the effective gain of the wanted signal. The underlying cause is third-order nonlinearity. It occurs when either the wanted signal is strong, or a weak wanted signal is accompanied by **one** strong out-of-band interferer. The effect in latter case is also referred to as desensitization or blocking;
- **Cross modulation:** mixing of wanted signal and **one** out-of-band interferer due to third-order nonlinearity. The resulting signal falls in the band of the wanted signal, causing interference;
- **Intermodulation:** mixing of **two** out-of-band interferers due to third-order nonlinearity. If  $f_c$ , wanted  $\approx 2f_c$ , interferer 1  $- f_c$ , interferer 2, the resulting signal falls in the wanted signal band.

The described effects occur when wanted signal is in the passband as well as in baseband.

It is immediately observed that the third-order nonlinearity is the main “culprit” in all of the most important nonlinearity effects. It is quantified by the input-referred third-order intercept point  $V_{\text{IIP}_3}$ , which is defined as the voltage level of the two interference signals in the “intermodulation” scenario for which the levels of wanted signal and intermodulation distortion are the same.

### 2.1.3 The case of out-of-band interferers

As seen in the preceding section, the interference signals that do not share the bandwidth with the wanted signal can still inflict damage and cause in-band interference in conjunction with nonlinearities. Apart from nonlinearity, there are numerous other scenarios in which out-of-band (OOB) interference can be detrimental to receiver performance, and some will be described in the later sections.

The effects of OOB interference are usually overlooked or ignored by communication-theoretic works. However, its impact is well-known to hardware designers and worst-case interference scenarios are carefully outlined in communication standards. For example, in the Long-Term Evolution (LTE) standard, depending on the scenario, systems need to be designed to deliver minimum performance in presence of OOB interference which is anywhere from 20 dB to 60 dB stronger than the wanted signal [17].

OOB interference originates in emissions of nearby wireless devices and other units that emit electromagnetic radiation. It can also occur on the same device when TX and RX are working in a frequency duplex setup and the TX signal leaks to the RX due to insufficient TX-RX isolation. The issues connected to OOB interference are located exclusively in the analog part of the receiver, for two main reasons. First, some of the analog circuits (as amplifiers or mixers) are broadband, so both wanted signal and OOB signals pass through. Secondly, even if analog circuits are specifically made narrowband (e.g. baseband or passband filters), they still have limited capabilities of combating very strong interference.

### 2.1.4 Chain rules for noise and nonlinearity

A common task in receiver design is calculating the overall noise and nonlinearity metrics for a chain of  $N$  analog blocks when the individual input-referred third-order intercept (IIP3) voltage  $V_{\text{IIP3}}$ , noise PSD  $\bar{V}_n^2$  or alternatively, noise factor  $F$  are known for each block.

The expressions given here are based on the ones found in [18]. Instead of more recognizable expressions based on power gains of individual blocks, these are based on voltage gains

$$A_{v,i} = \frac{V_{\text{out}}}{V_{\text{in}}}. \quad (2.4)$$

The use of voltage gains is more applicable in integrated RF designs, where the input impedance of one block might not be matched to the output impedance of the preceding block, thus making power and voltage gains different when measured in dB and causing possible ambiguity [16, Ch. 2]. To be more

specific, the chain rules from [18] are based on the *loaded* voltage gains

$$A_{vn,i} = A_{v,i} \left( \frac{R_{load,i}}{R_{load,i} + R_{out,i}} \right), \quad (2.5)$$

where  $R_{load,i}$  and  $R_{out,i}$  are the load and output impedances of the  $i$ th block, respectively.

The total noise figure  $F_{tot}$  of the chain can be found as

$$F_{tot} = 1 + \frac{1}{kT50} \left( \bar{V}_{n,1}^2 + \frac{\bar{V}_{n,2}^2}{A_{vn,1}^2} + \frac{\bar{V}_{n,3}^2}{A_{vn,1}^2 A_{vn,2}^2} + \dots \right), \quad (2.6)$$

and the chain rule for IIP3 is given as

$$V_{IIP3, tot}^2 = \left( \frac{1}{V_{IIP3,1}^2} + \frac{A_{vn,1}^2}{V_{IIP3,2}^2} + \frac{A_{vn,1}^2 A_{vn,2}^2}{V_{IIP3,3}^2} + \dots \right)^{-1}, \quad (2.7)$$

where  $k$  is the Boltzmann constant and  $T$  temperature in Kelvins and both input and output impedances of the chain are set to 50 Ohms. One can gather from (2.6) and (2.7) that the noise contribution of the front stages of the chain determines the overall noise figure, and that the total IIP3 is mostly determined by the IIP3 of the blocks at the end of the chain.

### 2.1.5 Power - performance tradeoffs in the analog front end

Based on the foregoing discussion, it is evident that noise and third order nonlinearity for the most part determine the performance of the analog part of the receiver. How do they relate to the power consumption? Looking for the answer to this questions starts by defining the dynamic range of an analog block:

$$DR \triangleq \frac{V_{IIP3}^2}{V_n^2}. \quad (2.8)$$

Starting from the single MOSFET transistor and moving on to more complicated structures such as a common-source LNA, a Gilbert-cell mixer and OTA-C low-pass filter, it can be shown [15], [18] that for all the aforementioned basic analog blocks,  $DR \propto I$ , where  $I$  is the bias current of the circuit. Since the power consumption is simply  $P = V_{DD}I$  with  $V_{DD}$  being the supply voltage, this yields a very simple but powerful relation between the power consumption and performance of an analog block:

$$P = \kappa DR, \quad (2.9)$$

with  $\kappa$  being a circuit-dependent proportionality parameter.

What about the power-performance tradeoff on the level of the entire analog chain? Based on (2.9), it can be shown [18] that if the total noise factor  $F_{\text{tot}}$  and total IIP3 voltage  $V_{\text{IIP3, tot}}$  are given, values of  $\bar{V}_{\text{n},i}^2$  and  $V_{\text{IIP3},i}$  that minimize the *total power consumption of the chain*  $P_{\text{tot}}$  can be found for each block  $i$ . Most importantly, the connection between this minimal power consumption and  $F_{\text{tot}}$  and  $V_{\text{IIP3, tot}}$  is given as

$$P_{\text{AFE, tot}}^* = \frac{V_{\text{IIP3, tot}}^2}{(F_{\text{tot}} - 1)kT50} \left( \sum_{i=1}^N \sqrt[3]{\kappa_i} \right)^3. \quad (2.10)$$

In the context of the general receiver design problem (2.1),  $V_{\text{IIP3}}$  and  $F$ , alternatively  $\bar{V}_{\text{n}}^2$  can be identified as the “tuning knobs”  $\boldsymbol{\tau}$ , either on the level of a single block or the entire analog section. Expressions (2.9) and (2.10) show how these parameters can be related to the power consumption of one block or the entire AFE. The dynamic range  $DR$  can serve as an elementary performance metric  $Q$ . Mapping to other metrics, such as the signal-to-noise-and-distortion ratio  $SNDR$  can be performed in a straightforward fashion.

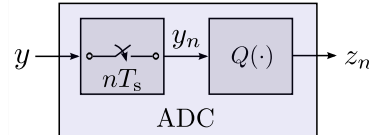
### 2.1.6 Tunable analog designs

Design of analog front ends that are able to dynamically change their structure with the goal of saving power has been a topic of interest in both academia and industry. A short overview of designs is presented here. In [19], total noise figure of the receiver is tuned by enabling gain tunability in the LNA. Paper [20] describes an LNA design where  $F$  and IIP3 are made orthogonally tunable, for maximal flexibility between scenarios featuring varying wanted signal power and OOB interference levels. The design described in [21] features mixers with adjustable IIP3 and  $F$ . Works [22–24] focus on flexible channel select filters that are able to tune their bandwidth to accommodate a varying symbol rate, or to adapt their dynamic range to the fluctuations in the OOB interference level or wanted signal power. Paper [25] features an entire analog front end with noise, linearity and selectivity adaptable to a varying OOB interference level. Finally, industrial patent [26] describes a receiver which switches between a high- and low-linearity implementations of a front end depending on the blocker level.

## 2.2 Analog-digital conversion

The section of the receiver performing the translation of analog signals into digital ones consists of a single block: analog-to-digital converter (ADC). The

process of converting an analog signal into a digital one can be split into two operations: sampling with frequency  $f_s$  (discretization in time) and quantization (discretization in amplitude), as illustrated in Fig. 2.4, where  $T_s = \frac{1}{f_s}$  and  $Q(\cdot)$  is used to denote the quantization operator. Based on the sampling frequency  $f_s$ , ADCs can be classified in two types. *Nyquist-rate* ADCs, such as flash, pipeline and successive approximation register (SAR) ADCs have  $f_s$  equal to the Nyquist sampling rate. On the other hand, *oversampling* ADCs such as the sigma-delta ( $\Sigma\Delta$ ) ADC have  $f_s$  which is much larger than the Nyquist rate.



**Figure 2.4:** A conceptual diagram of analog-to-digital conversion.

The fundamental design parameters of an ADC are

- Sampling frequency  $f_s$ ,
- Power consumption  $P_{\text{ADC}}$ , and
- Nominal bit resolution  $b$ .

Different ADC types, some of which have been listed above, offer different tradeoffs between these parameters, and some architectures are preferred over the others depending on the application.

The relations between the fundamental parameters of an ADC, as well as between the fundamental parameters and performance are commonly described using simplified heuristic expressions. These, however, can be shown to match well with theoretic analysis. All of these aspects will be covered in the following.

### 2.2.1 ADC performance. Modeling the effects of quantization

The corruption of the input signal by the ADC hardware block can be tracked down to different causes. Analog circuitry in the ADC will add thermal noise to the signal and nonlinearities in the circuitry can cause nonlinear distortion. Since the ADC is usually preceded by a chain of components with high composite gain, the effects of thermal noise can in general be neglected (cf. (2.6)). On the other hand, the effects of nonlinear distortion in an ADC can have a



significant impact on the performance (cf. (2.7)), especially in presence of a strong, poorly filtered OOB interferer. In any case, modeling the impact of thermal noise and nonlinearities in the ADC can be considered covered by the analysis in the preceding section.

Here, the focus is put on the inherent corruption of the input signal incurred by the discretization in amplitude, i.e. quantization. The quantization operation  $Q(\cdot)$  is heavily nonlinear in nature. However, for the purpose of embedding it in an analysis of a wider system setup it is beneficial to represent it by a classical “linear amplification plus noise” model and indeed, good models can be found that serve this purpose.

### The PQN model

The first commonly used model is in the simple form of

$$z_n \triangleq Q(y_n) = y_n + q_n, \quad (2.11)$$

where  $q_n$  is in general correlated with  $y_n$ . For example, if the quantizer  $Q(\cdot)$  is designed to minimize the mean square error (MSE)  $\mathbb{E}\{|y_n - z_n|^2\}$ ,  $q_n$  can be shown to *always* be correlated with  $y_n$  [27, Ch. 6]. Correlatedness of  $q_n$  with  $y_n$  might sound as bad news, taking into account that the calculation of most system performance metrics assumes additive noise that is uncorrelated with the signal. However, under some realistic assumptions that are readily encountered in practice, the following approximations regarding the model (2.11) can be adopted [27]:

1.  $y_n$  and  $q_n$  are uncorrelated, i.e.  $\mathbb{E}\{y_n q_n\} = 0$  for  $\mathbb{E}\{y_n\} = \mathbb{E}\{q_n\} = 0$ ;
2. Additive noise process  $q_n$  is uniformly distributed;
3.  $q_n$  is white.

Model (2.11) in conjunction with the above set of approximations is commonly referred to as the **pseudoquantization noise (PQN)** model. In order to analyze the conditions under which the PQN model is applicable it is assumed that the quantizer  $Q(\cdot)$  is uniform and operating on a zero-mean Gaussian input signal  $y_n$  with variance  $\sigma_y^2$ . A predefined symmetric input signal range of width  $2Y_{\max}$  is divided into  $N = 2^b$  quantization bins of width

$$\Delta = \frac{2Y_{\max}}{2^b}. \quad (2.12)$$

For the described setup, it is shown in [28] that

1.

$$\mathbb{E}\{y_n q_n\} = -2\sigma_y^2 \sum_{l=1}^{\infty} (-1)^l \exp\left(-\frac{2\pi^2 l^2 \sigma_y^2}{\Delta^2}\right); \quad (2.13)$$

2.  $q_n$  is approximately uniform for  $\frac{\sigma_y}{\Delta} > 0.5$ ;3.  $q_n$  is uncorrelated,  $\mathbb{E}\{q_n q_{n+l}\} = 0$  for an uncorrelated input process  $y_n$ .

Moreover, the variance of quantization noise  $q_n$  is shown to be equal to

$$\begin{aligned} \sigma_q^2 \triangleq \mathbb{E}\{q_n^2\} &= \frac{\Delta^2}{12} \left[ 1 + \frac{12}{\pi^2} \sum_{l=1}^{\infty} \frac{(-1)^l}{l^2} \exp\left(-\frac{2\pi^2 l^2 \sigma_y^2}{\Delta^2}\right) \right] \\ &\approx \frac{\Delta^2}{12} \quad \text{for } \frac{\sigma_y}{\Delta} > 0.5. \end{aligned} \quad (2.14)$$

It is evident that the PQN model applies well for high bit resolutions  $b$ . Additionally, even at lower values of  $b$  the PQN model can be shown to be applicable with carefully chosen values of the ADC input backoff  $\mu = Y_{\max}^2 / \sigma_y^2$ . For convenience, the variance of the quantization noise from (2.14) can be expressed as a function of bit resolution  $b$ :

$$\mathbb{E}\{q_n^2\} \approx \frac{1}{3} \mu \sigma_y^2 2^{-2b}. \quad (2.15)$$

For the purpose of further discussion it is useful to define the signal-to-quantization noise ratio  $SQR$  at the ADC output:

$$SQR \triangleq \frac{\sigma_y^2}{\sigma_q^2}, \quad (2.16)$$

and from (2.15) this metric is found to be

$$SQR \approx \frac{3}{\mu} 2^{2b}. \quad (2.17)$$

### The AQN model

In another model that is commonly used to describe the effects of quantization, the output of the quantizer is split into a wanted signal part and additive noise part uncorrelated with the wanted signal, i.e.

$$z_n = (1 - \beta)y_n + \epsilon_n, \quad (2.18)$$

where  $\mathbb{E}\{y_n\epsilon_n\} = 0$  for  $\mathbb{E}\{y_n\} = \mathbb{E}\{\epsilon_n\} = 0$ . Now, if the quantizer  $Q(y_n)$  is designed such that

$$z_n = \mathbb{E}\{y_n|z_n\}, \quad (2.19)$$

factor  $\beta$  is calculated as [29]

$$\beta = \frac{\mathbb{E}\{(y_n - z_n)^2\}}{\mathbb{E}\{y_n^2\}}, \quad (2.20)$$

and moreover, the variance of the additive noise term  $\epsilon$  is

$$\sigma_\epsilon^2 = \beta(1 - \beta)\sigma_y^2. \quad (2.21)$$

Condition (2.19), also known as the ‘‘centroid condition’’, is satisfied for  $Q(\cdot)$  designed to minimize the MSE [27] and moreover, it will be also satisfied for an uniform quantizer at high bit resolutions. Model (2.18) is referred to as the **additive quantization noise (AQN) model**. It should be noted that if the AQN model is derived under the assumption (2.19), no assumptions are needed on the distribution of the input  $y_n$ . Another common way of deriving model (2.18) is through the use of Bussgang theorem, as in e.g. [30], but in that case  $y_n$  must be assumed to be Gaussian, which however commonly holds in practice. From (2.20) and (2.11), it is seen that PQN and AQN models are connected by

$$\beta = \frac{\sigma_q^2}{\sigma_y^2}. \quad (2.22)$$

Values of  $\beta$  need to be found numerically, as is done in the seminal paper by Max [31] for  $\sigma_y^2 = 1$ . At high bit resolutions  $b$ , closed-form approximations and bounds for  $\beta$  are available under some assumptions commonly encountered in practice. For example, if input  $y_n$  is Gaussian and the quantizer is designed to minimize the MSE,  $\beta$  can be well approximated as [27]

$$\beta \approx \frac{\pi\sqrt{3}}{2}2^{-2b}, \quad (2.23)$$

and from (2.18), (2.21) and (2.23), the *SQR* in the high-resolution scenario can be found as (cf. (2.17))

$$SQR \triangleq \frac{(1 - \beta)^2\sigma_y^2}{\sigma_\epsilon^2} = \frac{1 - \beta}{\beta} \approx \frac{1}{\beta} = \frac{2}{\pi\sqrt{3}}2^{2b}. \quad (2.24)$$

On the other hand, for a uniform quantizer  $Q(\cdot)$  operating on a Gaussian input  $y_n$ , it is shown in [32] that

$$O(\beta) \geq O(b2^{-2b}). \quad (2.25)$$

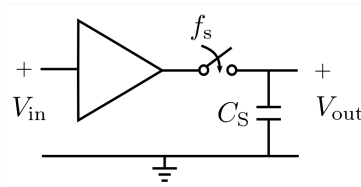
Overall, it can be concluded that at high bit resolutions  $b$ , with uniform and MMSE quantization and a Gaussian input signal, model (2.11) applies well and quantization noise and input signal can safely be assumed uncorrelated. Factor  $\mu$  from (2.15) can be optimized numerically, and is shown in Paper II to be of  $O(b)$ . Therefore, it can further be claimed that at high bit resolutions, quantization noise variance  $\sigma_q^2 = \gamma\sigma_y^2$ , where  $O(\gamma) \geq O(b2^{-2b})$  for both uniform and MMSE quantization and Gaussian input. At low resolutions, the deviations from this general model increase but it can still be considered robust enough to enable a reliable analysis.

## 2.2.2 ADC power consumption

The power  $P_{\text{ADC}}$  consumed in the process of analog-to-digital conversion is, for the most part, determined by the circuitry performing the sampling operation, cf. Fig. 2.4. For the purpose of describing important concepts, this circuitry can in the ideal case be represented by the sample-and-hold (S&H) circuit [33], shown in Fig. 2.5. The performance of this structure is limited by the thermal noise generated in the sampling capacitor  $C_s$ . For a class-B amplifier and a sinusoidal rail-to-rail input to the S&H circuit, the relation between the power consumed in the amplifier,  $P_{\text{amp}}$  and the  $SNR$  at the output of the S&H circuit is found to be ([34–36])

$$P_{\text{amp}} = 8kTf_s SNR. \quad (2.26)$$

The power consumption of the ideal S&H circuit can be considered to be the fundamental lower bound on  $P_{\text{ADC}}$ , and the power consumption of contemporary ADC designs is observed to be one or two orders of magnitude larger than (2.26) [34, 37].



**Figure 2.5:** Ideal sample-and-hold circuit.

The performance of the quantizer is determined by the  $SQR$ , whereas  $P_{\text{ADC}}$  (or bounds thereof) are described in terms of  $SNR$ . This is not strange, since quantization noise is intrinsic to digital signal processing and thermal noise to its analog counterpart. However, it is of interest to establish the relation

between performance of the ADC and its power consumption. More specifically, connecting  $P_{\text{ADC}}$  with the nominal bit resolution  $b$  is of particular interest.

As a preliminary, it is noted that the power dissipated in the sampling capacitor is  $\propto C_s$  and that the power of the thermal noise generated by the same capacitor is  $\propto 1/C_s$  [35]. Now, in order to gain an insight in connecting  $P_{\text{ADC}}$  with  $b$ , a high bit resolution can be assumed so  $SQR$  is also high and the ADC performance can be considered limited by thermal noise. In this regime,  $C_s$  is made large so as to limit the impact of the thermal noise. However, increasing  $C_s$  also has the effect of increasing power consumption. For simplicity, it can be assumed that  $SQR = SNR$  at the ADC output. Then, assuming that the sampler is ideal and dominates the power consumption and that additionally  $Q(\cdot)$  operates on an ideal sinusoidal input (to match the assumptions behind (2.26)), Sundström et. al. show in [38] that the power consumption of this “ideal” ADC is

$$P_{\text{ADC, ideal}} = 24kT2^{2b}f_s, \quad (2.27)$$

Hence, it can be established that the power consumption of an ideal ADC scales *exponentially* with bit resolution (more specifically, as  $2^{2b}$ ) and *linearly* with sampling rate  $f_s$  at high bit resolutions, where ADC performance is *limited by thermal noise*. The scaling laws between the fundamental ADC design parameters observed in (2.27) are shown to be remarkably accurate, especially for ADCs with a high nominal bit resolution.

At lower bit resolutions, quantization noise power increases, so thermal noise can also be relaxed in conjunction with this increase. This implies shrinking of the sampling capacitor and has a welcome consequence of decreasing  $P_{\text{ADC}}$ . However, value of  $C_s$  cannot be decreased below some minimum value  $C_{\text{min}}$ , which is determined by fundamental size limitations of the CMOS process if the ADC has an integrated circuit implementation. It can therefore be expected that the scaling of  $P_{\text{ADC}}$  with  $b$  has a rate lower than  $2^{2b}$  at lower bit resolutions, and ADCs designed to operate in the low bit resolution region are therefore sometimes referred to as being *process-limited* or *technology-limited*.

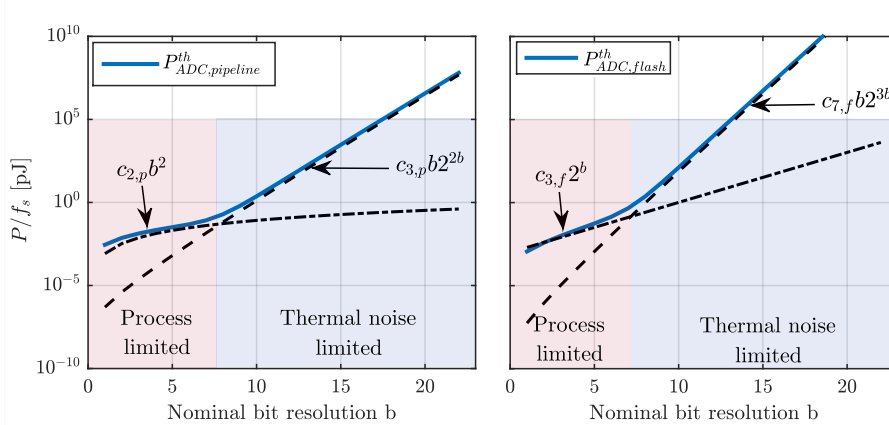
Using (2.27) as the starting point, Sundström et. al. in [38] derive lower bounds on  $P_{\text{ADC}}$  by taking into account all the circuitry needed for implementing actual ADCs and also considering the impact of process limitations. For such a detailed analysis to be meaningful, assumptions on the actual ADC structure need to be made, and in [38] the analysis focuses on flash and pipeline ADCs. The theoretical lower bounds on flash and pipeline ADC power con-

sumption from [38] are given as

$$P_{\text{ADC, flash}}^{\text{th}} = (c_{1,f} + c_{2,f}b + c_{3,f}2^b + c_{4,f}2^{2b} + c_{5,f}b2^{2b} + c_{6,f}2^{3b} + c_{7,f}b2^{3b}) f_s, \quad (2.28)$$

$$P_{\text{ADC, pipeline}}^{\text{th}} = (c_{1,p}b + c_{2,p}b^2 + c_{3,p}2^{2b} + c_{4,p}b2^{2b}) f_s, \quad (2.29)$$

where constants  $c_{i,f}$ ,  $c_{i,p}$  depend on various circuit parameters related to a particular CMOS technology generation.



**Figure 2.6:** Theoretical lower bounds on the power consumption of the pipeline (2.29) and flash ADCs (2.28), normalized by the sampling rate  $f_s$ .

A deeper look into the rather involved expressions for the bounds (2.29) and (2.28) reveals that, depending on  $b$ , different terms can be considered to dominate, which simplifies the expressions and enables drawing some interesting insights. The bounds are plotted in Fig. 2.6 together with the respective dominant terms, for the values of circuit parameters given in detail in [38] that correspond to the 90 nm CMOS process. The distinction between the process- and thermal noise limited regions of operation is clearly visible, with a distinct “knee” at the intersection of the two dominant terms. For pipeline ADCs, the lower bound on power scales roughly quadratically in the process-limited region and as  $b2^{2b}$  in the thermal noise limited region. On the other hand,  $P_{\text{ADC}}$  of flash ADCs scales approximately as  $2^b$  in the process-limited regime. A comparison of the calculated bounds with actual designs up to the year of publication (2009) was also made in [38] and it showed that the energy consumption of practical designs was 10 - 100 times larger than the bounds.

An alternative, semi-heuristic way of modeling the dependence of  $P_{\text{ADC}}$  on the bit resolution has emerged and gained popularity in the analysis of trends in ADC design. The attractiveness of this approach lies in its simplistic formulation and in the fact that it is agnostic to the choice of the ADC architecture. Namely, it consists of forming a figure-of-merit (FoM) that features the fundamental ADC parameters:  $P_{\text{ADC}}$ ,  $f_s$  and, instead of nominal bit resolution  $b$ , the signal-to-noise-and-distortion ratio  $SNDR$  at the ADC output, or alternatively, the effective number of bits

$$ENOB \triangleq \frac{SNDR_{\text{dB}} - 1.76}{6.02}. \quad (2.30)$$

The  $SNDR$  metric subsumes all measurable impairments occurring during the process of analog to digital conversion, i.e. quantization noise, thermal noise and nonlinear distortion. There are two FoM commonly used in literature. One is the so-called Walden FoM [39]:

$$FoM_{\text{W}} = \frac{P_{\text{ADC}}}{f_s 2^{ENOB}}, \quad (2.31)$$

and the other is the Schreier FoM [37, 40]:

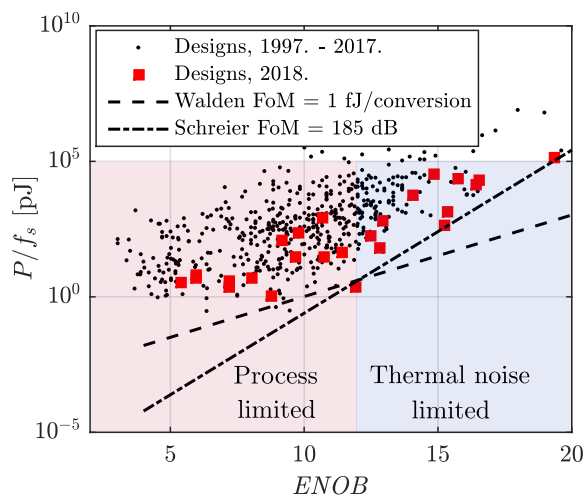
$$FoM_{\text{S}}(\text{dB}) = SNDR_{\text{dB}} + 10 \log_{10} \left( \frac{f_s/2}{P_{\text{ADC}}} \right). \quad (2.32)$$

The basic idea with FoM is to have a simple scalar metric that enables a fair comparison of different ADC designs, regardless of the intricacies of their particular designs. If assumed that FoM stays constant as a certain parameter changes, it can also serve as a “zeroth order” approximation of scaling trends between the fundamental parameters. Using this simplistic approach and with Schreier FoM converted to linear units, two scaling laws emerge:

$$P_{\text{ADC}} = FoM_{\text{W}} 2^{ENOB} f_s, \quad (2.33)$$

$$P_{\text{ADC}} = \frac{3}{4 FoM_{\text{S}}} 2^{2ENOB} f_s. \quad (2.34)$$

Energy consumption data for academic ADC designs, collected in [41] is shown in Fig. 2.7 together with scaling laws (2.33) and (2.34), for values of  $FoM_{\text{W}}$  and  $FoM_{\text{S}}$  chosen as the best fit for the state-of-the-art designs in 2018. In spite of being only a very crude approximation, scaling law (2.33) turns out to be a good fit for process-limited designs in the low- and mid-resolution range. Likewise, (2.34) fits well with high-resolution designs limited by thermal noise.



**Figure 2.7:** Measured energy consumption of ADC designs from ISSCC and VLSI Symposia, 1997. - 2018., together with scaling laws (2.33) (“Walden FoM”) and (2.34) (“Schreier FoM”).

### 2.2.3 Power - performance tradeoff in the ADC

Based on the foregoing discussion, it is evident that a unique exact rule on interdependence between performance and power consumption of the ADC cannot be derived. There is no “silver bullet” here, and in fact the same can be claimed for the analysis of this tradeoff anywhere in the receiver, due to the inherent intricacy of hardware designs. Different rules exist depending on the type of the ADC architecture, operating regime, metrics and parameters involved, and even then they might be derived heuristically. However, some approximate rules of thumb can be stated, and they work rather well. In light of (2.1), nominal bit resolution  $b$  can be chosen as the tuning/tradeoff parameter  $\tau$ .

For simplicity, an affine mapping

$$ENOB = b - \zeta \quad (2.35)$$

can be assumed, where  $\zeta \geq 0$  is a correction factor accounting for the influence of impairments other than quantization noise. For example,  $\zeta = 0.5$  when  $SNR = SQR$  and neglecting other sources of impairment [38]. With this assumption in place and taking into account the FoM-based modeling of power



consumption, it can be claimed that  $P_{\text{ADC}}$  scales as  $2^b$  for low and intermediate bit resolutions and as  $2^{2b}$  for high bit resolutions. The theoretical analysis represented by (2.29) and (2.28), though not being comprehensive and encompassing all possible ADC designs, adheres approximately to this conclusion in some of the selected cases. The analysis of performance presented in Sec. 2.2.1 likewise allows us to conclude that  $SQR$ , an idealized performance measure of the ADC, scales approximately as  $2^{2b}$ , with the approximation being tight at high bit resolutions.

## 2.3 Digital baseband

From the discussion in the preceding sections, an important meta-observation can be made, namely, that the modeling of the tradeoff between power/energy consumption and performance becomes more difficult as one progresses from the antenna to the digital baseband part. Due to the generally overwhelming complexity of digital hardware designs, any closed-form analysis of performance on block level would prove to be intractable, and the same holds for analyzing the power or energy. Therefore, these analyses, which are nevertheless important and needed, have to be performed heuristically or be very general in nature.

In digital processing blocks, computational complexity (here broadly defined as the number of computations needed to perform a certain calculation) is interchangeable with energy consumption and will be used as such in the remainder of the text. Taking a broader definition of performance encompassing bit error rate, throughput and latency, three possible groups of techniques that allow for trading off energy with performance in digital blocks can be identified. Each of these approaches corresponds to a particular layer of design abstraction which are, bottom to top, [42, 43]

- circuit,
- architecture and
- algorithm.

There are many interdependencies between the three groups of techniques and a truly optimal design is attained by cooptimization across layers. Common circuit techniques for trading off power/energy with performance are clock and power gating, dynamic voltage scaling and body biasing. On the other hand, architectural techniques specifically trade throughput and latency with energy, and classical techniques are pipelining and parallelization. Finally, computational complexity can be directly exchanged for error rate performance using a

plethora of algorithmic techniques, heavily dependent on the application and function of the particular block. One standard and general method, applicable in most scenarios is varying the wordlength in fixed-point implementations.

In this thesis, the focus is on application-specific algorithmic techniques for trading off energy/complexity with error rate performance. In particular, a low-density parity check (LDPC) channel decoder with tunable parameters in the decoding algorithm is analyzed. The mechanism of this tuning is described in more detail in the following.

### 2.3.1 LDPC decoding with tunable decoding parameters

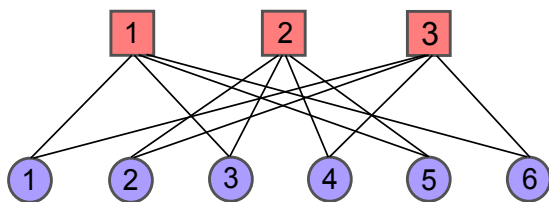
Owing to their excellent forward error correction capabilities, LDPC codes, introduced around 1960 by Gallager [44] have been adopted in a variety of communication standards, most recently in the New Radio (NR) standard for the cellular part of 5G [45,46]. For some LDPC code constructions, performance is fractions of a dB away from Shannon’s capacity.

LDPC codes are able to approach the Shannon limit due to large block sizes and a randomized structure, which can be considered to represent a practical embodiment of Shannon’s random coding argument. An LDPC code of block-length  $B$  and rate  $r = D/B$  is fully described by the  $P \times B$  binary parity check matrix  $\mathbf{H}$ ,  $P = B - D$  which describes parity relations between check bits and information bits in a codeword. For good codes, this matrix is extremely sparse, which is what gives the code its randomized structure. The relations between codeword bits and parity checks can be mapped from the parity check matrix to a bipartite graph, also commonly referred to as the Tanner graph [47], where codeword bits are represented by variable nodes and parity checks with check nodes. Fig. 2.8 shows an illustration of a Tanner graph for a  $B = 6$ ,  $P = 3$  example block code with parity check matrix

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$

Another key defining feature of LDPC codes which makes them practically implementable is an efficient decoding algorithm. The algorithm consists of passing probabilistic messages, or “beliefs” between variable nodes and check nodes in the Tanner graph, in an iterative manner. For computational convenience, the probabilities can be substituted by log-likelihood ratios (LLRs), and the decoding scheme then consists of following steps [48]:

1. *Initialization.* Calculate *a posteriori* LLRs  $L_v$  for each variable node  $v$  from the output of the channel. Initialize messages from variable nodes



**Figure 2.8:** Example bipartite graph.

$v$  to check nodes  $c$  as  $Q_{vc} = L_v$ .

2. *Check-node update.* For all check nodes  $c$  and all variable nodes  $v$  adjacent to  $c$  and contained in the set  $\mathcal{V}(c)$ , calculate the check-to-variable messages

$$R_{cv} = \left[ \prod_{v' \in \mathcal{V}(c) \setminus v} \text{sign}(Q_{v'c}) \right] 2 \tanh^{-1} \left[ \prod_{v' \in \mathcal{V}(c) \setminus v} \tanh \left( \frac{|Q_{v'c}|}{2} \right) \right]. \quad (2.36)$$

3. *Variable node update.* For all variable nodes  $v$  and all check nodes  $c$  adjacent to  $v$  and contained in the set  $\mathcal{C}(v)$ , calculate the variable-to-check messages

$$Q_{vc} = L_v + \sum_{c' \in \mathcal{C}(v) \setminus c} R_{c'v}. \quad (2.37)$$

Update the LLR of bit  $v$ :

$$Q_v = L_v + \sum_{c \in \mathcal{C}(v)} R_{cv}. \quad (2.38)$$

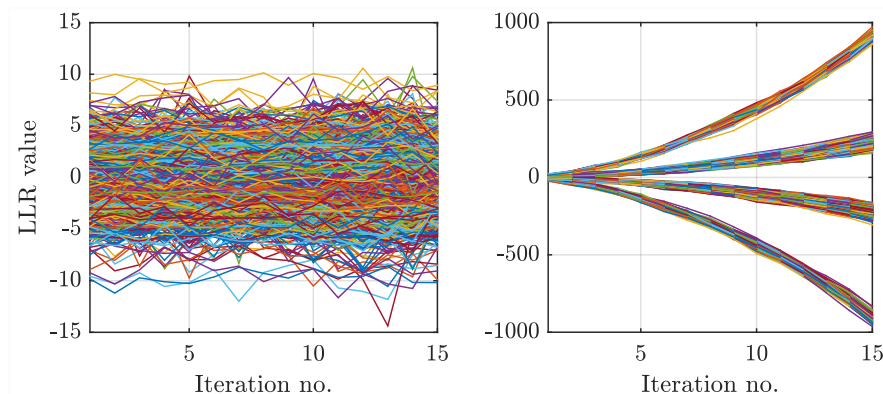
Go back to step 2 until the predefined maximum number of iterations  $I_{\max}$ .

This incarnation of the LDPC belief propagation decoding algorithm is referred to as the sum-product algorithm (SPA).

Step 2 of SPA involves computationally heavy nonlinear functions. In order to reduce complexity, the following useful approximation can be utilized:

$$R_{cv} \approx \left[ \prod_{v' \in \mathcal{V}(c) \setminus v} \text{sign}(Q_{v'c}) \right] \min_{v' \in \mathcal{V}(c) \setminus v} |Q_{v'c}|, \quad (2.39)$$

and if (2.36) is substituted by (2.39), the resulting algorithm is referred to as the min-sum algorithm (MSA). There is a minor performance degradation between SPA and MSA that can be ameliorated by scaling or offsetting the right hand side of (2.39) by a constant value, yielding the normalized/scaled or offset MSA, respectively.

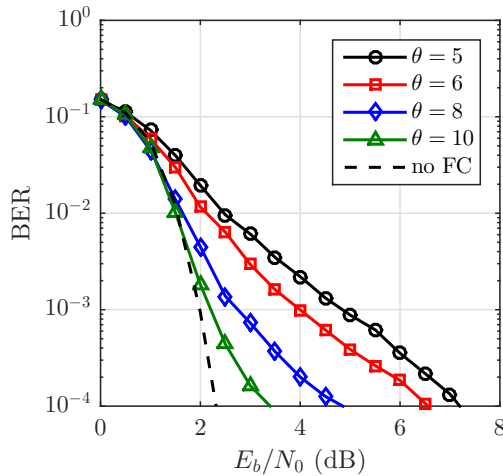


**Figure 2.9:** Evolution of LLR values  $Q_v$  for sample LDPC codewords of a  $R = 1/2$ ,  $B = 648$  LDPC code developed for the PHY layer of the IEEE 802.11n standard. Left:  $E_b/N_0 = 0$  dB, right:  $E_b/N_0 = 3$  dB.

Behavior of LLR values  $Q_v$  over iterations is illustrated in Fig. 2.9 for sample codewords transmitted over an AWGN channel with two different values of  $E_b/N_0$  and decoded using the offset MSA. At  $E_b/N_0 = 0$  dB, the a posteriori LLRs are of poor quality so soft parity checks (2.36) fail and  $Q_v$  do not increase in magnitude, i.e. there is no reinforcement of belief about whether a particular bit is 0 or a 1, altogether surely indicating a decoding failure. On the other hand, LLRs at  $E_b/N_0 = 3$  dB for the chosen sample are good enough to trigger convergence of beliefs, manifested as  $|Q_v|$  growing monotonically with iterations. While this only means that the decoder converges to a *valid codeword*, it does not necessarily mean that it is *the correct codeword*; however, probability of converging to a wrong codeword is marginally low for well-structured codes with large blocklengths  $B$ . From this basic example, one can intuitively gather that error rate performance will improve with each iteration. This indeed proves to be the case, but it is also observed that after a certain point, differential improvements in performance diminish rapidly with iterations [49], indicating that the decoding can be stopped at a certain finite  $I_{\max}$  once a certain performance requirement close to the optimal performance is met. Since

the total number of computations in the SPA/MSA is linear with  $I_{\max}$  as can be gathered from the algorithm description, this early termination can be used to save energy, and  $I_{\max}$  is therefore one simple example of a tunable parameter trading energy with performance in LDPC decoders.

A careful look into the structure of the algorithm combined with the intuition about LLR magnitude evolution for “good” codewords reveals more opportunities for energy-performance tradeoffs. Namely, if during the iterative message passing some  $|Q_v|$  is found to be larger than some predefined threshold  $\theta$ , that LLR/belief can be considered “good enough” and there is no need for its further updating, which means that steps 2 and 3 of the algorithm can be omitted for that v-node. If  $\theta$  is chosen too low, there is a risk that a node is “frozen” before its sign converges to its correct value. The error will propagate in the graph and cause a decoding failure. Threshold  $\theta$  is therefore another decoding parameter that can be used to trade complexity with performance. The described approach, referred to as “forced convergence” (FC) [50–52] uses a bit-level stopping criterion, in contrast with early termination which is a block-level stopping criterion.



**Figure 2.10:** BER performance of offset-MSA with forced convergence for different values of  $\theta$  and  $E_b = 1$ . LDPC code the same as for the results in Fig. 2.9.

The performance of offset MSA with FC is shown in Fig. 2.10. It exhibits a property similar to early termination of iterations, namely, the differential improvements of performance quickly diminish with increasing  $\theta$ .

### 2.3.2 Power - performance tradeoff in the digital base-band

Analyzing the tradeoff between the power and performance in digital circuits is not possible in closed form apart from some special cases. A multitude of different design techniques and the sheer complexity of digital designs prohibit such an analysis, and insights into the trends for particular techniques and applications need to be induced from a set of numerical simulations. Focusing on the particular case of LDPC decoding and different stopping criteria, it is seen that the complexity is linear in the number of iterations performed in the algorithm, while the performance is observed to improve monotonically and to saturate at a large number of iterations. Regarding the per-bit stopping criteria, complexity is a random variable with a generally unknown distribution, so dependence of both complexity and performance on the forced convergence threshold  $\theta$  needs to be determined through numerical simulations. Finding  $\theta$  that minimizes the *average* complexity under a performance constraint is possible through use of probabilistic optimization methods such as stochastic approximation, as discussed in more detail in Paper III.



## Chapter 3

# System Densification Under Hardware Constraints

As it was pointed out in Sec. 1.1, system densification, referring here to either a drastic increase of the number of BS antennas or an increase in spatial density of infrastructure nodes, has a beneficial influence on the energy efficiency and other cost efficiency metrics of a wireless network. However, great care needs to be taken when quantifying the benefits of densification. Insights into what constitutes an efficiently designed densified system can vary drastically depending on the choice of system parameters and cost metrics. Specifically, hardware-related constraints in practical system implementation are of primary importance here.

This chapter serves to provide a short overview of the basics of densified systems, and sheds a light on the determining influence that hardware constraints have on the benefits that these systems can provide. As elsewhere in this thesis, the main conclusion is that *being hardware-conscious is of paramount importance in system design*.

### 3.1 Massive MIMO

#### 3.1.1 Preliminaries

There exist slightly differing notions on what Massive MIMO (MaMI) is, and the definition of MaMI is rather flexible. In this work, the notion of MaMI corresponds to the one outlined in the work that introduced the MaMI concept, Marzetta's seminal paper [53]. This definition is sometimes referred to as the



“canonical MaMI”, as in the excellent monograph [54] which will be used as the main reference for the most important concepts described in this chapter. Canonical MaMI assumes a wireless network with  $L$  cells where

- the number of antennas  $M_j$  at the base station corresponding to cell  $j$  is much larger than the number of users in the cell,  $K_j$ . Typically,  $K_j$  is on the order of tens whereas  $M_j$  can be on the order of hundreds;
- space-division multiple access (SDMA) is used for multiplexing the  $K_j$  users in the same time-frequency resource;
- time division duplex (TDD) is used to separate uplink (UL) and downlink (DL) transmissions. Radio channel reciprocity of the UL and DL channels is assumed, which enables estimating the channel in the UL and using the estimates for formulating the DL precoder. UL and DL processing is linear, which is an asymptotically optimal strategy for  $M_j \rightarrow \infty$  [55].

Increasing the number of antennas  $M_j$  can, under proper conditions, lead to

- decrease of random variations of the channel from one user to all BS antennas [56], which is commonly referred to as **channel hardening** [57];
- channels from different users to BS becoming asymptotically orthogonal, usually referred to as **favorable propagation** [58].

Channel hardening leads to a gradual elimination of channel fading while favorable propagation eliminates interuser interference. Additionally, increasing the number of antennas serves to boost the array gain, which increases the post-processing  $SNR$ . As a result, spectral and energy efficiency of MaMI systems can be orders of magnitude larger than the conventional systems, as initial investigations have shown [56].

Initial investigations have, however, made use of an idealized channel model with spatially uncorrelated user channels. Recent theoretical results, supported by analyses conducted on measured channels, indicate that propagation properties of the channel, in particular spatial correlation, can diminish the favorable effects of MaMI and reduce the level of improvement over traditional systems [54, 59]. Moreover, the initial analyses only considered the transmitted power in calculating the energy efficiency, without taking into account power consumption spent on analog and digital signal processing. Finally, testbed architectures of MaMI BSs presented in [60–62] indicate that practical implementations of MaMI can be significantly constrained by hardware limitations, manifesting itself, among other things, as large overall power consumption and

high data interface throughput, both caused by the number of antennas being large.

In order to illustrate the importance of considering the limiting effects of hardware on the spectral efficiency (SE) and energy efficiency (EE) of MaMI systems, the analysis of the SE - EE tradeoff from [54] is given here in broad strokes. For more detailed information on different channel estimation techniques, linear processing schemes and technicalities of the mathematical analysis, the reader is referred to classical papers [55, 56, 63], as well as the monograph [54] and references given therein.

### 3.1.2 Energy efficiency - spectral efficiency tradeoff in massive MIMO

#### Asymptotic power scaling law

As an overture to the analysis of the EE - SE tradeoff in MaMI, a convenient scaling law is shortly discussed. It can serve to illustrate both the potential benefits of MaMI as well as the potential pitfalls in the analysis of these benefits when hardware effects are ignored. The initial version of the scaling law was derived for the UL in [56]; what is given here is its more general formulation from [54, sec. 5.2.1], derived for the DL.

To start with, the ergodic channel capacity of user  $k$  in cell  $j$  in the DL is lower bounded by [54, sec. 5.2.1]

$$\overline{SE}_{jk} = \frac{\tau_d}{\tau_c} \log_2 \left( 1 + \overline{SINR}_{jk} \right), \quad (3.1)$$

where  $\tau_d$  is the number of channel uses (symbols) dedicated to the DL data (payload),  $\tau_c$  the total number of symbols in a channel coherence block (encompassing UL channel estimation and UL and DL data transmission) and  $\overline{SINR}_{jk}$  is the ergodic signal-to-interference-and-noise ratio for the  $k$ th user in the  $j$ th cell.

Now, assume that  $M_j = M$  for all cells  $j$  and that per-user UL pilot power  $p_{jk}$  scales as  $M^{-\epsilon_1}$  and DL power per user  $\rho_{jk}$  as  $M^{-\epsilon_2}$ ,  $\epsilon_1, \epsilon_2 > 0$ . If maximum ratio transmission (MRT) is used, it can be shown that as  $M \rightarrow \infty$ ,  $\overline{SE}_{jk}$  has a nonzero asymptotic value if  $\epsilon_1 + \epsilon_2 < 1$ . This result indicates that, under the aforementioned conditions, the array gain is able to compensate for the scaling-down of pilot and DL transmit power. Therefore,  $p_{jk}$  and  $\rho_{jk}$  can e.g. simultaneously be scaled down *roughly* as  $1/\sqrt{M}$  and the lower bound on capacity will not go to zero as  $M \rightarrow \infty$ . This scaling law illuminates one of the benefits of MaMI, namely, that the per-antenna transmit powers in a MaMI system can be scaled down drastically compared to traditional cellular systems, resulting in power-efficient and “cheap” UEs and BSs.

How does this relate to the energy efficiency? In order to answer this question, energy efficiency needs to be defined first. In general terms, it can be defined as

$$EE = \frac{\text{Throughput [bit/s]}}{\text{Power consumption [W]}} \text{ [bits/Joule]}. \quad (3.2)$$

Now, for simplicity, only one cell with  $K$  users is considered. The throughput in (3.2) can be measured by the sum of the lower bounds (3.1), i.e. by the lower bound on *system sumrate*. Modeling the power consumption is not so straightforward, and the choice of what to include in the power consumption has a decisive impact on the conclusions that are made. If we decide to neglect the power consumption of analog and digital signal processing and only consider the transmitted power at the BS in the denominator of (3.2) (as it was done in [56]), the aforementioned scaling laws could lead us to believe that the energy efficiency grows unbounded as  $M \rightarrow \infty$ .

This is an attractive but misleading result. Power consumption of the circuits in the BS (and, if system EE is of concern, the power consumption of the UEs) cannot be neglected, since in a wide variety of cases it can dominate the transmitted power (which could e.g. happen in picocells where propagation distances and transmit powers tend to be small). Hardware power consumption modeling and its impact on the SE - EE tradeoff are addressed in the following section.

### EE - SE tradeoff for different power consumption models

Following the exposition in [54, sec. 5], in this section spectral efficiency (SE) is equivalent to the *lower bound on the ergodic achievable rate of a single user*. In setups where  $K = 1$ , SE will be used as a measure of throughput in (3.2); for  $K > 1$ , the throughput will be quantified by sumrate. As for the power consumption, the *hardware* power consumption (excluding the transmit power) is modeled using four different models, ordered by complexity and level of detail.

1. Hardware power consumption  $P_{\text{hardware}}$  is a fixed value  $P_{\text{fix}}$ , independent of  $M$ :

$$P_{\text{hardware}}^{(a)} = P_{\text{fix}}. \quad (3.3)$$

2.  $P_{\text{hardware}}$  consists of a fixed term and a term that scales linearly with  $M$ . The latter term can model the power consumption of all DL RF chains in the BS, where the consumption of a single chain is  $P_{\text{chain, BS}}$ :

$$P_{\text{hardware}}^{(b)} = P_{\text{fix}} + MP_{\text{chain, BS}}. \quad (3.4)$$

3.  $P_{\text{hardware}}$  subsumes a fixed term, a term that scales linearly with  $M$  and a term modeling the total power consumption of receiver hardware at the users, where each user consumes  $P_{\text{chain, user}}$ :

$$P_{\text{hardware}}^{(c)} = P_{\text{fix}} + MP_{\text{chain, BS}} + KP_{\text{chain, user}}. \quad (3.5)$$

4.  $P_{\text{hardware}}$  takes into account all power consumed in the RF and digital blocks, together with power consumed by backhaul and finally a fixed power consumption term:

$$P_{\text{hardware}}^{(d)} = P_{\text{fix}} + P_{\text{transceiver}} + P_{\text{channel estimation}} + P_{\text{coding/decoding}} \\ + P_{\text{load-dependent backhaul}} + P_{\text{spatial signal processing}}. \quad (3.6)$$

In the case of a single user,  $K = 1$  in a one-cell system,  $L = 1$ , the relation between  $EE$  and  $SE$  is given by [54, sec. 5.3.1]

$$EE = \frac{B \times SE}{(2^{SE} - 1) \frac{\nu}{M-1} + P_{\text{hardware}}}, \quad (3.7)$$

where  $B$  is system bandwidth and

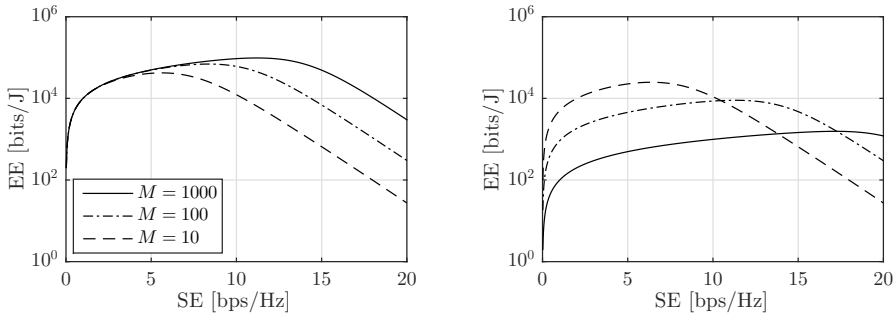
$$\nu = \frac{\sigma^2}{\mu\beta}, \quad (3.8)$$

with  $\sigma^2$  being the thermal noise power,  $\mu$  the efficiency of the power amplifier at the BS and  $\beta$  the average channel gain.

The tradeoff (3.7) is illustrated in Fig. 3.1 for  $K = 1$ ,  $L = 1$  and two different power consumption models  $P_{\text{hardware}}^{(a)}$  and  $P_{\text{hardware}}^{(b)}$ . Several things of interest can be observed. First, the  $EE$  -  $SE$  tradeoff is a unimodal function, i.e. it has a single distinct maximum ( $SE^*$ ,  $EE^*$ ). Second, the behaviour of this optimum on  $M$  depends heavily on which power consumption model is chosen. It can be shown [54, sec. 5.3.1] that

$$EE^* \approx \frac{eB}{1+e} \frac{\log_2(MP_{\text{hardware}})}{P_{\text{hardware}}}. \quad (3.9)$$

Relating this to hardware power consumption model (3.3), it can be seen that if the hardware is assumed to consume a constant amount of power,  $EE^*$  scales as  $\log_2(M)$ , i.e. grows unbounded with  $M$ . This is in line with the conclusions gathered from the scaling law in Sec. 3.1.2. On the other hand, if the more realistic power model (3.4) is used,  $EE^*$  is a unimodal function of  $M$ . This



**Figure 3.1:** EE - SE tradeoff (3.7) with hardware consumption model  $P_{\text{hardware}}^{(a)}$  used in the results on the left and  $P_{\text{hardware}}^{(b)}$  on the right hand side.  $\sigma^2/\beta = 24$  dBm,  $\mu = 0.4$ ,  $P_{\text{fix}} = 10$  W,  $P_{\text{chain}} = 1$  W.

means that the gains in spectral efficiency stemming from the increase of the number of antennas after a certain point get offset by the increase of hardware power consumption caused by the increased number of RF chains.

A similar conclusion can be made in the case  $K > 1$ ,  $L = 2$ , where numerical analysis shows that the optimal tradeoff  $(SE^*, EE^*)$  is unimodal with  $M/K$  [54, sec. 5.3.1]. Since  $SE$  depends on  $M$ , this implies that  $M$  and  $K$  maximizing  $EE$  can be found. This is done for a multicell, small-cell scenario with the detailed power consumption model (3.6) in [54, sec. 5.3.1] (a similar, more theoretical analysis is provided in [64]). For a wide range of parameters (different linear processing schemes, different generations of hardware with varying power consumption), it is shown that  $EE$  is optimized for  $K = 20$  to 30 and  $M/K$  ranging from 2 to 4.

In conclusion, optimizing the fundamental parameters (number of users and BS antennas) in MaMI networks heavily depends on the modeling of hardware power consumption. With a realistic power model, it can be shown that a) the energy efficiency does not grow unbounded as the number of antennas is increased (although it is still orders of magnitude better than in the traditional networks) and b) energy efficiency is maximized for a large, but *not extremely large* number of antennas.

### Massive MIMO and hardware impairments

In addition to being beneficial for system energy efficiency, MaMI also allows for improved hardware efficiency, i.e. use of hardware that introduces a higher level of impairments and consequently has lower cost and power consumption.

The level of impairments introduced by a general nonlinear block can be quantified using the hardware quality factor  $q$ , which is defined as follows. Assume that a zero-mean Gaussian signal  $x$  with variance  $p$  is fed to a memoryless nonlinear block. Applying the Bussgang theorem in this setup, the output signal of the block can be represented as

$$y = \sqrt{q}x + \eta, \quad (3.10)$$

where  $x$  and  $\eta$  are uncorrelated, and  $q$  is calculated as

$$q = \frac{\mathbb{E}\{yx^*\}}{\mathbb{E}\{|x|^2\}}. \quad (3.11)$$

Moreover, the additive distortion  $\eta$  is zero-mean and if the signal is processed such that  $\mathbb{E}\{|y|^2\} = p$ , its variance is equal to [54, Sec. 6]

$$\mathbb{E}\{|\eta|^2\} = (1 - q)p. \quad (3.12)$$

Nonlinear distortion at TX and RX chains of UEs and BSs in a MaMI system will result in various impairments affecting the spectral efficiency of the system. These impairments can be divided into two groups [54, Sec. 6]:

- Impairments that are noncoherently combined after linear processing, similarly to thermal noise or interuser interference in systems with perfect hardware. These impairments vanish asymptotically as  $M \rightarrow \infty$ .
- Impairments that combine coherently and therefore cannot be removed by the increase of number of antennas. These impairments depend exclusively on UE hardware quality.

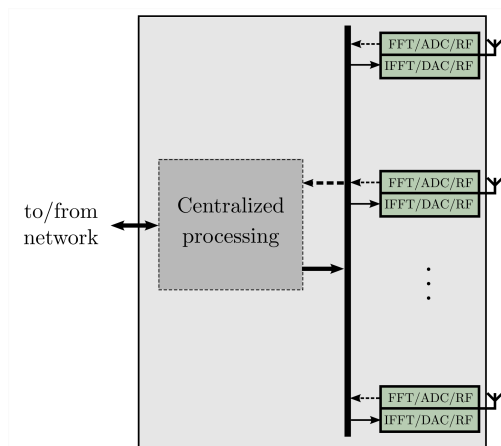
The asymptotic vanishing of noncoherently combined impairments gives rise to a hardware scaling law, described initially in [65]; the formulation given here is taken from [54, sec. 6.4]. This scaling law is similar to the power scaling law of Sec. 3.1.2, and is stated as follows. Assume that  $M_j = M$  for all cells  $j$ . Additionally, assume that the transmitter and receiver hardware quality factors  $q_t^{BS}$  and  $q_r^{BS}$  scale as  $M^{-\epsilon_1}$  and  $M^{-\epsilon_2}$ , respectively, where  $\epsilon_1, \epsilon_2 > 0$ . If maximum ratio transmission (MRT) is used, it can be shown that as  $M \rightarrow \infty$ ,  $\overline{SE}_{jk}$  has a nonzero asymptotic value if  $\epsilon_1 + \epsilon_2 < 1$ . The asymptotic value of  $\overline{SE}_{jk}$  is limited solely by UE hardware quality.

This scaling law implies that MaMI processing gains can compensate for an increased level of impairments caused by lowering the hardware quality. From the analyses in Sec. 2, it follows that low quality hardware can be expected to have a low power consumption. Hence, *there is a possibility* that the energy

efficiency of the system can be improved by reducing the hardware quality while increasing the number of antennas. This potential depends on the impact of hardware quality on  $SE$  for a finite  $M$ , the connection between hardware quality and the power consumption of a hardware block, and the power consumed in other parts of the system. An analysis of the effects of hardware quality on the energy efficiency in MaMI, focusing specifically on the energy-efficient choice of ADC resolution, is given in Paper II.

### 3.1.3 Interconnect throughput limitation

One implementational challenge that is peculiar to MaMI and is caused by hardware limitations is the high demand on the data throughput between the periphery of the BS (i.e. transceiver blocks close to the antennas) and the central data processing unit. This issue was observed already in the earliest attempts of MaMI BS implementation [60] and later BS test designs feature techniques particularly designed for overcoming it [62].



**Figure 3.2:** Conceptual diagram of a centralized base station architecture.

Defining and quantifying the interconnect throughput depends on the base station topology and on how the processing functions are divided between the periphery and center. For the purpose of illustration, we can assume a straightforward setup of a BS, shown in Fig. 3.2, for multicarrier MaMI. Peripheral units close to the antennas are assumed to perform all analog signal processing, AD/DA conversion, digital filtering and synchronization (not illustrated) and FFT/IFFT together with cyclic prefix removal/addition. Frequency domain

samples from/to all the antennas are then communicated over a bus to/from a centralized processing block, which will perform channel estimation, precoder/combiner calculation, precoding and combining, demodulation and channel decoding. It should be noted that maximum ratio combining/transmission does not ask for a centralized array topology since the channel from an antenna to all the users can be processed locally; on the other hand, the use of (regularized) zero forcing requires the channel state information to be present at a single central point in order for precoder/combiner to be formulated.

For both UL and DL transmission, in channel estimation as well as data transmission phases, each antenna will send/receive  $N_{\text{used}}$  complex samples over the bus, where  $N_{\text{used}}$  is the number of data-carrying subcarriers in an OFDM symbol. With  $w$  denoting the number of bits used to represent a complex sample and  $T_{\text{OFDM}}$  being the total duration of an OFDM symbol, the throughput of the bus is

$$R_{\text{bus}} = \frac{MwN_{\text{used}}}{T_{\text{OFDM}}}. \quad (3.13)$$

Assuming a LTE-like multicarrier setup used in [62] with sampling rate of 30.72 MHz,  $T_{\text{OFDM}} = 71.7 \mu\text{s}$  and  $N_{\text{used}} = 1200$ , and additionally assuming that  $w = 32$  (16 bits per I and Q sample), the required throughput of the bus for  $M = 128$  is 8.6 GBps.

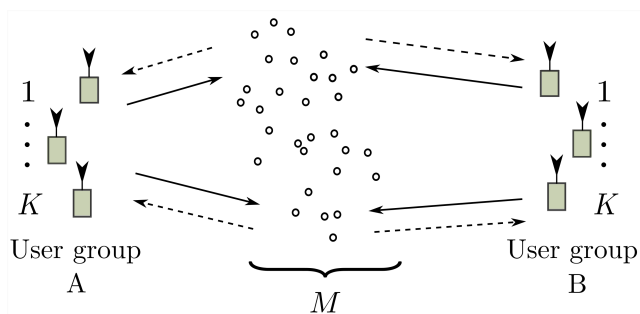
A throughput of this magnitude might prove to be unsupportable by standard radio interfaces such as the Common Public Radio Interface (CPRI) [66]. Even with advances in radio interface technology, the interconnect throughput limitation will never allow for an arbitrarily large number of antennas in a centralized base station topology. Alternative techniques of array data processing, as well as alternative array topologies need to be investigated with the goal of relaxing the interconnect throughput requirement. Paper IV presents a step in this direction.

## 3.2 Large-scale multipair two-way relay networks

As it was pointed out in Sec. 1.1, network energy efficiency can be boosted by increasing the number of infrastructure nodes, e.g. base stations or relays. When it comes to relays, one particular system setup involving relays that bears similarities to MaMI shows promising qualities. This system setup assumes two groups of  $K$  users, group A and group B, where each user from group A has a corresponding user from group B with whom it intends to exchange information but cannot do so via a direct link. Instead, the  $K$  user pairs are assisted in communication by a set of  $M$  single-antenna relays (or one relay



with  $M$  antennas). The traffic from one group to the other and vice versa is symmetric, and the exchange of information occurs in consecutive pairs of time slots. In the first slot of one pair, both groups transmit the information to the relays (UL transmission). The relays process the received information in an amplify-and-forward fashion and send the result to the intended users in the second slot of the pair (DL transmission). By keeping all the channels between users and relays active in every time slot, this method, referred to as two-way relaying, effectively doubles the throughput compared to traditional relaying. A sketch of a multipair two-way network is given in Fig. 3.3.



**Figure 3.3:** A multipair two-way network with  $2K$  users and  $M$  relays. Full arrows: UL transmission phase, dashed arrows: DL transmission phase.

Multipair two way networks where the number of relays  $M$  is large have some of the appealing properties of MaMI systems. A large scale multipair two way (LS - MTW) network where the  $M$  single-antenna relays are not exchanging CSI or data but only amplifying and forwarding their received scalar symbols is analyzed in [67]. It is shown that, under the condition of perfect CSI and as  $M \rightarrow \infty$ , transmit powers of users and relays can simultaneously scale as  $1/M$  and system sumrate will still tend to a nonzero value. The same scaling law was derived in [68] for a MaMI relay, which is equivalent to a setup where all single-antenna relays would send the CSI and received data to a central point where joint processing would be performed. Now, similarly to what was presented in Sec. 3.1.2, if only the transmit powers are used in the system energy efficiency metric, this scaling law could result in a conclusion that the energy efficiency of such a system will grow without bounds with  $M$ .

Although there does not seem to exist an exhaustive analysis of EE - SE tradeoff in LS - MTW networks, the similarities with MaMI strongly indicate that, as in MaMI, EE of the network will be maximized for some  $(M^*, K^*)$  when a realistic power consumption model is used. In this thesis, it is assumed

that the working point  $(M, K)$  is fixed according to some criterion and the focus is put on a different problem.

Namely, by comparing the numerical results in [67] and [68] for matching system parameters, it can be observed that if all  $M$  relays from Fig. 3.3 behave like a MaMI relay (where the CSI information and received data from all the relays is centrally processed), the sumrate is vastly improved compared to the case when each relay does its local amplification and forwarding. Intuitively, it makes sense that cooperation is rewarded by an improvement in performance. However, centralized processing of information does come with a price. In one way or another, some system resources will be spent on collecting the data in the central processing node and sending it back to the relays after it has been processed. For example, if relays are communicating wirelessly, some energy will be spent on sending the data from the relays to the central processing node and vice versa, with more energy being consumed by a relay system covering a larger area. Likewise, some bandwidth might be dedicated for communicating the overhead data, and it must not necessarily be in the same band, or of the same size as the bandwidth dedicated to UL and DL transmissions. The consumption of resources by this overhead communication can be regarded as another case of hardware-related (or, more precisely, implementation-related) limitation.

The idea examined in Paper V of this thesis is based on dividing the  $M$  relays into groups and making all relays in the group exchange information so that they operate as a MIMO relay. All such groups will process the information independently of other groups, without exchange of information. What is examined is how resource consumption for overhead communication trades off with performance, and in Paper V, the focus is on how system sumrate trades with bandwidth dedicated for overhead communication as the relay group size changes.



## Chapter 4

# Paper Summary and Discussion

This chapter presents a concise summary of work performed in each of the five original research papers contained in the thesis. For each paper, main findings and novel contributions are highlighted. Also, remarks are made on my personal contributions as the author in each of the papers.

### 4.1 Research contributions

#### 4.1.1 Paper I: Low Power Receiver Front Ends: Scaling Laws and Applications

The overarching idea of this work is bridging the gap between circuit and communication theory by combining known theoretical results from both fields in a unified theoretical framework for analyzing the tradeoffs between power consumption and performance of the analog front end (AFE). The starting point of the analysis is the circuit-theoretic result (2.10), linking the minimum power consumption of the AFE with noise PSD and IIP3 power. Based on this fundamental result, the main contributions of the paper are:

- Derivation of an approximate scaling law between (optimal) power consumption of the AFE and the  $SNDR$ . AFE power consumption is shown to approximately scale as  $SNDR^{3/2}$ ;
- Using the derived scaling law to determine how AFE power consumption scales with constellation size of the digital modulation and with the rate

and coding gain of error control codes. AFE energy efficiency is shown to always improve with decreasing size of square QAM constellation. Additionally, a numerical analysis utilizing extrapolated power numbers from implementations of error control decoders and the derived scaling laws suggests that codes with moderate coding gains and simple, noniterative decoding algorithms (such as convolutional codes) maximize the energy efficiency of the entire receiver;

- Using side results on scaling of AFE power with received power and OOB interference power, it is shown that a receiver that continuously adapts its structure to match the received power or interferer power level can achieve a vast reduction of power consumption (e.g. 20x) compared to a non-adaptive design. Practical implementations of such receivers, utilizing only two steps of adaptation (low power/high power) are proposed in a general form, and the loss in power savings compared to continuous adaptation is analyzed.

**Personal contribution:** I performed the complete theoretical analysis, structuring and organization of the material and wrote the paper.

#### 4.1.2 Paper II: When are Low Resolution ADCs Energy Efficient in Massive MIMO?

The main motivation behind the investigation presented in this work was a series of claims found in the academic literature that the use of 1-bit ADCs is a strategy that is beneficial from the point of view of energy efficiency in massive MIMO. A parametric power consumption model for the massive MIMO base station is set up, relating the power consumption of the ADCs to the power consumed by the rest of the constituent blocks in the base station. Circuit-theoretic models are used for the power consumption of the ADC, and in setting up the analysis, particular care is taken of including assumptions and parameters of practical relevance. Main findings of the paper are:

- Energy efficiency of the entire base station is not maximized at low bit resolutions unless the power consumed by the ADCs is comparable to the power consumed by all the other hardware blocks, which is unlikely in practice;
- For practical ratios of ADC power consumption and power consumed by other blocks, energy efficiency is typically optimized at 4 - 8 bits of resolution;

- Making the number of antennas extremely large makes lower resolutions energy-optimal only when the number of users is kept constant. On the other hand, if antenna and user numbers are scaled up simultaneously with keeping the antenna/user ratio constant, the energy-optimal resolution stays constant or even *increases* with this scaling;
- Energy-optimal bit resolution can significantly increase with presence of unfiltered OOB blockers, which is a common concern in practice.
- The overall findings advise against the use of low bit resolutions as means of improving system energy efficiency in the uplink of massive MIMO.

**Personal contribution:** Structuring and organization of the material have in most part been done by myself with additional input from the last author. I performed the complete theoretical and numerical analysis and wrote the paper.

### 4.1.3 Paper III: Modified Forced Convergence Decoding of LDPC Codes with Optimized Decoder Parameters

This paper looks into LPDC decoding with forced convergence (FC) and elaborates how the FC threshold  $\theta$  should be determined. It proposes setting the value of  $\theta$  such that the computational complexity of the decoding is minimized subject to a block error rate constraint. The main challenge identified with this approach is that the functional dependence of complexity and error rate on  $\theta$  cannot be determined in closed form but only estimated from observations in runtime. The main contributions of the work are:

- A small modification of the way bit-to-check messages are calculated is introduced, which results in a complexity reduction;
- The stochastic approximation (SA) algorithm is proposed to be used for determining the value of  $\theta$  that minimizes the average computational complexity under a block error rate constraint;
- The described structure (SA algorithm wrapped around a tunable LDPC decoding algorithm) is noted to be a natural way of implementing a decoder that adapts in real-time to the time variations in  $SNR$ . Such continuous tuning of the decoder to  $SNR$  is found to yield an additional 7 - 12 % reduction in complexity compared to an optimized but fixed LDPC decoder with FC, for the analyzed set of codes. Additionally, the settling

time of this adaptive structure is analyzed numerically, and it is concluded that the SA-based adaptation of the LDPC decoder is generally applicable in slowly varying channels.

**Personal contribution:** I performed the theoretical analysis and simulations and wrote the paper.

#### 4.1.4 Paper IV: Fully Decentralized Approximate Zero-Forcing Precoding for Massive MIMO Systems

The analysis in this paper is motivated by the fact that centralized massive MIMO processing at the base station side imposes extremely large throughput requirements on the interconnects/backhaul. A daisy chain of single-antenna units with local channel knowledge is investigated as an alternative base station topology. The main contributions of the paper are:

- Derivation of a decentralized linear precoding algorithm that suppresses interuser interference. The algorithm is specifically tailored to the limitations of the daisy chain topology;
- Theoretical analysis of the mechanisms of operation of the derived algorithm;
- The algorithm is shown to be close in error rate performance to zero-forcing precoding for a very large number of antennas at the base station, while also significantly outperforming maximum ratio transmission precoding. Latency of the precoder formulation is shown not to be limiting for the applicability of the algorithm when calculations are carefully scheduled over antennas and subcarriers in a multicarrier setting. For a very large number of antennas, the throughput of the links inside the daisy chain is shown to be much lower than the throughput needed for backhaul in a centralized topology;
- Overall, the results indicate that the daisy chain topology should be chosen for implementing massive MIMO base stations when the number of base station antennas is extremely large.

**Personal contribution:** I proposed the use of the daisy chain topology and derived the precoding algorithm together with the second author. I performed the theoretical analysis and simulations and wrote the paper.

### 4.1.5 Paper V: Impact of Relay Cooperation on the Performance of Large-scale Multipair Two-way Relay Networks

The paper analyzes relay networks where equally-sized groups of users exchange messages via a relay system with a large number of relays. The relay system can attain the form of a single multiantenna relay, or be embodied by single-antenna relays that individually perform information processing. As in **Paper IV**, the cost of centralized processing at the relays is a major implementation issue. The investigation in the paper focuses on clustering the relays in a number of groups that perform information processing individually and where the relays inside a group closely cooperate, without any exchange of information between the groups. Performance of such a system as a function of the group size is analyzed. The main contributions of the paper are:

- Derivation of a lower bound on the ergodic sumrate of the described system when zero-forcing processing is used at the relays to process the uplink and downlink signals;
- From the analysis of this lower bound, it is concluded that the performance (in this case sumrate) is approximately independent of the group size when the group size is much larger than the number of user pairs;
- The preceding observation indicates that a large centralized group of relays (alternatively, a massive MIMO relay) can be broken down into smaller constituent groups (smaller-sized MIMO relays) with only a marginal performance degradation;
- If the resource used to support the centralization of the data processing at each group is assumed reusable (e.g. bandwidth), then the resource efficiency is maximized for small sizes of the constituent group.

**Personal contribution:** Structuring and organization of the material have been done by myself. I also performed the theoretical analysis and simulations and wrote the paper.

## 4.2 Discussion and future work

This thesis represents a selection of topics in wireless communication system design where hardware-related implementation limitations are explicitly taken into account during the system design process. It examines the improvements



resulting from making the communication hardware adaptive to the environment, or by incorporating fundamental hardware limitations in the system design so that they become an integral part of the structure.

With this thesis work being a rather eclectic collection of topics, it is natural that all aspects of interest have not been covered in depth in all of the analyses since that would far exceed a format of an ordinary PhD thesis. Indeed, the list of things that could have additionally been looked into is extensive. Here is a short, subjective selection of the most interesting and attractive extensions of the work presented in the thesis that can form the topics of future work:

- The analysis in **Paper I** is for a single-antenna receiver. It would be of interest to use the laws derived there in a multi-antenna context. More specifically, employing these power-performance laws in the uplink of massive MIMO, at the base station and calculating the tradeoff of overall system power with performance as the number of antennas grows. Because of uplink/downlink duality, the scaling laws from **Paper I** could likewise be used at the single-antenna terminals. Massive MIMO is known to enable relaxations in terminal hardware, so these laws could give an estimate how much overall power in the system would be saved by scaling up the number of antennas and simultaneously relaxing the performance requirements in user hardware.
- There are many topics of interest that appear in the context of the daisy chain topology, with only the most basic ones actually having been analyzed in **Paper IV**. Aspects of uplink detection, clustering of antennas so that they share CSI, connecting the first and last antenna units directly so that they form a ring and then performing iterative processing on that ring are some of the most interesting topics. Also, after all the extensive analyses it would be good to see the benefits of the daisy chain topology verified in an actual massive MIMO testbed implementation.
- Similarly to **Paper IV**, **Paper V** has only “scratched the surface” when it comes to the analysis of relay grouping. The impact of inter-group synchronization on system performance is one aspect of practical relevance that can be analyzed in future work. Also, linear processing schemes other than zero-forcing can be considered.

# Bibliography

- [1] G. E. Moore, "Cramming More Components Onto Integrated Circuits," in *IEEE Solid-State Circuits Society Newsletter*, vol. 11, no. 3, pp. 33-35, Sept. 2006 (Reprinted from *Electronics*, volume 38, number 8, April 19, 1965, pp.114 ff.)
- [2] <https://data.worldbank.org/>
- [3] <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.pdf>
- [4] A. Fehske, G. Fettweis, J. Malmodin and G. Biczok, "The Global Footprint of Mobile Communications: The Ecological and Economic Perspective," in *IEEE Communications Magazine*, vol. 49, no. 8, pp. 55-62, August 2011.
- [5] "NGMN alliance 5G white paper," 2015 [Online]. Available: <https://www.ngmn.org/5g-white-paper/5g-white-paper.html>
- [6] J. G. Andrews *et al.*, "What Will 5G Be?," in *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065-1082, June 2014.
- [7] C. Han *et al.*, "Green Radio: Radio Techniques to Enable Energy-Efficient Wireless Networks," in *IEEE Communications Magazine*, vol. 49, no. 6, pp. 46-54, June 2011.
- [8] <https://www.iea.org/>
- [9] S. Buzzi, C. I, T. E. Klein, H. V. Poor, C. Yang and A. Zappone, "A Survey of Energy-Efficient Techniques for 5G Networks and Challenges Ahead," in *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 697-709, April 2016.
- [10] <http://www.oed.com>

- 
- [11] S. Cui, A. J. Goldsmith and A. Bahai, "Energy-Constrained Modulation Optimization," in *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349-2360, Sept. 2005.
- [12] J. C. Rudell *et al.*, "Recent Developments in High Integration Multi-Standard CMOS Transceivers for Personal Communication Systems," in *Proceedings of the 1998 International Symposium on Low Power Electronics and Design* (IEEE Cat. No.98TH8379), Monterey, CA, USA, 1998, pp. 149-154.
- [13] Y. Tsvividis, N. Krishnapura, Y. Palaskas and L. Toth, "Internally Varying Analog Circuits Minimize Power Dissipation," in *IEEE Circuits and Devices Magazine*, vol. 19, no. 1, pp. 63-72, Jan. 2003.
- [14] Y. Tsvividis, "Signal-to-Noise Ratio, Dynamic Range, and Power Dissipation: Paying Attention to Their Interrelation Can Greatly Benefit Analog Circuit Design," in *IEEE Solid-State Circuits Magazine*, vol. 10, no. 4, pp. 60-69, Fall 2018.
- [15] A. A. Abidi, G. J. Pottie and W. J. Kaiser, "Power-Conscious Design of Wireless Circuits and Systems," in *Proceedings of the IEEE*, vol. 88, no. 10, pp. 1528-1545, Oct. 2000.
- [16] B. Razavi, *RF Microelectronics*, 2nd ed. New York, NY, USA: Prentice-Hall, 2011.
- [17] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, 2nd ed. New York, NY, USA: Wiley, 2011.
- [18] W. Sheng, A. Emira and E. Sanchez-Sinencio, "CMOS RF Receiver System Design: a Systematic Approach," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 5, pp. 1023-1034, May 2006.
- [19] A. V. Do *et al.*, "An Energy-Aware CMOS Receiver Front End for Low-Power 2.4-GHz Applications," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 10, pp. 2675-2684, Oct. 2010.
- [20] S. Sen, D. Banerjee, M. Verhelst and A. Chatterjee, "A Power-Scalable Channel-Adaptive Wireless Receiver Based on Built-In Orthogonally Tunable LNA," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 946-957, May 2012.
- [21] A. Tasic *et al.*, "Design of Adaptive Multimode RF Front-End Circuits," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 2, pp. 313-322, Feb. 2007.

- [22] F. Behbahani *et al.*, “Adaptive Analog IF Signal Processor for a Wide-Band CMOS Wireless Receiver,” in *IEEE Journal of Solid-State Circuits*, vol. 36, no. 8, pp. 1205-1217, Aug. 2001.
- [23] M. T. Ozgun, Y. Tsvividis and G. Burra, “Dynamic Power Optimization of Active Filters with Application to Zero-IF receivers,” in *IEEE Journal of Solid-State Circuits*, vol. 41, no. 6, pp. 1344-1352, June 2006.
- [24] A. Yoshizawa and Y. Tsvividis, “A Channel-Select Filter With Agile Blocker Detection and Adaptive Power Dissipation,” in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 5, pp. 1090-1099, May 2007.
- [25] G. Hueber, J. Zipper, R. Stuhlberger and A. Holm, “An Adaptive Multi-Mode RF Front-End for Cellular Terminals,” in *Proceedings of the 2008 IEEE Radio Frequency Integrated Circuits Symposium*, Atlanta, GA, 2008, pp. 25-28.
- [26] S Sen *et al.*, “Real-Time Blocker-Adaptive Broadband Wireless Receiver for Low-Power Operation Under Co-Existence in 5G and Beyond”, U.S. Patent 9698838, Jul. 7, 2017.
- [27] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA, USA: Kluwer, 1992.
- [28] A. Sripad and D. Snyder, “A Necessary and Sufficient Condition for Quantization Errors to be Uniform and White,” in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 5, pp. 442-448, October 1977.
- [29] A. K. Fletcher, S. Rangan, V. K. Goyal and K. Ramchandran, “Robust Predictive Quantization: Analysis and Design Via Convex Optimization,” in *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 618-632, Dec. 2007.
- [30] C. Mollén, J. Choi, E. G. Larsson and R. W. Heath, “Achievable Uplink Rates for Massive MIMO With Coarse Quantization” in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, LA, 2017, pp. 6488-6492.
- [31] J. Max, “Quantizing For Minimum Distortion,” in *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7-12, March 1960.
- [32] J. Bucklew and N. Gallager, “Some Properties of Uniform Step Size Quantizers (Corresp.),” in *IEEE Transactions on Information Theory*, vol. 26, no. 5, pp. 610-613, September 1980.

- [33] C. Svensson, "Towards Power Centric Analog Design," in *IEEE Circuits and Systems Magazine*, vol. 15, no. 3, pp. 44-51, third quarter 2015.
- [34] B. Murmann, "Energy Limits in A/D Converters," in *Proceedings of 2013 IEEE Faible Tension Faible Consommation*, Paris, 2013, pp. 1-4.
- [35] B. J. Hosticka, "Performance Comparison of Analog and Digital Circuits," in *Proceedings of the IEEE*, vol. 73, no. 1, pp. 25-29, Jan. 1985.
- [36] E. A. Vittoz, "Future of Analog in the VLSI Environment," in *Proceedings of the IEEE International Symposium on Circuits and Systems*, New Orleans, LA, USA, 1990, pp. 1372-1375 vol.2.
- [37] B. Murmann, "The Race for the Extra Decibel: A Brief Review of Current ADC Performance Trajectories," in *IEEE Solid-State Circuits Magazine*, vol. 7, no. 3, pp. 58-66, Summer 2015.
- [38] T. Sundstrom, B. Murmann and C. Svensson, "Power Dissipation Bounds for High-Speed Nyquist Analog-to-Digital Converters," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 509-518, March 2009.
- [39] R. H. Walden, "Analog-to-Digital Converter Survey and Analysis," in *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539-550, April 1999.
- [40] R. Schreier and G. C. Temes, *Understanding Delta-Sigma Data Converters*. New York: Wiley, 2005.
- [41] B. Murmann. ADC performance survey 1997-2018. [Online]. Dec. 2018. Available: <http://web.stanford.edu/~murmam/adcsurvey.html>
- [42] D. Markovic, V. Stojanovic, B. Nikolic, M. A. Horowitz and R. W. Brodersen, "Methods for True Energy-Performance Optimization," in *IEEE Journal of Solid-State Circuits*, vol. 39, no. 8, pp. 1282-1293, Aug. 2004.
- [43] D. Markovic, B. Nikolic and R. W. Brodersen, "Power and Area Minimization for Multidimensional Signal Processing," in *IEEE Journal of Solid-State Circuits*, vol. 42, no. 4, pp. 922-934, April 2007.
- [44] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
- [45] T. Richardson and S. Kudekar, "Design of Low-Density Parity Check Codes for 5G New Radio," in *IEEE Communications Magazine*, vol. 56, no. 3, pp. 28-34, March 2018.

- 
- [46] 3GPP TSG RAN Meeting #71, RP-160671, “New SID Proposal: Study on New Radio Access Technology,” NTT DOCOMO Inc., Göteborg, Sweden, 7–10 Mar. 2016.
- [47] R. Tanner, “A Recursive Approach to Low Complexity Codes,” in *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533-547, September 1981.
- [48] J. Chen, A. Dholakia, E. Eleftheriou, M. P. C. Fossorier and Xiao-Yu Hu, “Reduced-Complexity Decoding of LDPC Codes,” in *IEEE Transactions on Communications*, vol. 53, no. 8, pp. 1288-1299, Aug. 2005.
- [49] A. Darabiha, A. Chan Carusone and F. R. Kschischang, “Power Reduction Techniques for LDPC Decoders,” in *IEEE Journal of Solid-State Circuits*, vol. 43, no. 8, pp. 1835-1845, Aug. 2008.
- [50] E. Zimmermann, P. Pattisapu, P. K. Bora, and G. Fettweis, “Reduced Complexity LDPC Decoding Using Forced Convergence,” in *Proceedings of the 7th International Symposium on Wireless Personal Multimedia Communications*, Padova, Italy, Sep. 2004, vol. 3, pp. 243246, WA2-2.
- [51] E. Zimmermann, P. Pattisapu and G. Fettweis, “Bit-Flipping Post-Processing for Forced Convergence Decoding of LDPC Codes,” in *Proceedings of the 2005 13th European Signal Processing Conference*, Antalya, Turkey, 2005, pp. 1-4.
- [52] A. Blad, O. Gustafsson and L. Wanhammar, “An Early Decision Decoding Algorithm for LDPC Codes Using Dynamic Thresholds,” in *Proceedings of the 2005 European Conference on Circuit Theory and Design*, 2005., Cork, Ireland, 2005, pp. III/285-III/288 vol. 3.
- [53] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” in *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, November 2010.
- [54] E. Björnson, J. Hoydis, and L. Sanguinetti, “Massive MIMO networks: Spectral, Energy, and Hardware Efficiency,” *Foundations and Trends in Signal Processing*, vol. 11, nos. 3–4, pp. 154–655, 2017.
- [55] F. Rusek et al., “Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays,” in *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40-60, Jan. 2013.

- [56] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," in *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436-1449, April 2013.
- [57] H. Q. Ngo and E. G. Larsson, "No Downlink Pilots Are Needed in TDD Massive MIMO," in *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2921-2935, May 2017.
- [58] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, "Aspects of Favorable Propagation in Massive MIMO," in *Proceedings of the 2014 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, 2014, pp. 76-80.
- [59] X. Gao, O. Edfors, F. Rusek and F. Tufvesson, "Massive MIMO Performance Evaluation Based on Measured Propagation Data," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3899-3911, July 2015.
- [60] C. Shepard, H. Yu, N. Anand, E. Li, T. Marzetta, R. Yang, and L. Zhong, "Argos: Practical Many-Antenna Base Stations," in *Proceedings of the 2012 ACM International Conference on Mobile Computing and Networking (Mobicom '12)*, pp. 53-64.
- [61] J. Vieira et al., "A Flexible 100-Antenna Testbed for Massive MIMO," in *Proceedings of the 2014 IEEE Globecom Workshops (GC Wkshps)*, Austin, TX, 2014, pp. 287-293.
- [62] S. Malkowsky et al., "The World's First Real-Time Testbed for Massive MIMO: Design, Implementation, and Validation," in *IEEE Access*, vol. 5, pp. 9073-9088, 2017.
- [63] H. Yang and T. L. Marzetta, "Performance of Conjugate and Zero-Forcing Beamforming in Large-Scale Antenna Systems," in *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 172-179, February 2013.
- [64] E. Björnson, L. Sanguinetti, J. Hoydis and M. Debbah, "Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer?," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059-3075, June 2015.
- [65] E. Björnson, M. Matthaiou and M. Debbah, "Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353-4368, Aug. 2015.

- 
- [66] [http://www.cpri.info/downloads/CPRI\\_v\\_7\\_0\\_2015-10-09.pdf](http://www.cpri.info/downloads/CPRI_v_7_0_2015-10-09.pdf)
- [67] H. Q. Ngo and E. G. Larsson, "Large-Scale Multipair Two-Way Relay Networks with Distributed AF Beamforming," in *IEEE Communications Letters*, vol. 17, no. 12, pp. 1-4, December 2013.
- [68] T. V. T. Le and Y. H. Kim, "Power and Spectral Efficiency of Multi-Pair Massive Antenna Relaying Systems With Zero-Forcing Relay Beamforming," in *IEEE Communications Letters*, vol. 19, no. 2, pp. 243-246, Feb. 2015.





## Part II

# Included Papers



# *Paper I*



# Low Power Receiver Front Ends: Scaling Laws and Applications

In this paper, we combine communication-theoretic laws with known, practically verified results from circuit theory. As a result, we obtain closed-form theoretical expressions linking fundamental system design and environment parameters with the power consumption of analog front ends for communication receivers. This collection of scaling laws and bounds is meant to serve as a theoretical reference for practical low power front end design. In one set of results, we first find that the front end power consumption scales at least as  $SNDR^{3/2}$  if environment parameters (fading and blocker levels) are static. The obtained scaling law is subsequently used to derive relations between front end power consumption and several other important communication system parameters, namely, digital modulation constellation size, symbol error probability, error control coding gain and coding rate. Such relations, in turn, can be used when deciding which system design strategies to adopt for low-power applications. For example, if error control coding is employed, the most energy-efficient strategy for the entire receiver is to use codes with moderate coding gain and simple decoding algorithms, such as convolutional codes. In another collection of results, we find how front end power scales with environment parameters if the performance is kept constant. This yields bounds on average power reduction of receivers that adapt to the communication environment. For instance, if a receiver front end adapts to fading fluctuations while keeping the performance above some given minimum requirement, power can theoretically be reduced at least 20x compared to a non-adaptive front end.



# 1 Introduction

Low power consumption is one of the main design targets for communication receivers, and its importance is especially high when it comes to wireless devices, which are often battery-powered and therefore energy-limited. At the same time, receivers also need to satisfy some performance requirement, such as minimum throughput and maximum bit error rate (BER).

Designing receivers that jointly meet power consumption and performance criteria tends to be predominantly based on the experience of hardware designers. Additionally, more often than not, receiver designs are optimized based on the worst-case scenario of operation (the most adverse possible combination of environment conditions under which satisfying performance must be delivered). The latter, conservative design trend in particular is what prevents hardware designs from exploiting their full potential for low-power operation.

It would be of significant interest to be able to theoretically predict how much power would be consumed by a receiver with certain performance requirements, with all system and environment constraints taken into consideration. Such a result would serve as a benchmark and motivation for both practical hardware and system design, an indicator of how low the power consumption can really be driven. If combined with the knowledge of the statistical properties of environment variables, it could also provide a measure of how much power can be saved if the receiver adapts to the communication environment.

The analog front end (AFE) (the chain of analog signal blocks of the receiver excluding the oscillator) typically has a defining impact on the overall performance of the receiver, while also consuming a substantial portion of its power. One of the main questions of low-power receiver design can thus be formulated as

*“How does the power consumption of an analog front end (AFE) of a receiver scale with performance?”*

If this question is answered, some important follow-up questions can be answered as well, such as:

- If the overall system design features techniques that serve to improve performance (e.g. by use of error control coding) and this opens up the possibility of relaxing the design of the AFE, how much power do we save by performing this relaxation?
- If the receiver is made adaptive to communication environment conditions (e.g. channel fading or out-of-band interference), how much AFE power can be saved, on average, compared to a design based on worst-case conditions?



The theoretical analysis of the relation between analog circuit power consumption and performance appears not to have gained a lot of attention in the scientific community. The relation between power consumption and performance for individual analog blocks is analyzed in [1] and [2]. It is found that power consumption grows linearly with the dynamic range of an analog circuit block<sup>1</sup>. Analysis of this relation for a chain of analog blocks becomes rather involved because performance metrics for the entire chain exhibit a complex dependence on gain, noise and linearity properties for individual blocks. Moreover, there are practically infinitely many combinations of per-block parameters that satisfy the performance requirements for the entire chain, with each combination resulting in a unique power consumption. A sensible approach is then to find the combination that yields the minimum power consumption, which then makes it possible to reveal the implicit or explicit connection between the performance requirement and the obtained optimal power consumption. In [3] and [4] this approach is adopted, with the focus being mostly on how to conveniently model the power - performance relation for individual blocks and how to solve the optimization problem. The analysis in [5] extends the ideas from [4], with the power - performance relation also being given some treatment in the context of communication systems.

There also exists a body of academic work [6] - [14] that examines the topic of environment-adaptive AFEs and receivers, with the focus being primarily on practical hardware implementations. It is demonstrated that adaptive receivers are implementable, and various implementation strategies are suggested. Furthermore, measured power numbers from these designs indicate that substantial power reduction is attainable if environment-adaptive receiver techniques are adopted.

What is found to be largely missing in the existing literature is a work that takes the power-performance laws from circuit theory and combines them with classical results from communication theory to formulate joint circuit-communication-theoretical laws of system behavior<sup>2</sup>. With such laws at hand, system design questions such as “if the BER requirement is relaxed from  $10^{-6}$  to  $10^{-3}$  and we redesign the receiver to meet the new requirements, how much power is this new receiver expected to consume?” can be answered in a precise and immediate fashion, without resorting to educated guessing or iterative hardware redesign and simulation/measurement cycles. Moreover, by taking into account the influence of environment conditions on front end power consumption, it would be possible to precisely determine power savings obtainable when making the receiver environment-adaptive.

---

<sup>1</sup> The definitions of dynamic range differ slightly between these two papers. As will be shown, here we adhere to the definition given in [1]. <sup>2</sup> A rare example is [5], but with the analysis limited only to the connection between throughput and relative level of the out-of-band blocking signal.

Here we aim at bridging this gap between circuit and communication theory. The idea is to obtain theoretical expressions that will describe how optimal AFE power consumption scales with important system and environment parameters. More specifically, we are interested in finding out the scaling of front end power with the signal-to-noise-and-distortion ratio (SNDR), representing system performance, when the environment parameters (fading and out-of-band interference) do not exhibit temporal changes. Conversely, we also aim to describe how AFE power scales with environment parameters when SNDR is kept constant. The obtained set of fundamental scaling laws can then be used to build up a more extensive system level analysis. We derive our scaling laws from a known relation between performance and minimum power consumption for AFEs, presented and verified in actual hardware implementations in [3]. This relation is modified so that it can be seamlessly combined with communication-theoretic laws. One set of results is based on a novel scaling law we obtain, namely, that AFE power consumption scales at least as  $SNDR^{3/2}$ . This result is then employed in finding closed-form expressions for AFE power scaling with QAM constellation size, symbol error rate and error control coding gain, which are further used to decide on appropriate system-level strategies for low-power design. In another line of results, we obtain power-law type relations between AFE power and environment parameters. These are combined with fading and blocker statistics, yielding important theoretical bounds on average power savings of environment-adaptive front ends, which demonstrate that substantial power savings are possible if the environment-adaptive design approach is adopted.

Throughout the course of our analysis, we rely on the fact that the fundamental results we build upon have been verified in practical front end implementations and we do not aim at recreating these verifications. Instead, we put the focus on laying out a general theoretical framework for low power receiver design and showing the advantages of environment-adaptive designs, which will hopefully make this work both a point of reference and motivation for future research efforts in the area of practical hardware implementations of such systems.

## 2 Optimal power consumption of analog front ends

Let us observe a chain of analog circuit blocks that form the front end of a communications receiver. One example of such a chain can be the direct conversion receiver with the structure LNA - downconversion mixer - channel select filter - variable gain amplifier. While the direct conversion receiver is

given as an example, we emphasize that the forthcoming analysis holds for any type of receiver chain.

Each of the blocks in the chain can be qualitatively characterized by noise and linearity properties (serving as performance quantifiers) and by an associated power consumption. Noise performance is commonly quantified by noise power spectral density  $\bar{V}_N^2$  [ $V^2/\text{Hz}$ ], and linearity by  $V_{\text{IIP3}}^2$  [ $V^2$ ], the input-referred third-order intercept voltage squared. Additionally, we denote by  $F_{\text{AFE}}$  the total noise factor and by  $V_{\text{IIP3, AFE}}^2$  the total IIP3 voltage squared of the AFE chain. These are usually set by performance requirements dictated from digital baseband. Given  $F_{\text{AFE}}$  and  $V_{\text{IIP3, AFE}}^2$ , one would preferably like to select  $\bar{V}_N^2$  and  $V_{\text{IIP3}}^2$  of individual blocks such that the power consumption of the entire chain is minimized.

In order to solve this task, we first need to look into the nature of the relation between the performance quantifiers and power consumption for each block. The dynamic range of a block with index  $j$  is defined as

$$DR_j \triangleq \frac{V_{\text{IIP3},j}^2}{\bar{V}_{N,j}^2}. \quad (1)$$

As presented in [1] and [3], for a wide range of the most common front-end blocks, the power consumption of a circuit is linear with the dynamic range as defined in (1), i.e.

$$P_j = P_{C,j} DR_j, \quad (2)$$

where  $P_{C,j}$  is a proportionality factor that can be taken as a natural figure-of-merit for analog blocks.

Starting from this simple but powerful relation, the authors in [3] have devised a method of finding  $\bar{V}_{N,j}^2$  and  $V_{\text{IIP3},j}^2$  that results in minimum power consumption of the whole AFE chain. Although proof is given in [3] that relation (2) holds for standard CMOS circuits (such as a common-source stage LNA, a double-balanced Gilbert cell mixer and an OTA-C baseband filter), the results of the optimization are valid for any chain of analog blocks that satisfy (2) and are hence not limited only to CMOS circuits. Moreover, [3] provides a comparison of theoretically optimal  $\bar{V}_{N,j}^2$  and  $V_{\text{IIP3},j}^2$  with measured noise PSD and IIP3 from an actual ‘‘hand-optimized’’ Bluetooth receiver implementation, with a good match between the two. This hardware verification naturally extends to our analysis, which considers optimally designed front ends in communication system settings.

What is important for our analysis is that the method from [3] provides the connection between the optimal power consumption of the entire AFE, denoted

by  $P_{\text{AFE}}^*$ , and  $V_{\text{IIP3, AFE}}^2$  and  $F_{\text{AFE}}$ , which reads [3, eq. (60)]

$$P_{\text{AFE}}^* = \frac{V_{\text{IIP3, AFE}}^2}{(F_{\text{AFE}} - 1)kT50} \left( \sum_{j=1}^N \sqrt[3]{P_{\text{C},j}} \right)^3, \quad (3)$$

where  $k$  is Boltzmann constant and  $T$  temperature in Kelvins. Remarkably, the optimal power consumption of the chain is independent of power/voltage gains of individual blocks.

If we are to use the result in (3) for drawing conclusions on the system-level behaviour of receivers, it would be convenient to “translate” this result to system designer parlance, so that it features power-related parameters:

- received wanted signal power at the antenna -  $p_{\text{S}}$ ,
- total input-referred thermal noise power -  $p_{\text{N}}$ ,
- power of the out-of-band (OOB) interfering signal at the antenna -  $p_{\text{I}}$ .<sup>3</sup>

As a first step, we can relate  $p_{\text{N}}$  and  $F_{\text{AFE}}$  through

$$p_{\text{N}} = kTB F_{\text{AFE}}, \quad (4)$$

with  $B$  being the noise-equivalent bandwidth of the system. On the other hand, IIP3 power and voltage can be related by

$$p_{\text{IIP3, AFE}} = \frac{V_{\text{IIP3, AFE}}^2}{R_{\text{in}}}, \quad (5)$$

where  $R_{\text{in}}$  is the input resistance of the receiver which we assume to be  $50 \Omega$  for simplicity. In order to directly assess the impact of third-order nonlinearity on system performance, we need to relate the IIP3 to  $p_{\text{IM3}}$ , the power of the in-band third-order intermodulation (IM3) distortion. A well-known relation linking  $p_{\text{IIP3}}$ ,  $p_{\text{I}}$  and  $p_{\text{IM3}}$  reads [15]

$$p_{\text{IIP3}} = \sqrt{\frac{p_{\text{I}}^3}{p_{\text{IM3}}}}. \quad (6)$$

For the purpose of notational convenience, we denote the last term in (3) as

$$\kappa_{\text{circuit}} \triangleq \left( \sum_{j=1}^N \sqrt[3]{P_{\text{C},j}} \right)^3 \quad (7)$$

<sup>3</sup> The results (1) and consequently, (3) were derived with the assumption of a two-tone interference model. For the sake of consistency, we maintain this model throughout our analysis, and  $p_{\text{I}}$  then denotes the total power of the two interfering tones. However, we conjecture that the obtained trends hold even in the case of modulated interferers.

and use (4), (5) and (6) in conjunction with (3) to obtain

$$P_{\text{AFE}}^* = \frac{F_{\text{AFE}}}{F_{\text{AFE}} - 1} B \frac{p_{\text{I}}^{3/2}}{p_{\text{N}} \sqrt{p_{\text{IM3}}}} \kappa_{\text{circuit}}. \quad (8)$$

For the analysis at hand it is of use to define the power ratio of intermodulation distortion and noise

$$\alpha_{\text{IM3}} \triangleq \frac{p_{\text{IM3}}}{p_{\text{N}}}, \quad (9)$$

which combined with (8) yields

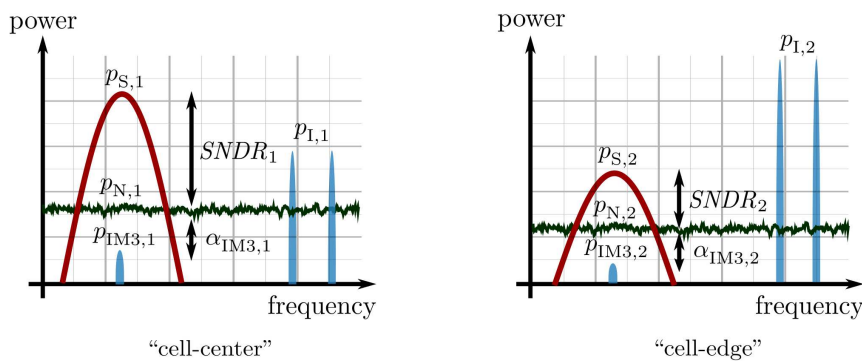
$$P_{\text{AFE}}^* = \frac{F_{\text{AFE}}}{F_{\text{AFE}} - 1} B \frac{1}{\sqrt{\alpha_{\text{IM3}}}} \left( \frac{p_{\text{I}}}{p_{\text{N}}} \right)^{3/2} \kappa_{\text{circuit}}, \quad (10)$$

with  $p_{\text{I}} > 0$  which follows from constraint  $V_{\text{IIP3, AFE}}^2 > 0$ . Equation (10) can be used as a basis for deriving simple but very useful scaling laws, as presented in the following section.

### 3 Scaling laws of AFE power consumption

A holistic receiver system design benefits greatly from the availability of closed form relations between receiver power consumption and other system parameters. This way, a mathematically tractable analysis of the tradeoffs encountered during receiver system design is made possible. When it comes to real-world hardware, obtaining such relations is not a trivial task, and there always exists a tradeoff between the accuracy of the functional dependencies and their analytical tractability. Ideally, they should appear in form of simple power laws. It turns out that (10), under some realistic assumptions, can yield such simple relations. The advantage of using (10) for this purpose is that it is soundly grounded in circuit theory which has also been verified against real-life receiver designs, so it enables striking a good balance between accuracy, simplicity and theoretical rigour.

To start with, a performance metric is needed that will provide a link between baseband metrics, like bit error rate (*BER*), via power-related system parameters, with circuit parameters  $F_{\text{AFE}}$  and  $V_{\text{IIP3, AFE}}^2$ . A commonly used



**Figure 1:** Illustration of all relevant system parameters for the cell center scenario (left: strong wanted signal, weak OOB interference, high  $SNDR$  requirement) and cell edge scenario (right: weak wanted signal, strong OOB interference, low  $SNDR$  requirement).

such metric is the signal-to-noise-and-distortion ratio<sup>4</sup>, which is defined as

$$SNDR \triangleq \frac{p_S}{p_N + p_{IM3}} = \frac{p_S}{(1 + \alpha_{IM3}) p_N}. \quad (11)$$

Now we focus our attention on four fundamental receiver design parameters, namely,  $SNDR$  and  $B$  (the values of which are determined by the particular application), and  $p_S$  and  $p_I$  (which describe the environment and are generally stochastic). The values of the fundamental parameters define distinct application-environment scenarios. We structure our analysis around a pair of such scenarios: an initial (pre-scaling) and target (post-scaling) scenario. An illustration of the relations between parameters of importance for an example scenario pair is given in Fig. 1. For each of the two scenarios—under practical constraints on parameter values—we assume that an analog front end with minimal power consumption is designed using the procedure described in [3]. Our aim is relating the scaling of fundamental parameter values between the two scenarios and the scaling of optimal front end power. To this end, we label variables corresponding to pre-scaling and post-scaling scenarios with indices 1

<sup>4</sup> It is commonly assumed that the third-order distortion is the dominant nonlinear impairment in analog systems. Therefore, along with thermal noise, we consider it a determining factor of system performance. All other possible impairments, such as second-order distortion, flicker noise, phase noise—either in-band or due to reciprocal mixing—are through appropriate design assumed to be dominated by thermal noise and third-order distortion in all scenarios considered.

and 2, respectively. The scaling of the optimal power consumption is denoted as

$$\varsigma_P \triangleq \frac{P_{\text{AFE},2}^*}{P_{\text{AFE},1}^*}. \quad (12)$$

The scaling factors of bandwidth,  $SNDR$ , signal and interference power are defined analogously to  $\varsigma_P$  and denoted respectively as  $\varsigma_B$ ,  $\varsigma_{SNDR}$ ,  $\varsigma_S$  and  $\varsigma_I$ . By using (10) and (11), the scaling of front end power reads

$$\varsigma_P = \varphi \delta \varsigma_B \varsigma_{SNDR}^{3/2} \varsigma_I^{3/2} \varsigma_S^{-3/2}, \quad (13)$$

where, for analytical convenience, we have introduced the factors

$$\varphi \triangleq \frac{F_{\text{AFE},2} F_{\text{AFE},1} - 1}{F_{\text{AFE},1} F_{\text{AFE},2} - 1} \quad (14)$$

and

$$\delta \triangleq \sqrt{\frac{\alpha_{\text{IM3},1}}{\alpha_{\text{IM3},2}}} \left( \frac{1 + \alpha_{\text{IM3},2}}{1 + \alpha_{\text{IM3},1}} \right)^{3/2}. \quad (15)$$

Expression (13) is a universal tool for calculating front end power scaling and can be used for all application-environment scenarios, under the condition that the corresponding front ends are implementable. However, one does need to use (13) in a careful and structured way due to interdependencies between the fundamental parameters ( $SNDR$ ,  $B$ ,  $p_S$ ,  $p_I$ ) and noise-distortion ratio  $\alpha_{\text{IM3}}$ , system-level design parameters ( $p_N$ ,  $p_{\text{IM3}}$ ,  $p_{\text{IIP3}}$ ) and circuit-level parameters ( $F_{\text{AFE}}$ ,  $V_{\text{IIP3}}^2$ ). More specifically, for a particular scenario,  $SNDR$ ,  $B$ ,  $p_S$ ,  $p_I$  and  $\alpha_{\text{IM3}}$  will through (6), (8) and (11) yield  $p_N$ ,  $p_{\text{IM3}}$  and  $p_{\text{IIP3}}$ , which through (4) and (5) result in  $F_{\text{AFE}}$  and  $V_{\text{IIP3}}^2$ . Combining  $F_{\text{AFE}}$  from pre- and post-scaling scenarios yields  $\varphi$  from (14), noise-distortion ratios  $\alpha_{\text{IM3}}$  give the value of  $\delta$  from (15), and the values of fundamental parameters result in respective scaling ratios, all of which is combined in (13) for the final result.

In order to isolate the scaling of power with only one of the fundamental parameters, we assume that the value of the parameter in question scales between the scenarios while other parameters remain constant. In this way, we obtain a restricted set of application-environment scenarios with high practical relevance, examined in detail in Section 4. Additionally, in all scenarios it is assumed that pre- and post-scaling  $\alpha_{\text{IM3}}$  values are the same, i.e. that input-referred thermal noise and third-order distortion levels are kept at a constant ratio. Available literature on systematic receiver design suggests that in practice, the value of  $\alpha_{\text{IM3}}$  is chosen to be small (typically on the order of 0.1) so that the third-order distortion is much weaker than the thermal noise, with the

choice being consistent over different application-performance scenarios [15, Ch. 13]. This consistency over scenarios is in line with our constant- $\alpha_{\text{TM3}}$  assumption.

The laws describing the scaling of front end power with fundamental parameters are given in Table 1, expr. (16)-(19). Each of the four rows of the table, corresponding to a particular scaling law, also provides a comprehensive list of application-environment constraints (columns 1-4), together with a list of resulting system/circuit design requirements (columns 4-8) needed for scaling laws to hold in practical implementations, obtained through (4)-(6), as discussed above. Note that the constraints on the scaling of  $B$  also double as explicit design requirements.



**Table 1:** Collection of fundamental scaling laws for front end power. Application and environment constraints (columns 1-4) are translated to front end design requirements (columns 4-8). Consequently, if the front ends are designed optimally, their power scales with a selected parameter (bandwidth,  $SNDR$ , received power, blocker power) as given in columns 9 and 10.

Constraints		Design requirements						Power scaling	Properties of $\varphi$
Performance	Environment	System			Circuit				
$SNDR_2$	$p_{S,2}$	$p_{1,2}$	$B_2$	$p_{N,2}$	$p_{IM3,2}$	$F_2$	$V_{IFP3,2}^2$	$\varsigma_P = \varphi \varsigma_B$ (16)	$\varphi < 1, 0 < \varsigma_B < 1$ $\varphi \geq 1, 1 \leq \varsigma_B < F_1$
$SNDR_1$	$p_{S,1}$	$p_{1,1}$	$\varsigma_B B_1$	$p_{N,1}$	$p_{IM3,1}$	$\frac{F_1}{\varsigma_B}$	$V_{IFP3,1}^2$	$\varsigma_P = \varphi \varsigma_B$ (16)	$\varphi < 1, 0 < \varsigma_{SNDR} < 1$ $\varphi \geq 1, 1 \leq \varsigma_{SNDR} < F_1$
$\varsigma_{SNDR} SNDR_1$	$p_{S,1}$	$p_{1,1}$	$B_1$	$\frac{p_{N,1}}{\varsigma_{SNDR}}$	$\frac{p_{IM3,1}}{\varsigma_{SNDR}}$	$\frac{F_1}{\varsigma_{SNDR}}$	$\sqrt{\varsigma_{SNDR}} V_{IFP3,1}^2$	$\varsigma_P = \varphi \varsigma_{SNDR}^{3/2}$ (17)	$\varphi < 1, 0 < \varsigma_{SNDR} < 1$ $\varphi \geq 1, 1 \leq \varsigma_{SNDR} < F_1$
$SNDR_1$	$\varsigma_S p_{S,1}$	$p_{1,1}$	$B_1$	$\varsigma_S p_{N,1}$	$\varsigma_S p_{IM3,1}$	$\varsigma_S F_1$	$\frac{V_{IFP3,1}^2}{\sqrt{\varsigma_S}}$	$\varsigma_P = \varphi \varsigma_S^{-3/2}$ (18)	$\varphi < 1, \varsigma_S > 1$ $\varphi \geq 1, \frac{1}{F_1} < \varsigma_S \leq 1$
$SNDR_1$	$p_{S,1}$	$\varsigma_1 p_{1,1}$	$B_1$	$p_{N,1}$	$p_{IM3,1}$	$F_1$	$\varsigma_1^{3/2} V_{IFP3,1}^2$	$\varsigma_P = \varsigma_1^{3/2}$ (19)	$\varphi = 1$

The four scaling laws can be stated in dB domain in form of convenient rules of thumb, as follows:

1. (*Power consumption-bandwidth scaling law*): For every 1 dB increase/decrease of system bandwidth, the power consumption of an optimally designed analog front end increases/decreases by at least 1 dB.
  - It is well known that the power consumption of standard analog blocks scales linearly with bandwidth [2]. This scaling law demonstrates that the linear power-bandwidth relation extends also to a chain of analog blocks.
2. (*Power consumption-SNDR scaling law*): For every 1 dB increase/decrease of *SNDR*, the power consumption of an optimally designed analog front end increases/decreases by at least 1.5 dB.
  - This novel scaling law serves as a fundamental relation for analyzing power-performance tradeoffs in analog receiver design, as analyzed more in-depth in Sections 4.2 and 4.3.
3. (*Power consumption-received power scaling law*): For every 1 dB increase/decrease of received wanted signal power, the power consumption of an optimally designed analog front end decreases/increases by at least 1.5 dB.
  - This relation will be useful in analyzing power savings of a front end that adapts to a fluctuating received signal level while maintaining constant performance, as will be presented in Section 4.4.
4. (*Power consumption-interference scaling law*): For every 1 dB increase/decrease of the out-of-band interference power, the power consumption of an optimally designed analog front end increases/decreases by 1.5 dB.
  - By defining the signal-to-interference ratio  $SIR = p_S/p_I$ , this scaling can be reformulated as  $\varsigma_P = \varsigma_{SIR}^{-3/2}$ , where  $\varsigma_{SIR}$  is the scaling of the *SIR*. An identical scaling law was presented in [5], where  $P_{AFE}$  was optimized for energy efficiency. Scaling law (19) is of importance when analyzing the power consumption of a front end that dynamically adapts its linearity to the interference level while maintaining constant performance. A detailed theoretical analysis of such a front end will be given in Section 4.5.

Laws (16) - (18) are characterized by the fact that the underlying scaling asks for tuning of the noise figure, which in turn makes the parameter  $\varphi$  scaling-dependent. More specifically, for  $\varsigma_*$ , where  $*$   $\in$  (B, SNDR),

$$\varphi = \frac{F_{\text{AFE},1} - 1}{F_{\text{AFE},1} - \varsigma_*}, \quad (20)$$

whereas for  $\varsigma_S$  we have

$$\varphi = \frac{F_{\text{AFE},1} - 1}{F_{\text{AFE},1} - \frac{1}{\varsigma_S}}. \quad (21)$$

The dependence of  $\varphi$  on the scaling parameters is outlined in the last column of Table 1. The constraint  $\varphi > 0$ , i.e. the fact that it is physically impossible to have a front end with  $F < 1$  imposes theoretical limitations on the values of scaling  $\varsigma$ . Furthermore, dependence of  $\varphi$  on  $\varsigma$  causes deviations from the ideal scaling of power (linear with bandwidth or following the 3/2 power law in case of *SNDR* and received power). In order to have proper scaling laws, it is necessary for  $\varphi$  to be independent of  $\varsigma$ . This condition is approximately satisfied in two cases:

- $F_{\text{AFE},1} \gg 1 \Rightarrow \varphi \approx 1$ ;
- $\varsigma_B, \varsigma_{\text{SNDR}} \ll 1$  or  $\varsigma_S \gg 1 \Rightarrow \varphi \approx \frac{F_{\text{AFE},1} - 1}{F_{\text{AFE},1}}$ .

At first, it can appear that the set of scenarios in which power scaling laws (16)-(18) are close to ideal ( $\varphi \approx 1$ ) is based on such a restrictive sequence of assumptions that their practical relevance is questionable. However, a closer look reveals that all the assumptions we used are commonplace in practice and/or of high practical interest for low-power design. To start with,  $F_{\text{AFE},1} \gg 1$  is typical for worst-case front end designs with a large OOB blocking signal present [15, Ch. 13], [16]. Moreover, radical scaling down of system bandwidth (e.g. going from a wideband to a narrowband system), drastic downscaling of *SNDR* requirement (due to e.g. use of power-efficient transmission techniques) or adaptation to wanted signal power that becomes much larger than worst-case (reference sensitivity) due to fading fluctuations are all use-cases of interest for low-power applications [17], [18].

## 4 Ramifications of the scaling laws

The scaling laws presented in the previous section constitute a set of tools which prove to be very useful in the design of receivers where power consumption is of high importance. Namely, as the laws in the preceding section formally show,

the power consumption of the analog front end can be lowered by using one (or more) of the following techniques:

- Intentionally degrading the bit/symbol error rate ( $SE_R$ ), which consequently reduces the  $SNDR$  requirement;
- Keeping  $BER$  or  $SE_R$  constant while applying some transmission technique that allows for lower  $SNDR$  (e.g. use of error control coding);
- Keeping the  $SNDR$  constant while making the AFE reconfigurable so that it adapts to the changes in the environment (e.g. fading level fluctuations, OOB interference level).

The scaling laws serve as a basis for estimates of the extent of power savings that can be achieved in the AFE if the aforementioned techniques are applied. System designers can then decide on which techniques to incorporate in their systems, and hardware designers are provided with general guidelines on how to increase the power efficiency of circuit designs.

#### 4.1 Preliminaries: limitations on hardware relaxation

Throughout the analysis that follows, we consider analog front ends designed for different target values of noise and distortion. When it comes to realistic hardware designs, however, it is reasonable to assume that the range of these values is limited. Naturally, there are fundamental physical constraints on the minimum noise (or distortion) level that a circuit can deliver, but, equally important, there are also upper bounds, imposed by either functionality or technology process constraints [2]. Hence, in line with considerations from the previous section, we establish a permissible tuning range  $\mu$  that applies to both noise figure and IP3. It is defined as the value of scaling of noise/linearity for which, given all architectural and physical limitations, the following holds:

- The noise figure  $F_{AFE}$  can be degraded from the reference value  $F_{AFE,1}$  to a maximum value of  $F_{AFE,2} = \mu F_{AFE,1}$ ,
- It is possible to degrade IP3 from the reference value  $V_{IP3,1}^2$  to a minimum possible value of  $V_{IP3,2}^2 = \frac{1}{\sqrt{\mu}} V_{IP3,1}^2$ .

#### 4.2 Power- and energy-efficient AFEs through intentional degradation of performance, uncoded case

In this section, we focus on systems using M-QAM without any error control coding. With the aim of saving power, System 2 either uses a lower QAM

constellation order  $M$  or operates at a higher symbol error probability  $P_e$ , formally,  $M_2 \leq M_1$  or  $P_{e,2} \geq P_{e,1}$ . As indicated in Section 3, the two systems are otherwise assumed to use the same bandwidth (and thus the same symbol rate  $R_s$ ), are affected by same OOB interference level and experience the same wanted signal power.

We assume that the classical matched-filter detector is employed at the receiver. If the thermal Gaussian noise dominates the IM3, i.e.  $\alpha_{\text{IM3}} \ll 1$ , the matched-filter receiver is optimal in the sense of maximum a posteriori detection. Under these circumstances, an upper bound on  $SEr$  for a square M-QAM ( $M = 2^{2k}$ ,  $k \in \mathbb{N}$ ) can be determined [19], which yields the inequality

$$SNDR \leq \rho \frac{M-1}{3 \log_2 M} \left[ Q^{-1} \left( \frac{P_e}{4} \right) \right]^2, \quad (22)$$

where  $\rho = R_b/B$  is the spectral efficiency of the uncoded system ( $R_b$  is the information bitrate) and  $Q^{-1}(\cdot)$  the inverse of the upper tail probability function of a unit-variance Gaussian random variable. At high  $SNDR$ s, the upper bound in (22) is tight.

We proceed by constructing a ratio of the upper bounds from (22) that apply to the two distinct scenarios under analysis. This ratio is given as

$$\varsigma_{\text{SNDR}} \geq \frac{M_2-1}{M_1-1} \left[ \frac{Q^{-1}(P_{e,2}/4)}{Q^{-1}(P_{e,1}/4)} \right]^2, \quad (23)$$

where the fact that  $B$  is the same for the two systems is used. Taking into account the practical limits on noise/linearity scaling, discussed in Section 4.1, along with law (17), the achievable scaling of front end power,  $\varsigma_{\text{P, a}}$ , is found to be

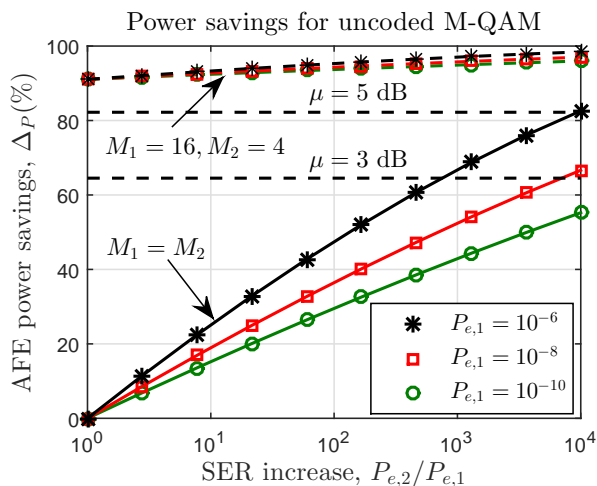
$$\varsigma_{\text{P, a}} < \max\{\varsigma_{\text{SNDR}}^{3/2}, \mu^{-3/2}\}. \quad (24)$$

By combining this together with (23) and the fact that the slack of the  $SEr$  upper bound increases with decreasing  $SNDR$ , we obtain the upper bound on the achievable AFE power downscaling:

$$\varsigma_{\text{P, a}} \leq \max \left\{ \left( \frac{M_2-1}{M_1-1} \right)^{3/2} \left[ \frac{Q^{-1}(P_{e,2}/4)}{Q^{-1}(P_{e,1}/4)} \right]^3, \mu^{-3/2} \right\}. \quad (25)$$

In other words, the AFE power can be decreased by *at least* the value of the right hand side of (25). For large  $SNDR$ s and large  $F_{\text{AFE},1}$ , the bound is tight.

The obtained bound enables the derivation of laws describing the performance-power consumption tradeoff in systems using uncoded QAM when there are no



**Figure 2:** Savings in AFE power consumption when symbol error probability and/or constellation order are degraded, for uncoded square QAM. For limited flexibility AFEs, the savings cap at values indicated by horizontal dashed lines.

limits on  $SNDR$  tuning,  $\mu \rightarrow \infty$ . In one case, we keep  $P_e$  constant but reduce the number of bits per symbol  $b = \log_2 M$  by  $\Delta_b = b_1 - b_2$ . This yields

$$\varsigma_P < \left( \frac{M_2}{M_1} \right)^{3/2} = 2^{-\frac{3}{2}\Delta_b}. \quad (26)$$

Therefore, the power consumption of an infinitely flexible AFE decreases at least exponentially with the difference in bits/symbol, or equivalently, with the difference in raw uncoded bitrate. In another setting, we assume  $M$  is the same between the two systems but target  $SE$ R is increased from  $P_{e,1}$  to  $P_{e,2}$ . Using the bound  $Q(x) \leq e^{-x^2/2}$ , we get

$$\varsigma_P \leq \left[ \frac{Q^{-1}(P_{e,2}/4)}{Q^{-1}(P_{e,1}/4)} \right]^3 \leq \left( \frac{1 - 1.66 \log_{10} P_{e,2}}{1 - 1.66 \log_{10} P_{e,1}} \right)^{3/2}. \quad (27)$$

Assuming additionally that the order of magnitude  $\omega_e = \log_{10} P_e$  of  $SE$ R is low enough, we get

$$\varsigma_P \leq \left( \frac{\omega_{e,2}}{\omega_{e,1}} \right)^{3/2}. \quad (28)$$

In other words, we can say that the power consumption of the AFE with infinite flexibility scales at least as  $O(\omega^{3/2})$ .

For convenience of presenting numerical results, we define the percentage savings of AFE power

$$\Delta_P \triangleq 100(1 - \varsigma_P) \quad [\%]. \quad (29)$$

These savings, represented in Fig. 2 imply that, if presented with a choice of whether to sacrifice bitrate or error rate in order to save power in the receiver, we should in general opt for the former. Taking into account hardware design limitations, substantial savings are achievable even when it is possible to scale down the *SNDR* by as little as e.g. 3 dB; naturally, in order to harvest the full potential of the savings, the AFE should be made as flexible as hardware constraints permit.

In order to provide a completely fair comparison between the systems, degradation of the performance and reduction of power consumption should be considered jointly. A joint metric for performance and power consumption is needed for this task, and one is readily found in the form of energy efficiency

$$\eta_{\text{AFE}} \triangleq \frac{R_b}{P_{\text{AFE}}} \quad [\text{bits/J}]. \quad (30)$$

In the case when constellation size  $M$  changes but error rate  $P_e$  stays fixed and with unlimited flexibility, the ratio of the two efficiencies is

$$\frac{\eta_{\text{AFE},2}}{\eta_{\text{AFE},1}} = \frac{1}{\varsigma_P} \frac{R_{b,2}}{R_{b,1}} \geq \left( \frac{M_1 - 1}{M_2 - 1} \right)^{3/2} \frac{\log_2 M_2}{\log_2 M_1}. \quad (31)$$

From here we can conclude that, for a fixed  $P_e$ ,  $\eta_{\text{AFE}}$  will always improve if the size of the square QAM constellation is reduced. As a quick proof, we consider the fact that for square QAM,  $M = 2^{2k}$ ,  $k \in \mathbb{N}$  and so for any  $k > 1$  we have  $2^{2k} - 1 > 1$ . This also means that for any  $k_1 > k_2$ ,  $k_1, k_2 \in \mathbb{N}$  it will hold that

$$\left( \frac{2^{2k_1} - 1}{2^{2k_2} - 1} \right)^{3/2} \frac{k_2}{k_1} > 1. \quad (32)$$

But the left hand side of (32) is equivalent to the right hand side of (31), which means that

$$\frac{\eta_{\text{AFE},2}}{\eta_{\text{AFE},1}} \geq 1 \quad (33)$$

for  $M_2 < M_1$ .

Therefore, the smaller the QAM constellation, the more energy efficient the AFE of the receiver. We note that, in the case when  $\eta$  is defined with respect

to *transmit signal power*, it is a well known fact that the energy efficiency increases with decreasing QAM constellation size [19]. With (33), however, we prove that this energy efficiency property of QAM constellations extends to the case of *power consumption of analog receiver hardware*.

### 4.3 Power- and energy-efficient AFEs through use of error control coding

Error control coding (ECC) techniques are used to improve reliability (error rate performance) of communication systems when  $SNDR$  is kept fixed. Seen from another angle, when the error rate is constrained to be the same for uncoded and coded systems, coding can be used to improve the power efficiency of communication systems as a consequence of relaxed requirements on  $SNDR$ . Here we analyze the case when this potential for increased power efficiency is used by the receiver (it can also be used by the transmitter, or be distributed between the two).

Power efficiency gain of coded systems is usually expressed in terms of the coding gain  $g_c$ . By assuming that  $\alpha_{\text{IM3}} \ll 1$ , we can approximate the PSD of the sum of all impairments by additive white Gaussian noise PSD  $N_0$  and define the ratio  $E_b/N_0$  of energy per bit  $E_b$  and  $N_0$ . Given the  $E_b/N_0$  values required to achieve the same error probability with and without coding, the coding gain is defined as

$$g_c \triangleq \frac{(E_b/N_0)_{\text{uncoded}}}{(E_b/N_0)_{\text{coded}}}. \quad (34)$$

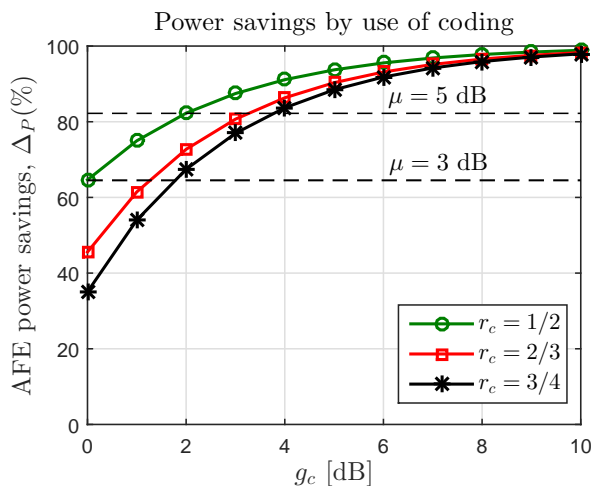
For finding the achievable AFE power reduction, we need to connect the coding gain  $g_c$  with the  $SNDR$  downscaling  $\varsigma_{\text{SNDR}}$ , where  $SNDR_1$  corresponds to the uncoded system and  $SNDR_2$  to the coded one. We do this by assuming that the system bandwidth is equal for both systems, which is a reasonable assumption for all applications where bandwidth is a limited resource. Consequently, using ECC will reduce spectral efficiency from  $\rho_{\text{uncoded}}$  to  $\rho_{\text{coded}} = r_c \rho_{\text{uncoded}}$ , where  $r_c$  is the coding rate. We additionally use the fact that  $E_b/N_0 = SNDR/\rho$  to obtain

$$\varsigma_{\text{SNDR}} = \frac{r_c}{g_c}, \quad (35)$$

and the associated achievable AFE power reduction (cf. (24)) is then given by

$$\varsigma_{\text{P, a}} < \max \left\{ \left( \frac{r_c}{g_c} \right)^{3/2}, \mu^{-3/2} \right\}. \quad (36)$$





**Figure 3:** Savings in AFE power consumption coming from use of error control coding. For limited flexibility AFEs, the savings cap at values indicated by horizontal dashed lines.

The savings function (29) for systems using coding is illustrated in Fig. 3. An important observation to make here is that a large portion of the power savings (in absolute power terms) is harvested by low to intermediate coding gains. Additional absolute power savings that are brought about by employing stronger codes with larger coding gains are only marginal. This point is further elaborated in the follow-up.

### Numerical example

Here we provide a system design scenario that serves to illustrate the potential savings of AFE power consumption when ECC is used, and also to give some system-level design guidelines. We assume a system with passband bandwidth of 40 MHz, BPSK modulation and single carrier transmission using raised cosine pulses with rolloff of 0.5 over a flat-faded channel. Total receiver power (AFE + decoder) is calculated for three versions of the system: one uncoded and two with different types of ECC. If coding is used, the AFE design is relaxed accordingly. Power consumption values used here are ballpark quantities based on actual hardware designs. For the decoders, the power numbers obtained from the designs are modified to match the information bitrate (assuming that a linear extrapolation of decoder power consumption is possible at lower bitrates) and scaled to the same process (65 nm CMOS) and voltage (1.2 V).

System parameters and calculated power numbers are listed out in Table 2. The use of coding allows for relaxation of the AFE by making it noisier and less linear, so its power consumption is ideally reduced as per (36). However, the overhead in power consumption stemming from the channel decoders also needs to be taken into account in order for the full story to be told. It can be seen that in the case of the system using convolutional codes (CC), a massive reduction of AFE power comes with a relatively small power overhead for the decoding. Using turbo codes allows for further reductions of AFE power compared to the CC case, but at a cost of a relatively high decoding power overhead, which is due to the iterative nature of the turbo decoder. Dividing the information bitrate with total power consumption yields the energy efficiency of the receiver, which indicates that coding indeed enables an improvement of the receiver energy efficiency, but the best strategy is to use “light” codes, with moderate coding gains and simpler decoders.

We note that the relation between error control coding and overall energy efficiency of the system is a long-standing research topic, examined both empirically and theoretically in, e.g., [24] and [25]. However, these papers analyze the combination of decoding power and *transmit power*, whereas we focus on the *total power of the receiver*, that is, the sum of decoding power and power consumed by supporting analog hardware. Here we have only touched upon this topic of high practical relevance, and a more thorough analysis is left for future work as it is out of the scope of this paper.

As for the energy efficiency of the AFE alone, it can be quickly shown that it always improves with coding. This is done by setting up the ratio of energy efficiencies (30) for the coded and uncoded system, which gives

$$\frac{\eta_{\text{coded}}}{\eta_{\text{uncoded}}} = \frac{1}{\varsigma_{\text{P}}} \frac{R_{\text{b,coded}}}{R_{\text{b,uncoded}}} = \frac{r_{\text{c}}}{\varsigma_{\text{P}}} > \frac{g_{\text{c}}^{3/2}}{\sqrt{\varsigma_{\text{P}}}} \quad (37)$$

in the case of infinite AFE flexibility. But the obtained ratio is always  $> 1$  for  $g_{\text{c}} > 1$  (for a properly designed code operating at a large enough *SNDR*).

Overall, the results in this section lead to the conclusion that low power applications that harness error control coding gains for the goal of relaxing the receiver favor simple codes with modest coding gains and simpler decoders over more powerful codes that ask for more involved decoding algorithms. Another, more general design guideline is that the power budget for the channel decoder must fit into the margin opened up by relaxing the AFE if the goal is to reduce the overall receiver power consumption. If we, on the other hand, consider solely the AFE, it can be shown that coding always improves its energy efficiency.

**Table 2:** System parameters and theoretical power numbers for AFEs and decoders in systems using error control coding

Inf. bitrate [Mbps]	Code	$r_c$	$g_c$ @ $BER = 10^{-3}$ [dB]	$P_{\text{AFE}}$ [mW]	$P_{\text{dec}}$ [mW]	$P_{\text{AFE}} + P_{\text{dec}}$ [mW]	Total energy efficiency [Gbits/J]
26.7	uncoded	-	-	35 (ref. [21])	0	35	0.76
13.35	convolutional (7, 5)	1/2	3.1 (ref. [22])	4.26	0.56 (ref. [22])	4.82	2.77
8.89	turbo N = 6144	1/3	6.1 (ref. [23])	0.82	8.3 (ref. [23])	9.12	0.96

#### 4.4 Power-efficient AFEs through adaptation to fading

In this section, we assume a single carrier transmission over a frequency flat wireless channel. Due to fading, received power  $p_S$  will be time varying and can be well described as a random process

$$p_S(t) = \beta\phi(t), \quad (38)$$

where  $\beta$  subsumes the transmit power, transmit and receive antenna gains, pathloss and large-scale fading, which are all assumed constant in this context. Additionally,  $\phi(t) = |h(t)|^2$ , where  $h(t)$  is a zero-mean unit-variance complex Gaussian random process, i.e. the small-scale fading adheres to the common Rayleigh fading model. It is well known that  $\phi(t)$  has an exponential pdf [20]

$$f_\Phi(\phi) = e^{-\phi}, \quad \phi \geq 0. \quad (39)$$

A common design parameter for wireless systems is the outage probability  $\Omega$ , defined as the probability that the normalized fading power  $\phi$  falls below some minimum acceptable level  $\phi_{\min}$  [20],

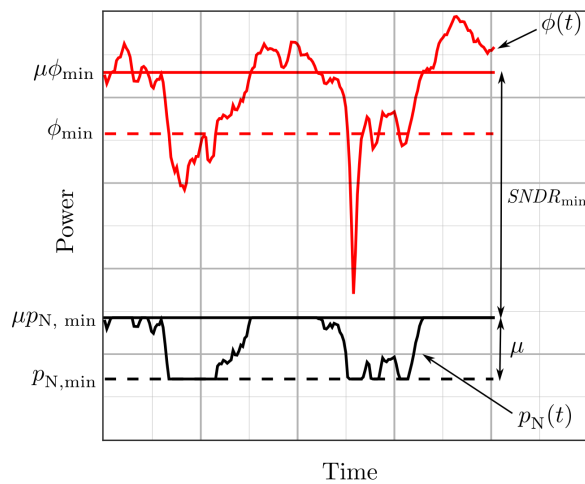
$$\Omega \triangleq \int_0^{\phi_{\min}} f_\Phi(\phi) d\phi. \quad (40)$$

In conjunction with  $\phi_{\min}$ , an outage  $SNDR$  is usually defined, which represents the minimum  $SNDR$  that provides acceptable performance. Using  $\phi_{\min}$  and

$SNDR_{\min}$ , a minimum (worst-case) thermal noise level is calculated as

$$p_{N, \min} = \frac{\beta \phi_{\min}}{(1 + \alpha_{\text{IM3}}) SNDR_{\min}}. \quad (41)$$

Therefore, a minimum noise level  $p_{N, \min}$  and a minimum third-order distortion  $p_{\text{IM3}, \min}$  need to be delivered by the AFE at least at the time instants where  $\phi(t) = \phi_{\min}$ . For all practical purposes, however, AFEs are built so that they deliver minimum noise and distortion *all the time*. Since the outage probability  $\Omega$  is typically chosen to be quite low (for example, on the order of  $10^{-2}$ ), this means that for the vast majority of time,  $SNDR$  delivered by these worst-case designs will be much larger than  $SNDR_{\min}$  and performance far better than the minimum acceptable one.



**Figure 4:** Illustration of time-varying fading and system parameters for the fading-adaptive front end design

Unless the variations in  $SNDR$  are leveraged for increasing throughput (via adaptive modulation and coding), having the front end operate in a fixed manner represents a waste of power. If a fixed throughput and error rate are acceptable for a particular application, front end noise and linearity can be tuned to track the variations of received power and maintain constant  $SNDR$  (effectively “equalizing” the channel). As indicated by results of Section 3, such an approach would result in a reduction of power consumed by the front end.

**Table 3:** Noise tuning parameters and normalized power consumption of fading-adaptive front ends with limited adaptation range

Fading level	$p_N(t)$	$\frac{P_{\text{AFE}}(t)}{P_{\text{AFE, wc}}}$	Remark
$\phi(t) \leq \phi_{\min}$	$p_{N, \min}$	$= 1$	Outage
$\phi_{\min} < \phi(t) \leq \mu\phi_{\min}$	$\frac{\beta\phi(t)}{(1+\alpha_{\text{IM3}})SNDR_{\min}}$	$< \left[\frac{\phi(t)}{\phi_{\min}}\right]^{-3/2}$	$SNDR = SNDR_{\min}$
$\phi(t) > \mu\phi_{\min}$	$\mu p_{N, \min}$	$< \mu^{-3/2}$	$SNDR > SNDR_{\min}$

We now turn to quantifying this reduction. Firstly, in line with considerations in Section 4.1, it is reasonable to assume that the noise level in an adaptive front end can be tuned only in a limited range ( $p_{N, \min}, \mu p_{N, \min}$ ) while being kept constant at the range boundaries for too small/large values of  $\phi(t)$ . The same logic extends to adapting the distortion level by means of tuning the nonlinearity, which yields the allowed range for the distortion of ( $p_{\text{IM3}, \min}, \mu p_{\text{IM3}, \min}$ ). The adaptation rule for thermal noise in a fading-adaptive front end with limited adaptation range is given in Table 3, with the most important parameters of interest illustrated in Fig. 4. Using relations (4) - (6),  $\alpha_{\text{IM3},1} = \alpha_{\text{IM3},2}$  and the set of constraints from the third row of Table 1, these rules can be easily translated to feature circuit design parameters.

We further denote by  $P_{\text{AFE, wc}}$  the power consumption of the non-adaptive, worst-case front end architecture, designed to deliver  $p_{N, \min}$  and  $p_{\text{IM3}, \min}$  throughout. Taking into account scaling law (18), the power consumption of the adaptive front end  $P_{\text{AFE}}(t)$  normalized by  $P_{\text{AFE, wc}}$  depends on  $\phi(t)$  and is given in Table 3. From there, the expected value of power scaling  $\varsigma_P$  for the adaptive front end can be calculated by assuming that  $\phi(t)$  is an ergodic process (so time averages can be substituted by ensemble averages) as

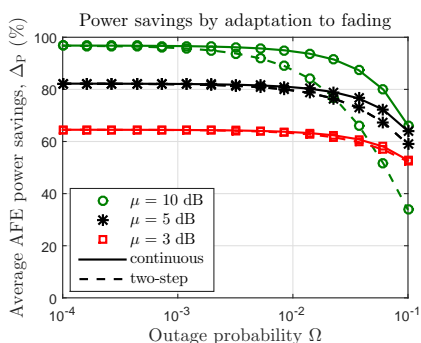
$$\mathbb{E}\{\varsigma_P\} \leq \int_0^{\phi_{\min}} e^{-\phi} d\phi + \phi_{\min}^{3/2} \int_{\phi_{\min}}^{\mu\phi_{\min}} \phi^{-3/2} e^{-\phi} d\phi + \mu^{-3/2} \int_{\mu\phi_{\min}}^{\infty} e^{-\phi} d\phi, \quad (42)$$

which yields

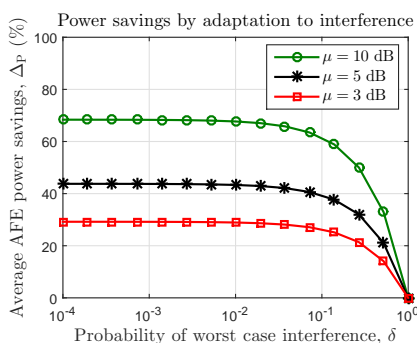
$$\mathbb{E}\{\varsigma_P\}_{\text{continuous}} \leq 1 - e^{-\phi_{\min}} + \quad (43)$$

$$2 \left\{ \phi_{\min} e^{-\phi_{\min}} \left( 1 - \frac{1}{\sqrt{\mu}} e^{(1-\mu)\phi_{\min}} \right) + \phi_{\min}^{3/2} \left[ \Gamma\left(\frac{1}{2}, \mu\phi_{\min}\right) - \Gamma\left(\frac{1}{2}, \phi_{\min}\right) \right] \right\} + \mu^{-3/2} e^{-\mu\phi_{\min}},$$

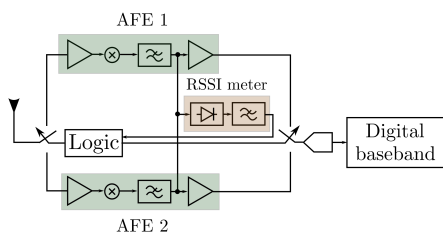
where  $\Gamma(a, x)$  denotes the upper incomplete gamma function [26].



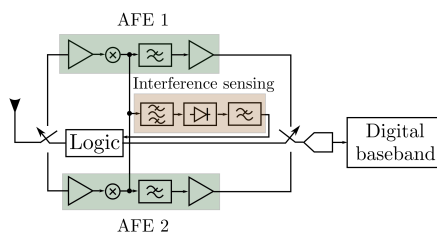
a) Average power savings of continuous and two-step fading-adaptive AFEs



b) Average power savings of two-step interference-adaptive AFEs



c) Example architecture for two-step adaptation to fading. Front end 1 has low noise figure and high IP3; front end 2 high noise figure and low IP3.



d) Example architecture for two-step adaptation to interference. Front end 1 has high IP3; front end 2 low IP3.

**Figure 5:** Theoretical power savings and conceptual illustrations of architectures for adaptive receivers

Achieving continuous tuning of noise and linearity can be challenging in practical implementations. Apart from adapting to the environment, the issue of random PVT (process, voltage, temperature) variations also needs to be accounted for. There exist solutions for jointly solving these practical problems, such as the one presented in [10], where LNAs with orthogonally tunable noise and linearity are combined with a simple online optimization algorithm, yielding substantial power savings. An alternative way of tackling this issue is to form a bank of front ends that are optimally designed for different noise and linearity settings. During operation, the receiver would switch between different front ends based on the measured received power, keeping one front

end active and switching off the rest. In the most basic case, such a bank would consist of only two front ends. A switching rule for this two-step adaptive front end that guarantees  $SNDR \geq SNDR_{\min}$  can be defined as

$$p_N = \begin{cases} p_{N, \min}, & \phi(t) \leq \mu\phi_{\min}, \\ \mu p_{N, \min}, & \phi(t) > \mu\phi_{\min}. \end{cases} \quad (44)$$

Average power downscaling for the two-step front end is found to be

$$\mathbb{E} \{ \zeta_P \}_{\text{two-step}} \leq 1 - \left( 1 - \mu^{-3/2} \right) e^{-\mu\phi_{\min}}. \quad (45)$$

Average power scaling for flexible and two-step front ends is converted to average savings as per (29) and shown in Fig. 5a). When the tuning range  $\mu$  is small, normalized signal power  $\phi(t)$  is either in outage or above  $\mu\phi(t)$  for most of the time, so continuous and two-step front ends have similar power savings. As the tuning range increases, more power can be saved, but in the case of large outage probability,  $\phi(t)$  is rarely larger than  $\mu\phi(t)$ . This means that in the case of the two-step front end, the noisy, nonlinear, low power front end rarely gets activated and the power savings are significantly lower compared to continuous adaptation. In any case, the obtained savings are substantial<sup>5</sup>, which should serve as a motivation for implementing fading-adaptive front ends in practice. In the case of two-step adaptation, such implementations can have an appealing simplicity. As means of illustration, we provide a high-level conceptual sketch of how they might look like, shown in Fig. 5c). Under the condition that the channel select filter removes most of the OOB interference, the wanted signal power can be measured in the baseband by a simple power detector. This information, properly calibrated to account for in-band gains, can be used by a logic circuit which will drive the switching between the two front ends.

#### 4.5 Power-efficient AFEs through adaptation to out-of-band interference

The analysis of practical implications of the AFE power scaling laws is concluded by looking into how much power can be saved if the AFE adapts its linearity to the OOB interferer level. It is assumed that the wanted signal, whose level does not change, is accompanied by two interferers with total power  $p_I$  and equal, slowly time varying amplitudes, so that they can be well approximated by two tones.

<sup>5</sup> We reiterate that the front end power can be scaled down by *at least* the values given by the right hand side of (43) and (45), i.e. Fig. 5a) illustrates a lower bound on possible savings!

We analyze a receiver structure that is able to adjust its linearity in two discrete steps and in doing so, adapt to the fluctuating interference level. To this end, suppose that we have two analog front end designs at our disposal. One of them is designed for the worst-case interference level  $p_{I, \text{wc}}$  (a value commonly prescribed in communication standards) and its linearity is equal to  $p_{\text{IIP3,wc}}$ . On the other hand, the IP3 of the other design has been degraded down to the limits of implementability and is equal to  $p_{\text{IIP3,wc}}/\sqrt{\mu}$ <sup>6</sup>. Otherwise, the bandwidth and noise figure of the two front ends are the same.

The task of the receiver is to track the interference power and switch between the two front ends so that a minimum performance requirement is always satisfied,  $\text{SNDR} \geq \text{SNDR}_{\text{min}}$ , or equivalently, that the intermodulation distortion is always kept below a certain level:

$$p_{\text{IM3}} \leq p_{\text{IM3,wc}} = \frac{p_{I, \text{wc}}^3}{p_{\text{IIP3,wc}}^2}. \quad (46)$$

Condition (46) is met by a receiver which will tune its IP3 by switching between the described front ends in line with the following rule:

$$p_{\text{IIP3}} = \begin{cases} p_{\text{IIP3,wc}}, & \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}} < p_I \leq p_{I, \text{wc}}, \\ \frac{1}{\sqrt{\mu}} p_{\text{IIP3,wc}}, & p_I \leq \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}}, \end{cases} \quad (47)$$

with one front end with desired linearity being on and the other one switched off.

In order to characterize average power savings, it is not necessary to have the knowledge of the actual distribution of  $p_I$ . It is sufficient to assume that the probability of  $p_I > p_{I, \text{wc}}$  is negligible (which is why this case is not covered by the adaptation rule), and that only the probability  $\delta$  of interference being “high” is known, i.e.

$$\begin{aligned} \Pr \left\{ \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}} < p_I \leq p_{I, \text{wc}} \right\} &= \delta, \\ \Pr \left\{ p_I \leq \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}} \right\} &= 1 - \delta. \end{aligned} \quad (48)$$

As in the preceding section, we normalize the power consumption of the adaptive receiver with the power consumed by a non-adaptive receiver that utilizes only the high linearity front end. By using (19), we obtain

$$\frac{P_{\text{AFE, adaptive}}}{P_{\text{AFE, fix}}} = \begin{cases} 1, & \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}} < p_I \leq p_{I, \text{wc}}, \\ \frac{1}{\sqrt{\mu}}, & p_I \leq \frac{1}{\sqrt[3]{\mu}} p_{I, \text{wc}}, \end{cases} \quad (49)$$

<sup>6</sup> This value is chosen in line with considerations from Section 4.1 and provides a fair comparison with other results in this section.



which, combined with (48), yields

$$\mathbb{E} \{ \zeta_P \} = \delta + \frac{1 - \delta}{\sqrt{\mu}}. \quad (50)$$

Average power savings of such a receiver are shown in Fig. 5b). For example, given that  $\mu = 10$  dB, the range of OOB interferer values for which the high linearity AFE is activated (worst-case interference) is  $(0.46 p_{I, wc}, p_{I, wc})$ . If the interference power is inside this range for 10% of the time, the low linearity AFE would be used for the remaining 90% of the time and the average power savings compared to a non-adaptive design are 60%. Taking the ballpark power numbers for a front end from [21], this signifies a reduction of average front end power from 35 mW to 14 mW. Paper [21] also suggests a practical implementation of the interference sensing circuit, consisting of a passband filter and an energy detector. We include this sensor in the high-level conceptual illustration of an interference-adaptive receiver, shown in Fig. 5d). The sensor from [21] consumes 10 mW, which combined with the reduced average AFE power consumption (and neglecting the consumption of the logic circuitry) yields 24 mW, which is still 30% less than the power consumed by the non-adaptive receiver.

## 5 Conclusion

Based on a known result from circuit theory that has also been verified in practice, we determine scaling laws between performance and power consumption of an analog front end (AFE). The power consumption of the AFE is found to scale as  $SIR^{-3/2}$  and at least as  $SNDR^{3/2}$ . These simple scaling laws can be used in a wide variety of communication-theoretic contexts, and some of the most important ones are explored. Namely, the power-SNR scaling law is extended to find the scaling laws between AFE power consumption and QAM constellation size, symbol error probability for QAM and error control coding gain and rate. Some general rules for low-power system design can be drawn from these laws: one example rule is that low-power applications favor “light” channel codes with moderate coding gains (such as simple convolutional codes) over more powerful ones, like turbo codes. Moreover, we derive laws that describe how front end power scales with environment parameters when performance is kept constant. Combined with fading and out-of-band blocker statistics, this enables us to determine theoretical average power savings of AFEs that adapt to the environment. The impressive results (about one order of magnitude reduction of power consumption in some cases) indicate that designing the front end so that it adapts to the environment is definitely a worthwhile effort.

## Acknowledgement

The authors would like to thank the Swedish Foundation for Strategic Research (SSF), which provided the funding of this research in the scope of the Digitally Assisted Radio Evolution (DARE) project.



# Bibliography

- [1] A. A. Abidi, G. J. Pottie and W. J. Kaiser, "Power-conscious design of wireless circuits and systems," in *Proceedings of the IEEE*, vol. 88, no. 10, pp. 1528-1545, Oct. 2000.
- [2] C. Svensson, "Towards power centric analog design," in *IEEE Circuits and Systems Magazine*, vol. 15, no. 3, pp. 44-51, 2015.
- [3] W. Sheng, A. Emira and E. Sanchez-Sinencio, "CMOS RF receiver system design: a systematic approach," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 53, no. 5, pp. 1023-1034, May 2006.
- [4] P. G. M. Baltus and R. Dekker, "Optimizing RF front ends for low power," in *Proceedings of the IEEE*, vol. 88, no. 10, pp. 1546-1559, Oct. 2000.
- [5] J. H. C. van den Heuvel, Y. Wu, P. G. M. Baltus, J. P. P. M. G. Linnartz and A. H. M. van Roermund, "Front end power dissipation minimization and optimal transmission rate for wireless receivers," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 5, pp. 1566-1577, May 2014.
- [6] M. Meghdadi and M. Sharif Bakhtiar, "Two-dimensional multi-parameter adaptation of noise, linearity, and power consumption in wireless receivers," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 8, pp. 2433-2443, Aug. 2014.
- [7] A. V. Do, C. C. Boon, M. A. Do, K. S. Yeo and A. Cabuk, "An energy-aware CMOS receiver front end for low-power 2.4-GHz applications," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 10, pp. 2675-2684, Oct. 2010.
- [8] G. Hueber, J. Zipper, R. Stuhlberger and A. Holm, "An adaptive multi-mode RF front-end for cellular terminals," 2008 IEEE Radio Frequency Integrated Circuits Symposium, Atlanta, GA, 2008, pp. 25-28.

- 
- [9] R. Senguttuvan, S. Sen and A. Chatterjee, "Multidimensional adaptive power management for low-power operation of wireless devices," in *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 55, no. 9, pp. 867-871, Sept. 2008.
  - [10] S. Sen, D. Banerjee, M. Verhelst and A. Chatterjee, "A power-scalable channel-adaptive wireless receiver based on built-in orthogonally tunable LNA," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 946-957, May 2012.
  - [11] D. Banerjee, S. Sen, A. Banerjee, and A. Chatterjee, "Low-power adaptive RF system design using real-time fuzzy noise-distortion control," *ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED)*, Jul. 2012, pp. 249-254.
  - [12] D. Banerjee, A. Banerjee and A. Chatterjee, "Adaptive RF front-end design via self-discovery: using real-time data to optimize adaptation control," *2013 26th International Conference on VLSI Design and 2013 12th International Conference on Embedded Systems*, Pune, 2013, pp. 197-202.
  - [13] D. Banerjee, S. K. Devarakond, X. Wang, S. Sen and A. Chatterjee, "Real-time use-aware adaptive RF transceiver systems for energy efficiency under BER constraints," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 8, pp. 1209-1222, Aug. 2015.
  - [14] D. Banerjee, B. Muldrey, X. Wang, S. Sen and A. Chatterjee, "Self-learning RF receiver systems: process aware real-time adaptation to channel conditions for low power operation," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 1, pp. 195-207, Jan. 2017.
  - [15] B. Razavi, *RF Microelectronics*, 2nd ed. New York, NY, USA: Prentice-Hall, 2011.
  - [16] J. Borremans et al., "A 40 nm CMOS 0.4–6 GHz Receiver Resilient to Out-of-Band Blockers," in *IEEE Journal of Solid-State Circuits*, vol. 46, no. 7, pp. 1659-1671, July 2011.
  - [17] J. Xu, J. Yao, L. Wang, Z. Ming, K. Wu and L. Chen, "Narrowband Internet of Things: Evolutions, Technologies, and Open Issues," in *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 1449-1462, June 2018.
  - [18] S. Sen, "Invited: Context-aware energy-efficient communication for IoT sensor nodes," *2016 53rd ACM/EDAC/IEEE Design Automation Conference (DAC)*, Austin, TX, 2016, pp. 1-6.

- [19] J. G. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw Hill, 2014.
- [20] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.
- [21] M. Abdulaziz, W. Ahmad, A. Nejedel, M. Törmänen and H. Sjöland, "A cellular receiver front-end with blocker sensing," 2016 IEEE Radio Frequency Integrated Circuits Symposium (RFIC), San Francisco, CA, 2016, pp. 238-241.
- [22] C. Studer, S. Fateh, C. Benkeser and Q. Huang, "Implementation trade-offs of soft-input soft-output MAP decoders for convolutional codes," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 11, pp. 2774-2783, Nov. 2012.
- [23] S. Belfanti, C. Roth, M. Gautschi, C. Benkeser and Q. Huang, "A 1Gbps LTE-advanced turbo-decoder ASIC in 65nm CMOS," 2013 Symposium on VLSI Circuits, Kyoto, 2013, pp. 284-285.
- [24] S. L. Howard, C. Schlegel, and K. Iniewski, "Error control coding in low-power wireless sensor networks: when is ECC energy-efficient?," in *EURASIP Journal on Wireless Communications and Networking*, pp. 1-14, 2006.
- [25] P. Grover, K. Woyach and A. Sahai, "Towards a communication-theoretic understanding of system-level power consumption," in *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1744-1755, September 2011.
- [26] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, 1970.



*Paper II*





# When are Low Resolution ADCs Energy Efficient in Massive MIMO?

Massive MIMO (MaMI) is often promoted as a technology that will enable the use of low-quality, cheap hardware. One particular component that has been in the focus of MaMI-related research is the analog-to-digital converter (ADC), and use of very low resolution ADCs has been proposed. However, studies about whether this strategy is justified from an energy-efficiency point of view have largely been inconclusive. In this work, we choose system setup and models that reflect the hardware implementation reality as close as possible and perform a parametric analysis of uplink energy efficiency as a function of ADC resolution. If antenna scaling and decrease of ADC resolution are considered independently, the energy efficiency is shown to be maximized at intermediate ADC resolutions, typically in the range of 4 - 8 bits. Moreover, optimal ADC resolution does not decrease when more antennas are used except in some specific cases, and when it does, the decrease is approximately logarithmic in the number of antennas. In the case when antenna scaling and ADC degradation are coupled through a constant-performance constraint, it is shown that energy efficiency cannot improve with reduced bit resolution unless the power consumption of blocks other than ADCs scales down with the up-scaling of antennas at a fast enough rate. Altogether it is concluded that in MaMI, intermediate ADC resolutions are optimal in energy efficiency sense, and, except in some special cases, scaling up the antennas to very large numbers does not change this conclusion.



## 1 Introduction

Wireless engineers and researchers are increasingly recognizing the potential of equipping base stations with a large number of antennas. Introduced in [1] and most often referred to as Massive MIMO (MaMI), this technique promises substantial increase in system throughput while simultaneously allowing for reduced radiated power both at the base station and at user terminals [2]. Another revolutionary benefit of MaMI is that the use of simple linear processing in the uplink and downlink becomes asymptotically optimal [3].

MaMI also offers resilience to hardware impairments [4] - [6] and this feature indicates that the quality of the hardware can be reduced as the number of antennas is upscaled. A decrease in hardware quality can be utilized to reduce the power consumption of individual hardware components, since performance and power consumption are tightly connected. However, given that the number of antennas and corresponding RF chains grows, the overall power consumption (calculated by taking all hardware components into account) may decrease, stay the same, or grow, all depending on the exact relation between the performance and power consumption of individual components. A general overview of hardware scaling laws in MaMI is given in [6].

One hardware component whose function in MaMI has attracted particular attention is the analog-to-digital converter (ADC). Such interest is motivated by the fact that the power consumption of ADCs grows at least linearly with the sampling rate [26]. Therefore, ADCs might form a power consumption bottleneck when employed in MaMI systems with large bandwidth. However, a reduction of ADC power consumption could be achieved by reducing bit resolution. Though doing so would introduce additional distortion in the system, the aforementioned resilience of MaMI to hardware impairments means that this distortion is annulled. Moreover, reduction in the quality of the ADCs is followed by a reduction in their individual cost, potentially leading to cheaper base station receiver systems if the benefits of MaMI are leveraged in the right way. Following this baseline motivation, some analyses of the impact of reduced ADC resolution on the performance of MaMI have been performed, a significant portion of which focuses on the extreme case of using 1 bit quantization [7] - [13].

It is not clear, however, whether choosing ADCs with extremely low bit resolutions is justified from the point of view of overall energy efficiency of MaMI, defined as the ratio of system sumrate and power consumption. Moreover, analyses of this important problem are scarce and somewhat contradictory. The issue is partially analyzed in a generalized setting in [14], where energy efficiency of a general MIMO receiver is maximized by choosing the optimal distribution of bit resolutions across receiver chains, in combination with antenna selection.

For the chosen system setup, the average of optimal bit resolutions decreases as the number of antennas scales up at low SNRs, while remaining constant at high SNR. Furthermore, it is shown that with a large number of antennas, very low average bit resolution can be used at low SNRs (approximately 1.5 bits at -30 dB of SNR). The connection between bit resolution and energy efficiency in MaMI is briefly mentioned in [13], where the study concludes that energy efficiency is maximized when 1-bit ADCs are used. In stark contrast to these two works, the analysis in [15] concludes that using very low bit resolutions is not optimal in an energy efficiency sense, and that 4 - 5 bits of ADC resolution are optimal.

Obviously, differing assumptions on system setup have inevitably led to variation in conclusions concerning the connection between ADC resolution and energy efficiency. Hence, there is a need for a structured parametric analysis that will help reveal the underlying effects that determine the energy efficiency aspects of ADCs in MaMI, in connection to the most important system parameters (number of antennas, number of users, SNR, etc.). Moreover, such an analysis should employ models of hardware behavior that are realistic enough, which would help hardware and system designers reach a consensus on the design goals for ADCs to be used in MaMI base stations.

This contribution, which is an extension of the work presented in [16], employs such an analysis, offering answers to following important questions:

- Under which conditions does a reduction of ADC resolution lead to improved energy efficiency of the receiver system in the uplink, and which parameters play a decisive role here?
- In particular, will increasing the number of antennas make low ADC resolutions more energy efficient?

Principal findings of the work show that

- The value of ADC resolution that maximizes energy efficiency primarily depends on how much power is consumed by other blocks in the receiver. Optimal bit resolution increases as other blocks become more power consuming;
- As the number of antennas increases, the behavior of optimal ADC resolution is determined by what happens with the number of users. If the number of users is kept constant, then the optimal resolution decreases with the increase of number of antennas, and this decrease is slow (approximately logarithmic). If the number of users increases linearly with antennas, optimal resolution stays constant or even grows, depending on which linear processing scheme is used;

- Presence of a poorly filtered out-of-band interferer can drastically affect the choice of optimal resolution. Namely, for each 10 dB increase of interference power, optimal resolution increases by approximately one bit;
- If the antennas are scaled up and simultaneously the quality of all the receiver hardware (including ADCs) is degraded, a decrease of bit resolution will not yield an improvement of energy efficiency unless the power consumption of all receiver blocks other than ADCs is scaled down at a fast enough rate.

As pointed out previously, an important feature of this analysis is that the models and system setup are chosen so they are as close as possible to hardware implementation reality. In particular, the effect of automatic gain control (AGC) and its dependence on bit resolution are explicitly modeled; power consumption model for the ADC is based on results from circuit theory; and the impact of out-of-band interference on the performance is taken into account.

## 2 Preliminaries: ADC and AGC

### 2.1 ADC and AGC: principles of operation and performance measures

This work considers scalar Nyquist-rate ADCs having bit resolution  $b$  and performing uniform quantization with  $2^b$  quantizer output levels. Uniform quantization was chosen because it is commonly encountered in practical ADC designs.

Quantization  $Q(y)$  is a nonlinear mapping of  $y \in \mathbb{R}$  to a discrete set that results in additive distortion

$$q = Q(y) - y. \quad (1)$$

The nature of distortion  $q$  can be described as twofold, depending on the relation between the magnitude of  $y$  and an overload level  $Y_{ol}$ : if  $|y| > Y_{ol}$ , we say that the signal is “clipped” and consequently,  $q$  is referred to as *clipping* or *overload distortion* with variance  $\sigma_{ol}^2$ . On the other hand, if  $|y| \leq Y_{ol}$ , distortion  $q$  is referred to as *granular noise*.

Assume that signal  $y$  is Gaussian and that its dynamic range is set such that the overload distortion can be neglected and standard deviation of  $y$  is larger than the width of one quantization bin. For a uniform quantizer operating

on such input  $y$ , distortion  $q$  can be well approximated as being uniformly distributed, uncorrelated with the input and white [18], with

$$\mathbb{E}\{q^2\} \approx \frac{1}{3} Y_{oi}^2 2^{-2b} \triangleq \sigma_{\text{PQN}}^2. \quad (2)$$

This model is usually referred to as the pseudoquantization noise (PQN) model.

In practical systems, the dynamic range of input signal  $y$  is typically adjusted by an automatic gain control (AGC) variable gain amplifier that precedes the ADC. A commonly used design parameter for the AGC is input backoff  $\mu = Y_{oi}^2/\mathbb{E}\{y^2\}$ , and various performance criteria are used for determining values of  $\mu$ , with practical solutions often targeting to minimize the effects of overload distortion. In this work,  $\mu$  is set so that the deviation  $\delta\sigma_{\text{PQN}}^2 = |\mathbb{E}\{q^2\} - \sigma_{\text{PQN}}^2|/\sigma_{\text{PQN}}^2$  is equal to some predefined small value (which can typically be -10 to -20 dB). With  $\mu$  set in such a way, all the conditions for applying the PQN model will be satisfied. The resulting  $\mu^*(b)$ , obtained numerically, is approximated by a chord as

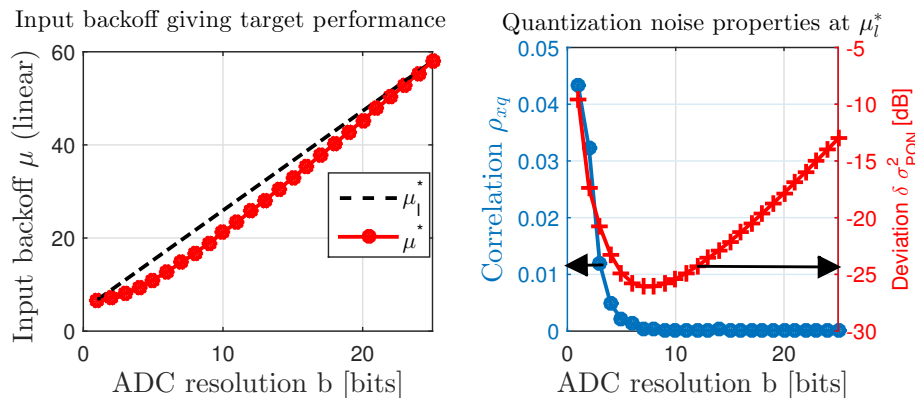
$$\mu^*(b) \approx \mu_i^*(b) = \theta_0 + \theta_1 b. \quad (3)$$

Deviation  $\delta\sigma_{\text{PQN}}^2$  and input-distortion crosscorrelation

$$\rho_{yq} = \mathbb{E}\{yq\} / \left( \sqrt{\mathbb{E}\{y^2\}} \sqrt{\mathbb{E}\{q^2\}} \right)$$

were obtained by simulations for  $b \in [1, 25]$  and  $\mu_i^*(b)$  and target  $\delta\sigma_{\text{PQN}}^2$  of -13 dB. The results are shown in Fig. 1 and illustrate how the PQN model applies well even for very low bit resolutions (1 bit) if AGC backoff is set properly.

Finally, we make a brief comparison of the PQN model with another commonly used signal model for ADCs. This model, referred to as additive quantization noise model (AQN), is employed in [13], among other works. It is derived under the assumption that the ADC - uniform or nonuniform - is designed to be optimal in MMSE sense. MMSE-optimal ADCs always have a nonzero correlation between the input and noise [17]. The key step in deriving the AQN is then the application of the Bussgang theorem, which results in a linear model with additive noise that is uncorrelated with the input, but with a compressive gain factor that effectively depends on input-noise correlation. Since the ADC considered in this work is designed based on criteria other than MMSE, the input signal and noise can still safely be assumed to be uncorrelated while there is no compressive gain factor involved, and this helps reduce the computational clutter in the analysis. However, a comparison of the results from works employing AQN and nonuniform ADCs and a matching subset of results from this work reveals that there are no significant differences in the fundamental conclusions.



**Figure 1:** Left: input backoff  $\mu$  with target deviation from PQN model of -13 dB. Right: deviation from PQN model and input-distortion correlation when linear approximation  $\mu_l^*$  is used.

## 2.2 ADC power consumption modeling

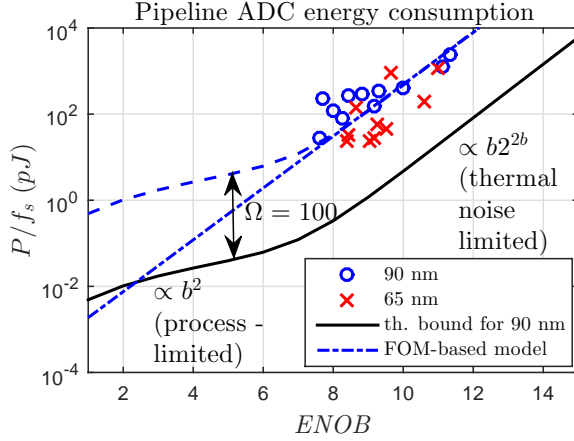
When it comes to choosing a model for the power consumption of the ADCs, we follow the general theme of this work - adopting system models that are relevant in practice. To this end, we assume a particular type of Nyquist-rate ADC that is likely to be used in practical base station implementations. We then perform a minor modification of an existing hardware-theoretic model for the power consumption of the chosen ADC type, and use this modified model as a realistic and representative model for the ADC power consumption.

The ADC type chosen for this purpose is the pipeline ADC. Pipeline ADCs are typically designed for intermediate bit resolutions and medium to high sampling rates  $f_s$ , with designs generally having power consumption that is comparatively superior to other types of ADCs when observed over a wide range of operating resolutions [21], [22], [23]. Moreover, a comparison of theoretical bounds on  $P_{\text{ADC}}$  between pipeline and other common types of Nyquist-rate ADCs - namely, flash and SAR ADCs - in [24] and [25] reveals that 1) flash ADCs have  $P_{\text{ADC}}$  that can be orders of magnitude higher than that of pipeline ADCs, and 2) power consumption of SAR ADCs follows the same functional trends as pipeline, while pipeline has overall lower  $P_{\text{ADC}}$ . These facts further corroborate the motivation to base the modeling of  $P_{\text{ADC}}$  on pipeline ADCs.

The basis for the model used here is a theoretical bound on power dissipation of pipeline ADCs presented in [24]. Figure 2 gives a comparison between this bound and pipeline ADC designs collected in [26], for the same values



of effective number of bits (ENOB). When it comes to the designs, ENOB is calculated from the measured SNDR, and for the bound, ENOB is assumed to be equal to  $b - 0.5$ . As the figure clearly shows, functional dependency in the bound matches the trend exemplified by state-of-the-art pipeline architectures. Notwithstanding, there is a gap (about two orders of magnitude wide) between the bound and the designs. Based on this observation, we modify the bound by applying a multiplicative factor  $\Omega$  and use this modification as an estimate of the power consumption of state-of-the-art ADC designs, as illustrated. The



**Figure 2:** ADC power consumption model, compared with actual pipeline ADC designs.

theoretical ADC power consumption model is therefore a modified version of the bound from [24], formulated as

$$P_{\text{ADC}}^{\text{th}} = \Omega (c_1 b + c_2 b^2 + c_3 2^{2b} + c_4 b 2^{2b}) f_s, \quad (4)$$

where factors  $c_1$  through  $c_4$  are given for completeness here as  $c_1 = 2C_{\min} V_{\text{FS}}^2$ ,  $c_2 = 12 \ln 2 V_{\text{eff}} V_{\text{FS}} C_{\min}$ ,  $c_3 = 216kT$ ,  $c_4 = 432 \ln 2 kT V_{\text{eff}} / V_{\text{FS}}$ . In the preceding expressions,  $V_{\text{eff}}$  is the effective voltage of the CMOS transistor (typically 80 - 100 mV),  $V_{\text{FS}}$  the full-scale range of the ADC,  $C_{\min}$  the minimum input capacitance of an inverter (CMOS process dependent, about 1 fF for 90 nm CMOS) and  $k$  and  $T$  are Boltzmann's constant and temperature in Kelvins, respectively. An important feature of this model is that the behavior of power consumption for low and intermediate bit resolutions is determined by CMOS process size through  $C_{\min}$ ; the relation between  $P_{\text{ADC}}$  and  $b$  in this region is approximately quadratic. At higher bit resolutions, functional properties of

power consumption are limited by thermal noise, and in this region of operation  $P_{\text{ADC}}$  is superexponential in  $b$ .

In addition to the model based on circuit theory, we also present a model for  $P_{\text{ADC}}$  based on the ADC figure of merit (FOM):

$$P_{\text{ADC}}^{\text{FOM}} = FOM_{\text{const}} 2^{2b} f_s, \quad (5)$$

with  $FOM_{\text{const}}$  that is extracted from state-of-the-art designs and assumed to be independent of  $b$ . This type of model is often employed in existing works on ADCs in MaMI, e.g. [15]. FOM-based model is also illustrated in Fig. 2. It is introduced in this work on the basis of its popularity and for the purpose of comparison with  $P_{\text{ADC}}^{\text{th}}$ ; since the latter is based on circuit theory, we consider it closer to reality and it is given larger weight when conclusions are drawn.

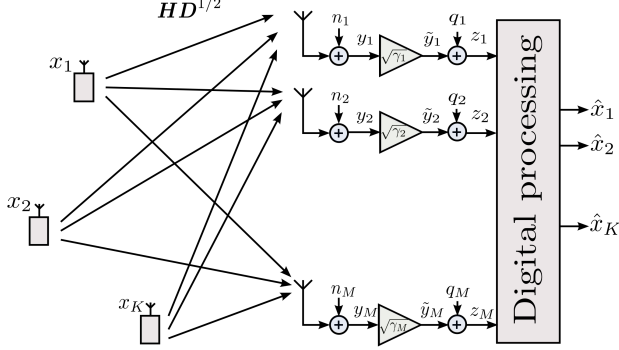
Finally, we note that in both models,  $P_{\text{ADC}}$  is linear in sampling rate  $f_s$ . The same trend is observed in actual ADC designs [26] up to very high sampling rates (on the order of 400 - 500 MHz).

### 3 System model

As an initial step in the energy efficiency analysis, we formulate the system model of the MaMI uplink that explicitly includes models of AGCs and ADCs. System setup assumed in this work is the following:

- Uplink of a single-cell MaMI system with  $M$  antennas and  $K$  single-antenna users;
- Narrowband, single-carrier transmission over bandwidth  $B$ . The system model can also represent one subcarrier in a multicarrier system, under the assumption that the quantization noise between different subcarriers is independent and has identical properties, and additionally that the input to the ADC and quantization noise are uncorrelated, as postulated by the ADC signal model;
- i.i.d. Rayleigh block fading over  $T$  symbols;
- Least-squares/maximum likelihood channel estimation performed using spatially orthogonal pilot sequences of length  $\tau$  in the uplink. Although suboptimal, ML channel estimation does not require any knowledge of channel statistics and is therefore favorable from the point of view of implementation complexity;
- Linear receiver processing using estimated channels - maximum ratio combining (MRC) and zero-forcing (ZF) receivers are considered.

An illustration of the uplink system model, where AGCs precede ADCs and ADCs are substituted by quantization noise sources based on the PQN, is given in Fig. 3.



**Figure 3:** Uplink system model with quantization noise

The PQN model is applied based on the assumption that the input signal to the ADCs is Gaussian, which holds in the case when the number of users is large or SNR is low. The complex baseband signal at the input of the digital processing unit is represented as

$$\mathbf{z} = \sqrt{p_u} \widetilde{\mathbf{H}} \mathbf{x} + \tilde{\mathbf{n}} + \mathbf{q}, \quad (6)$$

where  $p_u$  is the uplink transmit power,  $\mathbf{x}$  is the vector of user symbols with  $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{I}_K$  and  $\mathbf{q}$  is the quantization noise vector. Furthermore, the composite channel is represented by the matrix  $\widetilde{\mathbf{H}} = \mathbf{\Gamma}^{1/2} \mathbf{H} \mathbf{D}^{1/2}$ , where  $\mathbf{D}^{1/2} = \text{diag}(\sqrt{\beta_1} \dots \sqrt{\beta_K})$  is a diagonal matrix of effective amplitude path gains that subsumes the effects of geometric pathloss, large scale fading (LSF) and uplink power control;  $\mathbf{H}$  is the standard iid small-scale fading (SSF) matrix with unit variance elements; and  $\mathbf{\Gamma}^{1/2} = \text{diag}(\sqrt{\gamma_1} \dots \sqrt{\gamma_M})$  is the diagonal matrix of amplitude AGC gains. Individual AGC power gains are formed by combining the average received signal power at the input of the AGC with a proper backoff:

$$\gamma_m = \frac{2}{\mu_{l_m}^* \left( p_u \sum_{k=1}^K \beta_k + p_n + p_i \right)}. \quad (7)$$

The term  $p_i$  in (7) is the average power of an out-of-band (OOB) interfering signal that is assumed to be present at the input of the AGC due to limited

capabilities of analog filtering. The OOB interference signal is typically completely removed by digital baseband filter and therefore does not form a part of the digital baseband signal. However, its presence at the AGC input alters the AGC gain and consequently reduces the dynamic range of the useful signal at the ADC input. This can drastically affect the performance of the ADC, a fact that to the best of our knowledge has not been considered in traditional MaMI system level analyses focusing on the impact of ADCs, although it is all too familiar to ADC hardware designers. Lastly,  $\tilde{\mathbf{n}} = \mathbf{\Gamma}^{1/2}\mathbf{n}$ , where  $\mathbf{n}$  is the thermal noise vector with covariance  $\mathbb{E}[\mathbf{n}\mathbf{n}^H] = p_n\mathbf{I}_M$ .

Pilot sequences for channel estimation are contained in matrix  $\Phi = \sqrt{p_u\tau}\Psi$ , with  $\Psi\Psi^H = \mathbf{I}_{K \times K}$ ;  $\Phi$  is optimal for least-squares pilot-based channel estimation [19]. The least-squares channel estimate of  $\widetilde{\mathbf{H}}$  is of the form  $\widehat{\widetilde{\mathbf{H}}} = \widetilde{\mathbf{H}} + \widehat{\mathbf{H}}_\epsilon$ , with the impact of thermal and quantization noise modeled by  $\widehat{\mathbf{H}}_\epsilon$ . Channel estimates are used to formulate linear processing matrices  $\mathbf{A}$  for MRC and ZF, which can be split into a sum of two terms: one based on the actual channel and the other an error term. The split is exact for MRC and approximate for ZF, where the approximation holds if the SNR is sufficiently high [20]:  $\widehat{\mathbf{A}}_{\text{MRC}} = \mathbf{A}_{\text{MRC}} + \mathbf{A}_{\text{MRC},\epsilon} = \widetilde{\mathbf{H}} + \widehat{\mathbf{H}}_\epsilon$ ,  $\widehat{\mathbf{A}}_{\text{ZF}} \approx \mathbf{A}_{\text{ZF}} + \mathbf{A}_{\text{ZF},\epsilon} = \widetilde{\mathbf{H}}^\dagger - \widetilde{\mathbf{H}}^\dagger \widehat{\mathbf{H}}_\epsilon^H \widetilde{\mathbf{H}}^\dagger$ . Finally, the estimate of user symbols is obtained as  $\hat{\mathbf{x}} = \widehat{\mathbf{A}}^H \mathbf{z}$ .

The decomposition of  $\widehat{\mathbf{A}}$  allows for splitting the estimate of  $x_k$ , pertaining to the  $k$ th user, into a wanted signal term and several noise terms:

$$\hat{x}_k = \underbrace{\sqrt{p_u} \mathbf{a}_k^H \tilde{\mathbf{h}}_k}_{x_k^{(w)}} x_k + \underbrace{\sqrt{p_u} \sum_{j=1, j \neq k}^K \mathbf{a}_k^H \tilde{\mathbf{h}}_j}_{w_{\text{IUI}}} x_j + \underbrace{\mathbf{a}_k^H \tilde{\mathbf{n}}}_{w_n} + \underbrace{\mathbf{a}_k^H \mathbf{q}}_{w_q} + \underbrace{\sqrt{p_u} \mathbf{a}_{k,\epsilon}^H \widehat{\mathbf{H}} \mathbf{x}}_{w_{\text{IUI},\epsilon}} \quad (8)$$

$$+ \underbrace{\mathbf{a}_{k,\epsilon}^H \tilde{\mathbf{n}}}_{w_{n,\epsilon}} + \underbrace{\mathbf{a}_{k,\epsilon}^H \mathbf{q}}_{w_{q,\epsilon}},$$

where  $\mathbf{a}_k$ ,  $\mathbf{a}_{k,\epsilon}$  and  $\tilde{\mathbf{h}}_k$  are  $k$ th columns of  $\mathbf{A}$ ,  $\mathbf{A}_\epsilon$  and  $\widetilde{\mathbf{H}}$ , respectively. Note that in (8) we have implicitly defined the noise/interference terms, with subscripts IUI,  $n$  and  $q$  denoting interuser interference, thermal and quantization noise, respectively. Additionally, subscript  $\epsilon$  denotes that a particular term originated from imperfect CSI knowledge, and with perfect CSI, these terms are zero. Finally, assuming that  $M$  is large, the central limit theorem applies and all noise terms can be assumed zero-mean Gaussian.

## 4 System sumrate

With the system model well established and described, we move on to the next step of the energy efficiency analysis, namely, a study of uplink performance as function of the main system parameters. The metric used for quantifying the performance is the uplink sumrate.

### 4.1 Calculating the Sumrate

Using the post-processing noise and interference terms defined in (8), the signal-to-interference-and-thermal-and-quantization-noise-ratio (*SINQR*) for the  $k$ th user can be calculated as

$$SINQR_k = \frac{\mathbb{E} \left\{ |x_k^{(w)}|^2 \right\}}{\mathbb{E} \left\{ |w_{IUI}|^2 + |w_n|^2 + |w_q|^2 + |w_{IUI,\epsilon}|^2 + |w_{n,\epsilon}|^2 + |w_{q,\epsilon}|^2 \right\}}. \quad (9)$$

A simple summing up of the powers of noise terms in the denominator of (9) is possible because the PQN model applies, so data and thermal noise become uncorrelated with quantization noise. As the next step in sumrate calculation, we assume that bit resolution and AGC gain are same across all receiver chains, so  $b_m = b$  and  $\gamma_m = \gamma$ , and that quantization noise is uncorrelated across receiver chains. The assumption of no correlation across chains applies in all cases except when the channel is very highly correlated and pre-processing SNR is extremely high, but we do not consider these particular cases since they are not of practical interest. Finally, ergodic per-user rates  $R_k$  are calculated by averaging the per-block-and-user rates  $\log_2(1 + SINQR_k)$  over small-scale fading realizations  $\mathbf{H}$ . Standard application of Jensen's inequality and results from random matrix theory for central complex Wishart matrices [2], [27] yield

$$R_k = \mathbb{E}_{\mathbf{H}} \{ \log_2(1 + SINQR_k) \} \geq \log_2 \left( 1 + \frac{1}{\mathbb{E}_{\mathbf{H}} \left\{ \frac{1}{SINQR_k} \right\}} \right) \quad (10)$$

with

$$\begin{aligned} & \mathbb{E}_{\mathbf{H}} \left\{ \frac{1}{SINQR_k} \right\} \\ &= \frac{1}{p_u} \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \{ |w_{IUI}|^2 \}_{\text{eff}} + \mathbb{E} \{ |w_n|^2 \}_{\text{eff}} + \mathbb{E} \{ |w_q|^2 \}_{\text{eff}} \right. \\ & \quad \left. + \mathbb{E} \{ |w_{IUI,\epsilon}|^2 \}_{\text{eff}} + \mathbb{E} \{ |w_{n,\epsilon}|^2 \}_{\text{eff}} + \mathbb{E} \{ |w_{q,\epsilon}|^2 \}_{\text{eff}} \right\}. \end{aligned} \quad (11)$$

Individual terms under the  $\mathbb{E}_{\mathbf{H}}$  operator in (11) are given in Table 1. They represent the *effective* contributions of interuser interference, thermal noise,

etc. after linear processing. Some of these terms are (tightly) bounded by simpler ones owing to the fact that  $M$  is large, with relation operators indicating whether the expression is exact or a bound.

Very importantly for this analysis, the effects of quantization noise in Table 1 are represented by the *effective quantization noise* term (cf. (2), (3) and (7))

$$\tilde{p}_q = \frac{1}{3} \left( p_u \sum_{k=1}^K \beta_k + p_n + p_i \right) (\theta_0 + \theta_1 b) 2^{-2b}. \quad (12)$$

Consequently, effective quantization noise grows linearly with number of users, thermal noise and OOB interference power.

Since the bound in (10) and bounds in Table 1 have conflicting relation operators, per-user rates are expressed as approximations, which are very tight at large  $M$ :

$$R_k^{\text{MRC, ZF}} \approx \log_2 \left( 1 + \text{SINQR}_{\text{erg},k}^{\text{MRC, ZF}} \right), \quad (13)$$

where ergodic *SINQR* for MRC and ZF is given in Table 2. Finally, sumrate is calculated as

$$R = \frac{T - \tau}{T} \sum_{k=1}^K R_k \quad [\text{bps/Hz}]. \quad (14)$$

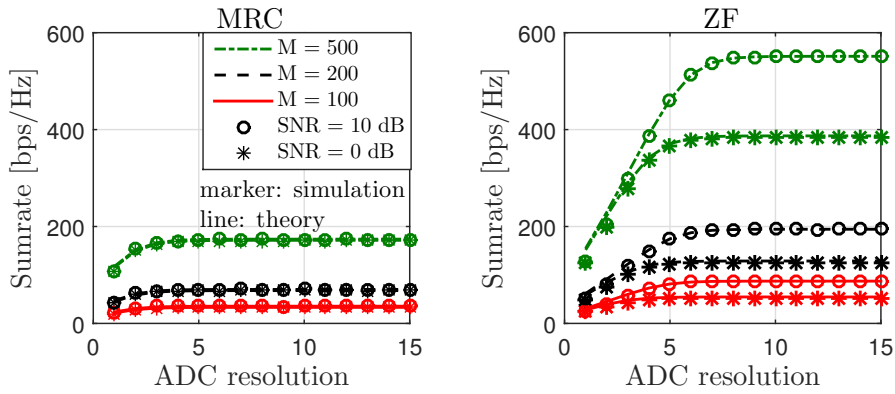
The performance model given by (14) aims to give a concise description of how system sumrate depends on the most relevant parameters. It is also open to further simplifications, if those are needed for the sake of clarity. One such simplification is the assumption that all effective amplitude path gains  $\beta_k$  are equal to 1; this effectively means that perfect power control is performed in the uplink. Since it is reasonable to assume that some form of uplink power control will be performed in an actual MaMI system as means of boosting the performance, we apply the perfect power control assumption throughout the analysis that follows.

**Table 1:**  $SINQR_k$  terms for MRC and ZF after averaging over channel realizations

	MRC	ZF
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_{\text{IUI}} ^2 \right\}_{\text{eff}} \right\}$	$= \frac{1}{\beta_k} \frac{1}{M-1} p_u \sum_{j=1, j \neq k}^K \beta_j$	$= 0$
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_n ^2 \right\}_{\text{eff}} \right\}$	$= \frac{1}{\beta_k} \frac{1}{M-1} p_n$	$= \frac{1}{\beta_k} \frac{1}{M-K} p_n$
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_q ^2 \right\}_{\text{eff}} \right\}$	$= \frac{1}{\beta_k} \frac{1}{M-1} \tilde{p}_q$	$= \frac{1}{\beta_k} \frac{1}{M-K} \tilde{p}_q$
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_{\text{IUI}, \epsilon} ^2 \right\}_{\text{eff}} \right\}$	$> \frac{1}{\beta_k^2} \left( \sum_{j=1}^K \beta_j \right) \frac{1}{\tau} \frac{1}{M-1} (p_n + \tilde{p}_q)$	$= \frac{1}{\beta_k} \frac{1}{\tau} \frac{K}{M-K} (p_n + \tilde{p}_q)$
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_{n, \epsilon} ^2 \right\}_{\text{eff}} \right\}$	$> \frac{1}{\beta_k^2} \frac{1}{\tau} \frac{p_n}{M-1} p_n (p_n + \tilde{p}_q)$	$> \frac{1}{\beta_k} \left( \sum_{j=1}^K \frac{1}{\beta_j} \right) \frac{1}{\tau} \frac{p_n}{(M-K)^2} p_n (p_n + \tilde{p}_q)$
$\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{  w_{q, \epsilon} ^2 \right\}_{\text{eff}} \right\}$	$> \frac{1}{\beta_k^2} \frac{1}{\tau} \frac{\tilde{p}_q}{M-1} p_n (p_n + \tilde{p}_q)$	$> \frac{1}{\beta_k} \left( \sum_{j=1}^K \frac{1}{\beta_j} \right) \frac{1}{\tau} \frac{\tilde{p}_q}{(M-K)^2} p_n (p_n + \tilde{p}_q)$

**Table 2:** Total ergodic  $SINQR_k$  for MRC and ZF

	MRC	ZF
$SINQR_{\text{erg}, k}^{\text{MRC, ZF}} \approx$	$\frac{p_u \beta_k (M-1)}{(p_u \sum_{j=1, j \neq k}^K \beta_j + p_n + \tilde{p}_q) \left( 1 + \frac{p_n + \tilde{p}_q}{p_u \beta_k \tau} \right) + \frac{p_n + \tilde{p}_q}{\tau}}$	$\frac{p_u \beta_k (M-K)}{\left\{ 1 + \frac{1}{\tau} \left[ K + \frac{p_n + \tilde{p}_q}{p_u (M-K)} \sum_{j=1}^K \frac{1}{\beta_j} \right] \right\} (p_n + \tilde{p}_q)}$



**Figure 4:** Ergodic sumrate as a function of SNR and ADC resolution  $b$ , simulated and theoretical.  $\lambda_s = 0.1$  [users/antenna],  $\lambda_t = 0.01$  [users/block],  $\tau = K$ .

## 4.2 Model validation and some preliminary observations regarding sumrate

The proposed model for system sumrate was compared to simulated ergodic sumrate. In order to reduce the dimensionality of the analysis, two auxiliary parameters were introduced, namely spatial loading,  $\lambda_s = K/M$  and temporal loading,  $\lambda_t = K/T$  of the system. System parameters of primary interest are ADC resolution  $b$ , number of antennas  $M$  and preprocessing SNR, defined as  $SNR = p_u/p_n$ . In all simulations,  $SNR$  and bit resolution during training and data transmission phases are set to be the same.

Results are shown in Figure 4 and show overall good agreement between theory and simulations. Interestingly, only 5 - 7 bits of ADC resolution in the receiver are sufficient to achieve almost full uplink sumrate; this observation is in line with some recent research [28]. It can be observed that this “saturation resolution” is not affected by  $SNR$  in the case of MRC, whereas it increases with  $SNR$  in the case of ZF. Moreover, it appears to be independent of  $M$  when MRC is used and to increase with  $M$  when ZF is used. Independence of the saturation point from  $SNR$  in the MRC case implies that IUI dominates over thermal noise for the chosen SNRs, and thus the sumrate is primarily influenced by the relation between IUI and quantization noise. On the other hand, in the ZF case, IUI is zero (if we neglect the interference leakage term due to use of imperfect CSI), so the interplay between thermal noise and quantization noise determines the behavior of the sumrate. The dependence of the saturation



resolution on  $SNR$  and  $M$  will be rigorously analyzed in the upcoming sections of the paper.

## 5 Energy efficiency analysis

Analysis of the impact of ADC resolution on performance (here represented by sumrate) tells only a part of the whole story. Namely, reduced ADC resolution will also lead to reduced power consumption of the entire system, and a comprehensive analysis needs to take both performance and power consumption into account through a scalar metric. A convenient metric is energy efficiency, which is calculated as

$$\eta = BR/P_{\text{tot}} \text{ [bits/J]}, \quad (15)$$

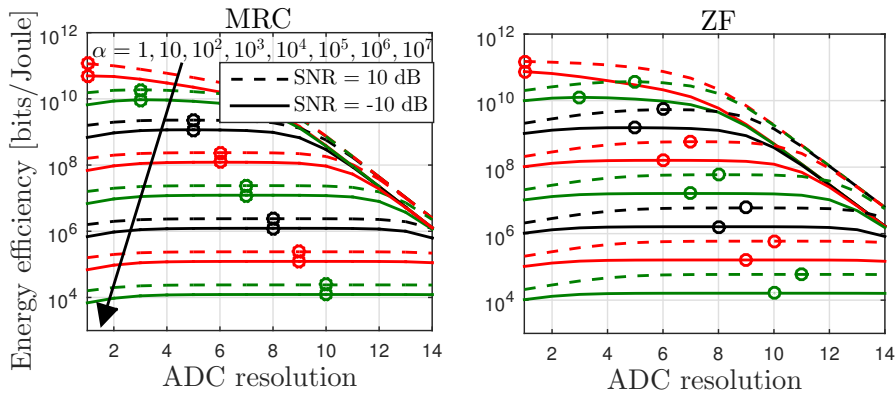
where  $B$  is system bandwidth,  $R$  is the uplink sumrate, calculated as per (14), and  $P_{\text{tot}}$  is total power consumption of the base station receiver. Note that here the focus is on the energy efficiency of the base station and the uplink power consumption of the users is left out of the analysis.

A power consumption model for the ADCs was already chosen in Section 2.2. What is left to do in order to obtain the complete  $P_{\text{tot}}$  is to model the power consumption of other receiver blocks, both analog and digital. Finding a general power consumption model that is both tractable and close to hardware design reality proves to be a challenging task, due to wide variations between system architectures, various techniques of practical implementation and a lack of unifying theoretical analysis. Therefore, a parametric approach is adopted in modeling  $P_{\text{tot}}$ .

To this end, power consumption of all the blocks excluding ADCs, denoted by  $P_{\text{rest}}$ , is normalized by  $P_{\text{ADC, ref}}$  - ADC power consumption calculated at a reference bit resolution  $b_{\text{ref}}$  - summed over all receiver chains. This normalized version of  $P_{\text{rest}}$ , denoted as

$$\alpha = \frac{P_{\text{rest}}}{2MP_{\text{ADC, ref}}} \quad (16)$$

is hereafter referred to as the *architecture parameter*. The primary goal with introducing the architecture parameter is to enable a parameterized analysis that covers a wide range of system architectures. It is given in a normalized form in order to better illustrate how the power consumption of the ADCs relates to the power consumption of the rest of the blocks, something that would be harder to see if we were working with  $P_{\text{rest}}$  given in absolute terms. Note that the choice of reference bit resolution  $b_{\text{ref}}$  is arbitrary.



**Figure 5:** Energy efficiency as a function of ADC resolution, SNR and architecture parameter. Channel estimation is performed using  $\tau = K$ .  $M = 100$ ,  $\lambda_s = 0.1$  [users/antenna] and  $\lambda_t = 0.01$  [users/block].  $b_{\text{ref}} = 2$ ,  $\Omega = 100$ .

Total power consumption of the base station receiver in the uplink is therefore calculated as

$$P_{\text{tot}} = 2MP_{\text{ADC}} + P_{\text{rest}} = 2M(P_{\text{ADC}} + \alpha P_{\text{ADC, ref}}) [W]. \quad (17)$$

The energy efficiency function  $\eta$ , calculated using (14) and (15) is plotted in Fig. 5, with the goal of gaining an initial insight in the behavior of  $\eta$  in a subspace of system parameters: ADC resolution  $b$ , SNR and architecture parameter  $\alpha$ . Processing using imperfect CSI is taken into account. Additionally, it was taken that  $\Omega = 100$  in the ADC power consumption model, and  $b_{\text{ref}}$  was set to be equal to 2.

The results show a general trend of degradation of  $\eta$  at very low and very high bit resolutions for  $\alpha = 10$  and larger. Degradation at low  $b$  is due to sum-rate being degraded when extremely low bit resolutions are used: on the other hand,  $\eta$  degrades at high bit resolutions due to the increase of  $P_{\text{ADC}}$ . Optimal bit resolutions have been obtained using a simple linear search: for most of the cases, intermediate bit resolutions (4 - 10 bits) are optimal. Generally, as  $P_{\text{rest}}$  becomes comparatively closer to power consumption of all the ADCs, lower ADC resolutions should be chosen to optimize total energy efficiency.

The initial results shown in Fig. 5 call for a more thorough and rigorous analysis of the dependence of  $\eta$  on  $b$  over the subspace of the most important system parameters -  $M$ , SNR,  $\alpha$  and  $K$ . Two different system setups are of particular interest:

- $M$  scales up, but there are no constraints on the effective contribution of thermal and quantization noise post-processing, i.e. the quality of hardware is not directly coupled with the increase of the number of antennas;
- Antennas scale up and the hardware quality is intentionally degraded so that the effective influence of thermal and quantization noise before and after scaling remains the same.

Both of the described setups are commonly encountered in existing work analyzing different aspects of low-resolution ADCs in MaMI. The first setup is employed in works that aim to find particular values of ADC resolution that either give a satisfying performance or maximize energy efficiency, such as [15]. On the other hand, the second setup is used in works that look into how ADC resolution (or hardware quality in general) scales with antennas when performance is fixed, e.g. [6]. As pointed out before, here we build upon this existing body of work by covering a wider range of system parameters with the aim of providing a more general analysis.

### 5.1 Hardware quality not coupled to scaling of $M$

In this section, we analyze how the optimal bit resolution in energy efficiency sense,

$$b_{\text{opt}} = \underset{b}{\operatorname{argmax}} \eta, \quad (18)$$

behaves as a function of  $M$ ,  $SNR$ ,  $\alpha$  and  $K$ . Although  $b_{\text{opt}}$  can be found numerically, such an approach does not provide much insight in its behavior. Rather, we choose to use the properties of constituent functions of  $\eta$  - sumrate  $R$  and power consumption  $P_{\text{tot}}$  - to find a tight approximation of  $\eta$ . By way of a rigorous mathematical analysis of this approximation, we find the lower and upper bounds of its optimum and in this way provide some valuable insights on the behavior of the optimum of the original energy efficiency function. The bounding approach will also yield some interesting side results on the behavior of sumrate as a function of ADC resolution (hints of which were seen in Section 4.2), which can serve as system design guidelines on their own.

**General assumptions:** we assume the variable  $b$  to be continuous instead of discrete. Additionally, we consider that CSI is perfectly known, so all terms in sumrate  $R$  that stem from channel estimation errors are 0. The perfect CSI assumption is introduced for the sake of improving the tractability of the analysis, and as it will soon be shown, the observations made using this assumption are valid also in the case of imperfect CSI.

Given these simplifying assumptions, the lower bound on sumrates for MRC and ZF is given compactly as

$$R^{\text{MRC, ZF}} > \tilde{R}_b^{\text{MRC, ZF}} \triangleq K \log_2 \text{SINQR}_{\text{pCSI}}^{\text{MRC, ZF}}(M, K, p_u, p_n, p_i, b), \quad (19)$$

where

$$\begin{aligned} \text{SINQR}_{\text{pCSI}}^{\text{MRC, ZF}}(M, K, p_u, p_n, p_i, b) = \\ \frac{p_u(M - K^{1-\mathbb{I}_{\text{MRC}}})}{\mathbb{I}_{\text{MRC}}p_u(K-1) + p_n + \frac{1}{3}(p_uK + p_n + p_i)(\theta_0 + \theta_1b)2^{-2b}} \end{aligned} \quad (20)$$

and  $\mathbb{I}_{\text{MRC}}$  is an indicator function

$$\mathbb{I}_{\text{MRC}} = \begin{cases} 1, & \text{MRC,} \\ 0, & \text{ZF.} \end{cases} \quad (21)$$

Having established the approximation of the sumrate, we now start the analysis by examining general properties of the sumrate and power consumption functions. In the sequel, energy efficiency in (15) is analyzed as a product of the sumrate and the reciprocal of power consumption; representing  $\eta$  in this manner will add to the clarity of the analysis.

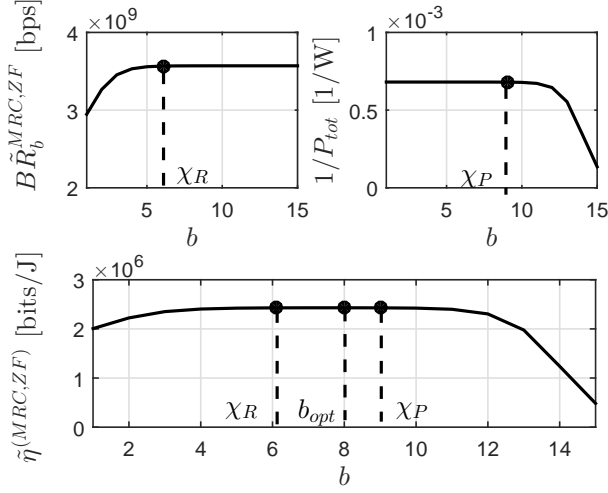
**Observation 1:** Sumrate lower bound  $\tilde{R}_b^{\text{MRC, ZF}}$  is monotonically increasing in  $b$ . Reciprocal of total power consumption,  $1/P_{\text{tot}}$ , is monotonically decreasing in  $b$ . Moreover, both  $\tilde{R}_b^{\text{MRC, ZF}}$  and  $1/P_{\text{tot}}$  exhibit saturating behavior, i.e. become practically constant for  $b > \chi_R$  (in case of  $\tilde{R}_b^{\text{MRC, ZF}}$ ) or  $b < \chi_P$  (in case of  $1/P_{\text{tot}}$ ), where  $\chi_R$  and  $\chi_P$  are some conveniently chosen values of ADC resolution.

**Proof of Observation 1:** The monotonical increase of  $\tilde{R}_b^{\text{MRC, ZF}}$  is shown by first establishing that the function  $p_{\text{q, ker}} = \left(1 + \frac{\theta_1}{\theta_0}b\right) 2^{-2b}$  is monotonically decreasing for  $b \geq 1$  and all  $\frac{\theta_1}{\theta_0} > 0$  and then by using the fact that a function of the form  $a/(b + f(x))$  is monotonically increasing if  $f(x)$  is monotonically decreasing. Monotonical decrease of  $1/P_{\text{tot}}$  is quickly proved by noting that

$$P_{\text{tot}} = 2M\Omega (c_1b + c_2b^2 + c_32^{2b} + c_4b2^{2b}) f_s + P_{\text{rest}} \quad (22)$$

is monotonically increasing with  $b$  and that its reciprocal is in turn monotonically decreasing. The saturating behavior of  $\tilde{R}_b^{\text{MRC, ZF}}$  is formally proved by showing that  $\lim_{b \rightarrow \infty} p_{\text{q, ker}} = 0$ . Hence, for a large enough  $\chi_R$  and  $b > \chi_R$ ,

$$\tilde{R}_b^{\text{MRC, ZF}} \approx K \log_2 \left[ \frac{p_u(M - K^{1-\mathbb{I}_{\text{MRC}}})}{\mathbb{I}_{\text{MRC}}p_u(K-1) + p_n} \right] \triangleq \tilde{R}_{\text{const}}^{\text{MRC, ZF}}. \quad (23)$$



**Figure 6:** Illustration of sumrate  $R$ , reciprocal of  $P_{\text{tot}}$  and energy efficiency  $\eta$ , shown together with saturation points  $\chi_R$  and  $\chi_P$  as well as  $b_{\text{opt}}$  for the case  $\chi_R < \chi_P$

To see that  $1/P_{\text{tot}}$  saturates, it suffices to observe from (22) that for large enough  $P_{\text{rest}}$  and small enough  $\chi_P$ ,

$$1/P_{\text{tot}} \approx 1/P_{\text{rest}}, \quad b < \chi_P. \quad \blacksquare \quad (24)$$

The energy efficiency function from (15), where  $R$  is substituted by  $\tilde{R}_b^{\text{MRC,ZF}}$ , is shown together with its constituent functions  $B\tilde{R}_b^{\text{MRC,ZF}}$  and  $1/P_{\text{tot}}$  in Fig. 6. It is clear that the shape of the energy efficiency function follows the saturating shapes of its two constituent functions. This enables us to define a very useful approximation of  $\eta$ .

**Observation 2:** Assume that the values  $\chi_R$  and  $\chi_P$  are given. Conditioned on their relative order, two piecewise approximations of the energy efficiency function  $\eta$  can be defined (symbols  $\nearrow$ ,  $\searrow$  and  $-$  represent monotonical increase, monotonical decrease and constant behavior of a function, respectively):

- For  $\chi_R < \chi_P$ :

$$\tilde{\eta}^{(\text{MRC,ZF})} = \begin{cases} \tilde{R}_b^{(\text{MRC,ZF})} \frac{1}{P_{\text{rest}}}, & b \leq \chi_R \quad (\nearrow) \\ \tilde{R}_{\text{const}}^{(\text{MRC,ZF})} \frac{1}{P_{\text{rest}}}, & \chi_R < b < \chi_P \quad (-) \\ \tilde{R}_{\text{const}}^{(\text{MRC,ZF})} \frac{1}{P_{\text{tot}}}, & b \geq \chi_P. \quad (\searrow) \end{cases} \quad (25)$$

- For  $\chi_P < \chi_R$ :

$$\tilde{\eta}^{(\text{MRC,ZF})} = \begin{cases} \tilde{R}_b^{(\text{MRC,ZF})} \frac{1}{P_{\text{rest}}}, & b < \chi_P \quad (\nearrow) \\ \tilde{R}_b^{(\text{MRC,ZF})} \frac{1}{P_{\text{tot}}}, & \chi_P \leq b \leq \chi_R \quad (\nearrow \times \searrow) \\ \tilde{R}_{\text{const}}^{(\text{MRC,ZF})} \frac{1}{P_{\text{tot}}}, & b > \chi_R \quad (\searrow) \end{cases} \quad (26)$$

**Proof of Observation 2:** Expressions (25) and (26) follow from establishing the relative order of  $\chi_R$  and  $\chi_P$  and then applying (19) and (22), together with their respective approximations (23) and (24) to (15). ■

Values  $\chi_R$  and  $\chi_P$  prove to be of essential importance to the analysis of the behavior of  $b_{\text{opt}}$  as various parameters change. Both of these terms will now be formally defined, and their properties examined.

**Definition 1:** Let  $\Delta_R \in (0, 1)$  denote the normalized deviation

$$\frac{\tilde{R}_{\text{const}}^{\text{MRC,ZF}} - \tilde{R}_b^{\text{MRC,ZF}}}{\tilde{R}_{\text{const}}^{\text{MRC,ZF}}} = \Delta_R. \quad (27)$$

Additionally, let  $\Delta_P \in (0, 1)$  be the normalized deviation

$$\frac{1/P_{\text{rest}} - 1/P_{\text{tot},b}}{1/P_{\text{rest}}} = \Delta_P. \quad (28)$$

The saturation point for the sumrate,  $\chi_R^{\text{MRC,ZF}}$  is defined as the value of  $b$  at which the normalized deviation of  $\tilde{R}_b^{\text{MRC,ZF}}$  from  $\tilde{R}_{\text{const}}^{\text{MRC,ZF}}$  is equal to some arbitrarily chosen value. Likewise, the saturation point for the power consumption,  $\chi_P$  is defined as the value of  $b$  at which the normalized deviation of  $1/P_{\text{tot},b}$  from  $1/P_{\text{rest}}$  is equal to an arbitrarily chosen value.

**Remark:** Point  $\chi_R$  is of great interest in practical system design since it tells us explicitly how many bits of resolution are sufficient if we can tolerate some level of normalized sumrate degradation  $\Delta_R$ . Its properties and dependence on the most important system parameters, together with the properties of  $\chi_P$ , are given in the following

**Observation 3:** Define  $\phi(b) \triangleq c_1 b + c_2 b^2 + c_3 2^{2b} + c_4 b 2^{2b}$ . The saturation point  $\chi_P$  is found as a solution of the transcendental equation

$$\phi(b) = \frac{\Delta_P}{1 - \Delta_P} \alpha \phi(b_{\text{ref}}). \quad (29)$$

Assume that  $p_i = 0$  and  $M$  and  $K$  are large. Saturation points  $\chi_R^{\text{MRC}}$  and  $\chi_R^{\text{ZF}}$  are then found as solutions of the following transcendental equations:

$$\text{MRC : } (\theta_0 + \theta_1 b) 2^{-2b} = 3 \left( \frac{M}{K + \frac{1}{\text{SNR}}} \right)^{\Delta_R} - 1, \quad (30)$$

$$\text{ZF : } (\theta_0 + \theta_1 b)2^{-2b} = 3 \frac{(M - K)^{\Delta_R}}{K} \text{SNR}^{\Delta_R - 1}. \quad (31)$$

With regards to  $\chi_R$ , the following important trends can be observed:

- For practical values of  $\text{SNR}$ ,  $\chi_R^{\text{MRC}}$  is independent of  $\text{SNR}$ , whereas  $\chi_R^{\text{ZF}}$  increases with increasing  $\text{SNR}$ ;
- When spatial loading  $\lambda_s$  is kept constant as the number of antennas  $M$  increases,  $\chi_R^{\text{MRC}}$  is independent of  $M$ , whereas  $\chi_R^{\text{ZF}}$  increases with increasing  $M$ ;
- When number of users  $K$  is kept constant as  $M$  increases, both  $\chi_R^{\text{MRC}}$  and  $\chi_R^{\text{ZF}}$  decrease with increasing  $M$ .

Additionally, the saturation point  $\chi_P$  is observed to increase with increasing  $\alpha$ .

**Proof of Observation 3:** The transcendental equation in (29) results directly from (22), (24) and (28). Likewise, (30) and (31) follow from plugging in (19) and (23) into (27) and applying the large- $M$ , large- $K$  assumption.

To prove the observed behavior of  $\chi_R$  with  $\text{SNR}$ , we examine the right hand sides (RHS) of the equations in (30) and (31), labeled here as  $\text{RHS}_{\text{MRC}}$  and  $\text{RHS}_{\text{ZF}}$ . When  $K \gg 1/\text{SNR}$ ,  $\text{RHS}_{\text{MRC}}$  is essentially independent of  $\text{SNR}$ . On the other hand,  $\text{RHS}_{\text{ZF}} \propto \text{SNR}^{\Delta_R - 1}$ , which decreases with  $\text{SNR}$  because  $\Delta_R < 1$ . Now, taking into account the fact that left hand sides (LHS) of equations in (30) and (31) are decreasing functions of  $b$ , we can conclude that  $\chi_R$  as the argument of LHSs is essentially independent of  $\text{SNR}$  in the case of MRC, whereas it increases with  $\text{SNR}$  when ZF is used.

For the case  $\lambda_s = \text{const.}$ , we note that  $\lim_{M \rightarrow \infty} \text{RHS}_{\text{MRC}} \approx 3(1/\lambda_s^{\Delta_R} - 1)$ , which does not depend on  $M$ , and that  $\text{RHS}_{\text{ZF}} \propto M^{\Delta_R - 1}$ , which decreases with  $M$ . Therefore, we can conclude that  $\chi_R$  stays constant as  $M$  increases in case of MRC, and increases with  $M$  in case of ZF. On the other hand, when  $K = \text{const.}$ , directly from (30) and (31) we see that  $\text{RHS}_{\text{MRC}}$  and  $\text{RHS}_{\text{ZF}}$  grow with  $M$  for both MRC and ZF, so then both  $\chi_R^{\text{MRC}}$  and  $\chi_R^{\text{ZF}}$  decrease with  $M$ .

Finally, to show that  $\chi_P$  indeed grows with  $\alpha$ , it suffices to notice that  $\phi(b)$  increases with  $b$ . As the RHS of the equation in (29) increases with  $\alpha$ ,  $\chi_P$  as the argument of  $\phi(b)$  then also has to increase with  $\alpha$ . ■

In practical systems, the choice of adequate ADC resolution is heavily influenced by the level of unfiltered out-of-band (OOB) interference  $p_i$ , as illustrated by the following

**Observation 4:** define the signal-to-interference ratio  $\text{SIR} = p_u/p_i$ . Assuming  $M \gg 1$ ,  $K \gg 1$  and a  $\text{SIR} \ll 1$  (so that OOB interference dominates over the useful signal),  $\chi_R$  for both MRC and ZF grows as  $\text{SIR}$  decreases.

**Proof of Observation 4:** By plugging (19) and (23) in (27) and applying the assumptions,  $\chi_R$  for MRC and ZF is found as the solution of transcendental equations

$$\text{MRC : } (\theta_0 + \theta_1 b)2^{-2b} = 3 \text{ SIR } K \left[ \left( \frac{M}{K + \frac{1}{\text{SNR}}} \right)^{\Delta_R} - 1 \right], \quad (32)$$

$$\text{ZF : } (\theta_0 + \theta_1 b)2^{-2b} = 3 \text{ SIR } [\text{SNR}^{\Delta_R} (M - K)^{\Delta_R}]. \quad (33)$$

Both RHS in the equations in (32) and (33) decrease as  $\text{SIR}$  decreases, and since the LHS decrease with  $b$ ,  $\chi_R$  as the argument of the LHS increases with the decrease of  $\text{SIR}$ . ■

Although saturation points are important for system analysis on their own, they also serve a convenient purpose of bounding  $b_{\text{opt}}$ . This means that the behavior of  $b_{\text{opt}}$  in conjunction with important system parameters is directly determined by how  $\chi_R$  and  $\chi_P$  behave. These important facts are formally stated in

**Observation 5:** The value of ADC resolution that maximizes the approximate energy efficiency  $\tilde{\eta}$ , denoted by  $\tilde{b}_{\text{opt}}$ , is always found between saturation points  $\chi_R$  and  $\chi_P$ , formally:

$$\min\{\chi_R, \chi_P\} \leq \tilde{b}_{\text{opt}} \leq \max\{\chi_R, \chi_P\}. \quad (34)$$

If it is assumed that  $\chi_P$  is independent of  $M$  and that  $M$  and  $K$  are large, then the following properties hold:

- For  $\chi_R < \chi_P$ ,  $\tilde{b}_{\text{opt}}$  decreases with decreasing  $\alpha$ ;
- In the case when ZF is used and  $\chi_R < \chi_P$ ,  $\tilde{b}_{\text{opt}}$  increases with increasing  $\text{SNR}$ .

Additionally, the following behavior of  $\tilde{b}_{\text{opt}}$  with increasing  $M$  is observed, depending on how  $K$  relates to  $M$ :

- If  $\lambda_s = \text{const.}$ ,  $\tilde{b}_{\text{opt}}$  *cannot decrease* with increasing  $M$ ;
- If  $K = \text{const.}$ ,  $\tilde{b}_{\text{opt}}$  *cannot increase* with increasing  $M$ .

Finally, it can be observed that  $\tilde{b}_{\text{opt}}$  cannot decrease with decreasing  $\text{SIR}$ .

**Proof of Observation 5:** For the case  $\chi_R < \chi_P$ , we refer to (25) and find that  $\tilde{\eta}$  is maximized for  $\chi_R < b < \chi_P$ , and therefore all  $b \in (\chi_R, \chi_P)$  maximize  $\tilde{\eta}$ . On the other hand, when  $\chi_P < \chi_R$ , we can refer to (26) and focus on the case  $\chi_P \leq b \leq \chi_R$ . As  $\tilde{\eta}$  is continuous, the extreme value theorem applies for



$\chi_P \leq b \leq \chi_R$ , and therefore  $\tilde{\eta}$  will have a local maximum  $\tilde{b}^* \in [\chi_P, \chi_R]$ . Since  $\tilde{\eta}$  is increasing for  $b < \chi_P$ ,  $\tilde{\eta}(\tilde{b}^*) > \tilde{\eta}(b)$ ,  $\forall b < \chi_P$ . On the other hand, since  $\tilde{\eta}$  decreases for  $b > \chi_R$ ,  $\tilde{\eta}(\tilde{b}^*) \geq \tilde{\eta}(b)$ ,  $\forall b > \chi_R$ . Therefore,  $\tilde{b}^*$  maximizes  $\tilde{\eta}$  over the entire range of  $b$ . Overall, we conclude that  $\tilde{b}_{\text{opt}}$  can always be found between  $\chi_R$  and  $\chi_P$ , regardless of their positions relative to one another.<sup>7</sup>

As for the behavior of  $\tilde{b}_{\text{opt}}$  with different parameters, we start by noting that in the case  $\chi_R \leq \tilde{b}_{\text{opt}} \leq \chi_P$ ,  $\tilde{b}_{\text{opt}}$  must grow as its lower bound grows and, likewise, must decrease if its upper bound decreases. Since  $\chi_P$  decreases with  $\alpha$  and  $\chi_R$  increases with  $SNR$  (in the case when ZF is used), as shown in Observation 3,  $\tilde{b}_{\text{opt}}$  needs to follow their decrease/increase accordingly.

Before we carry on to the final and most important observations on the connection between  $\tilde{b}_{\text{opt}}$  and  $M$ , we first turn our attention to the assumption that  $\chi_P$  is independent of  $M$ . From (29), we see that this claim is equivalent to saying that  $\alpha$  is independent of  $M$ , and from (16) this in turn implies that  $P_{\text{rest}}$  is linear in  $M$ . A deeper look on the power consumption model for MaMI base stations presented in [29] shows that the dominant part of base station power consumption indeed scales linearly with  $M$ , which serves as a confirmation of our linearity assumptions.

**Case 1 ( $\lambda_s = \text{const.}$ ):**  $\chi_R$  was shown in Observation 3 to either remain constant (in case of MRC) or increase with  $M$  (in case when ZF is performed). Therefore, when  $\chi_R \leq \tilde{b}_{\text{opt}} \leq \chi_P$ ,  $\tilde{b}_{\text{opt}}$  cannot decrease with  $M$ . Likewise, when  $\chi_P \leq \tilde{b}_{\text{opt}} \leq \chi_R$  and  $\chi_P$  is assumed to not change with  $M$ ,  $\tilde{b}_{\text{opt}}$  cannot decrease with  $M$  since it would otherwise conflict with its lower bound.

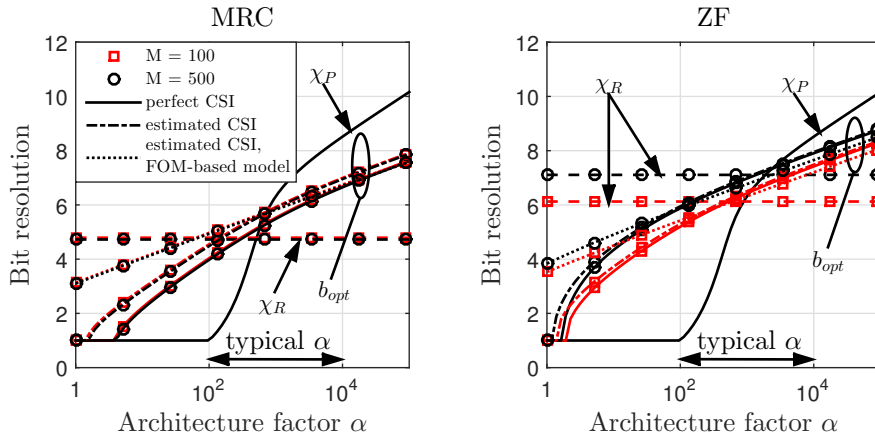
**Case 2 ( $K = \text{const.}$ ):** when  $\chi_R \leq \tilde{b}_{\text{opt}} \leq \chi_P$ ,  $\tilde{b}_{\text{opt}}$  cannot increase with  $M$  since it would eventually conflict with its upper bound. On the other hand, when  $\chi_P \leq \tilde{b}_{\text{opt}} \leq \chi_R$ ,  $\tilde{b}_{\text{opt}}$  must decrease with  $M$  since its upper bound is decreasing with  $M$ .

Dependence of  $\tilde{b}_{\text{opt}}$  on  $SIR$  can be proven using the same arguments as dependence on  $M$ . ■

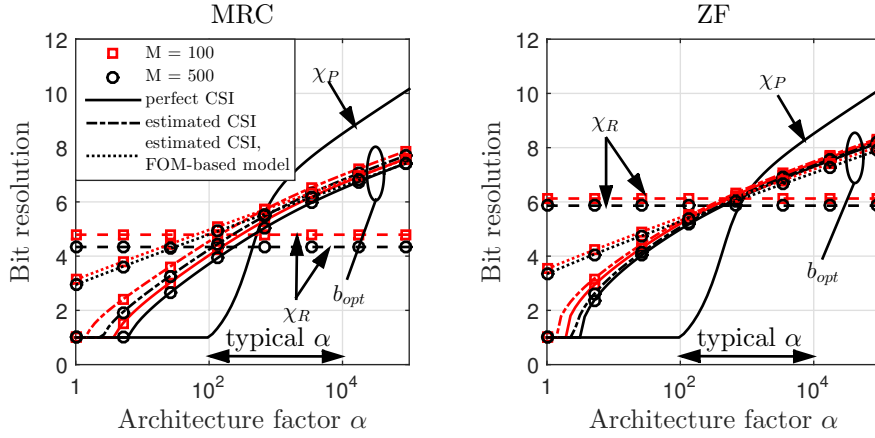
**Numerical results:** as an illustration of the analysis presented in Observations 1 to 5,  $b_{\text{opt}}$  have been found numerically for both the perfect-CSI and estimated-CSI cases. In the first set of results,  $M$  changes, and  $K$  either scales linearly with  $M$  or remains constant (illustrated in Figures 7 and 8, respectively), whereas in the second set, presented in Figure 9, different values of  $SIR$  are tested. Values of saturation points  $\chi_R$  and  $\chi_P$  were calculated numerically as well for the value of normalized deviation  $\Delta_P = \Delta_R = 5 \cdot 10^{-3}$ .

The value of  $b_{\text{ref}}$  was again taken to be 2. Using the model from [29], and

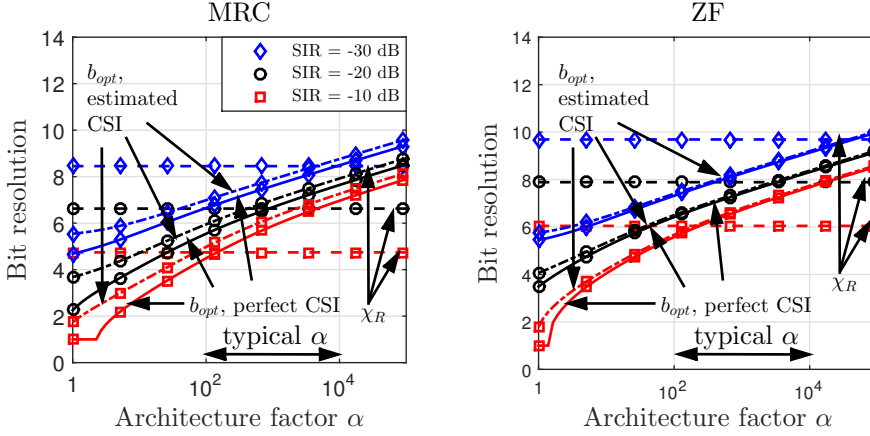
<sup>7</sup> It needs to be noted that this proof tacitly neglects the fact that  $\tilde{\eta}(b)$  is not smooth exactly at  $\chi_P$  and  $\chi_R$ . However, since  $\tilde{\eta}(b)$  is smooth in the limit as  $\Delta$  approaches zero, the proof is valid in this limit.



**Figure 7:** Bit resolution that maximizes receiver energy efficiency, together with saturation points, for varying  $M$  and  $\lambda_s = \text{const}$ . Channel estimation performed using  $\tau = K$ .  $SNR = 0 \text{ dB}$ ,  $\lambda_s = 0.1$  [users/antenna] and  $\lambda_t = 0.01$  [users/block].  $b_{\text{ref}} = 2$ ,  $\Omega = 100$ .



**Figure 8:** Bit resolution that maximizes receiver energy efficiency, together with saturation points, for varying  $M$  and  $K = \text{const}$ . Channel estimation performed using  $\tau = K$ .  $SNR = 0 \text{ dB}$ ,  $K = 10$  and  $\lambda_t = 0.01$  [users/block].  $b_{\text{ref}} = 2$ ,  $\Omega = 100$ .



**Figure 9:** Bit resolution that maximizes receiver energy efficiency, together with saturation points, for varying  $SIR = p_u/p_i$ . Channel estimation performed using  $\tau = K$ .  $SNR = 0$  dB,  $M = 100$ ,  $\lambda_s = 0.1$  [users/antenna] and  $\lambda_t = 0.01$  [users/block].  $b_{\text{ref}} = 2$ ,  $\Omega = 100$ .

assuming  $P_{\text{ADC}}^{\text{th}}$  with  $\Omega = 100$ , typical values of  $\alpha$  corresponding to  $b_{\text{ref}} = 2$  were calculated, and they turn out to be  $\approx 10^4$  for wide ranges of different system parameters. On the other hand, it is not unreasonable to assume that in the future, base station hardware will be implemented using integrated CMOS techniques. If we use known power consumption values for user equipment receiver chains implemented in CMOS and assume that similar power numbers might hold also for future MaMI base stations, then the base station power consumption might go down by one or two orders of magnitude compared to the model in [29]. Overall, admitting the fact that changes in technology, advances in hardware design etc. can change the initial result obtained from [29], we assume that values in the range  $10^2 - 10^4$  can be considered “typical” for  $\alpha$ .

In Observation 5, behavior of approximate optimum  $\tilde{b}_{\text{opt}}$  was analyzed, whereas the results shown in Figures 7 - 9 present the true optimum,  $b_{\text{opt}}$ . Nevertheless, numerical results are in perfect accordance with the theoretical analysis. Non-infinitesimal values of  $\Delta_R$  and  $\Delta_P$  can be accounted for the fact that  $\chi_P$  and  $\chi_R$  do not “sandwich”  $b_{\text{opt}}$  exactly around their crossing point, but this will not impact the results of the analysis in any significant way.

The architecture factor  $\alpha$  can be identified as the primary influence on which values of  $b$  will maximize energy efficiency. As predicted by theory,  $b_{\text{opt}}$  decreases with increasing  $M$  only in the case  $K = \text{const}$ , and this decrease is

slow (approximately logarithmic, as deduced from (30) and (31)). Therefore, considering a large-but-finite- $M$  regime and for practical values of  $\alpha$ , intermediate ADC resolutions (3 - 8, depending on scenario) are optimal in energy efficiency sense.

It is also of interest to analyze the values of  $\chi_R$ : they span the range 4 - 7 bits, which coincides with the values suggested in some works on ADC in MaMI (e.g. [28]) for ADC resolutions that could be used in MaMI with acceptable performance degradation. However, as with  $b_{\text{opt}}$ , we see that any possible reductions of  $\chi_R$  introduced by scaling up  $M$  to very large values are minuscule.

On the other hand, a very practical concern when deciding on a proper ADC resolution is poorly filtered OOB interference. As seen in Fig. 9, OOB interference can have a detrimental impact on  $b_{\text{opt}}$ ; when its power increases by 10 dB,  $b_{\text{opt}}$  increases by roughly 1 bit at lower values of  $\alpha$ .

Finally, we can shortly reflect on the influence of the model for  $P_{\text{ADC}}$  on  $b_{\text{opt}}$ . In Figs. 7 and 8,  $b_{\text{opt}}$  was calculated using the FOM-based model for the ADCs but with the same  $P_{\text{rest}}$  as for the theoretical model, to allow for a fair comparison. General trends of  $b_{\text{opt}}$  are invariant to the choice of ADC power consumption model, but using  $P_{\text{ADC}}^{\text{th}}$  proves important if choice of the ADC resolution needs to be fine-tuned; this choice would be off (under- or over-estimated, depending on  $\alpha$ ) if it were based on the more simplistic FOM-based model.

**Discussion and main takeaways:** Observation 5 brings forth the main message of this subsection and one of the main messages of the entire work. Namely, in the case when spatial loading  $\lambda_s$  is kept constant, increasing the number of antennas alone does not make ADCs with lower bit resolutions optimal in energy efficiency sense. ADCs with a very low resolution sometimes do maximize overall energy efficiency, but the reason for this is not the fact that a very large number of antennas is being used; rather, it is either due to preprocessing SNR or power consumption of other blocks. The prerequisite for making ADCs with low bit resolutions optimal by increasing the number of antennas is that the number of users remains constant as the number of antennas is increased. Even then, the decrease of  $b_{\text{opt}}$  with  $M$  is only logarithmic.

Furthermore, as shown in Observation 4, any potential decrease of optimal ADC resolution that could have been harvested by increasing the number of antennas can be reversed by the presence of a poorly filtered out-of-band interferer. This behavior is explained by the linear increase of effective quantization noise with OOB interference power.

The analysis in this subsection also produced a valuable side result regarding the values of “good enough” ADC resolution for acceptable sumrate degradation due to low precision ADC. This quantity may be of interest for

system designs where the net information flow from all the ADCs to the digital processing block should be reduced as much as possible without damaging the performance. Results in Observation 3 show that scaling up the antennas to very large numbers will not allow for the use of lower bit resolutions, unless the number of users is kept constant as antennas scale.

We conclude that, in the case when antenna scaling and ADC resolution decrease are not coupled, scaling up the antennas to very large numbers does not make a significant impact on the choice of ADC resolution. However, it is of interest to also investigate what happens when antenna and ADC scaling are coupled through a fixed-performance constraint, and this investigation is performed in the next subsection.

## 5.2 Hardware quality coupled with scaling of M

We now consider the case where the number of antennas is scaled up from  $M_1$  to  $M_2$  and, simultaneously, the quality of receiver chain hardware is degraded. Increasing the number of antennas allows for higher levels of pre-processing noise and distortion coming from lower quality hardware, because noise and distortion are effectively “averaged out” by signal processing, as indicated by the terms in Table 1. Higher levels of pre-processing distortion are expected to result in lower cost/power consumption of individual hardware components, and this is the reason why MaMI is promoted as being friendly to low-cost, low-power-consumption hardware [4] - [6]. However, the number of receiver chains also grows, so using low-quality hardware per receiver chain does not guarantee that the overall cost/power consumption will be reduced; also, there is possibly a residual impact on the performance. Again, the energy efficiency metric should be used to join together the performance and power consumption parts of the story.

The initial step of the analysis is establishing some simplifying assumptions. Again it is taken that all  $\beta_k = 1$  and that the CSI is perfectly known. Additionally, it is assumed that antennas and hardware scale such that the total effective postprocessing noise/distortion remains the same before and after the scaling, formally (from Table 1)

$$\begin{aligned} \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_n|^2 \right\}_{\text{eff}} \right\} (M_1) + \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_q|^2 \right\}_{\text{eff}} \right\} (M_1) = \\ \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_n|^2 \right\}_{\text{eff}} \right\} (M_2) + \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_q|^2 \right\}_{\text{eff}} \right\} (M_2). \end{aligned} \quad (35)$$

In order to further simplify the analysis, the condition (35) is substituted by a set of sufficient conditions pertaining individually to quantization and thermal

noise:

$$\begin{aligned}\mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_q|^2 \right\}_{\text{eff}} \right\} (M_1) &= \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_q|^2 \right\}_{\text{eff}} \right\} (M_2), \\ \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_n|^2 \right\}_{\text{eff}} \right\} (M_1) &= \mathbb{E}_{\mathbf{H}} \left\{ \mathbb{E} \left\{ |w_n|^2 \right\}_{\text{eff}} \right\} (M_2).\end{aligned}\quad (36)$$

The constraints in (36) can be written out explicitly for MRC and ZF as

$$\frac{\tilde{p}_{q2}^{\text{MRC, ZF}}}{\tilde{p}_{q1}} = \frac{p_{n2}^{\text{MRC, ZF}}}{p_{n1}} = \frac{M_2 - K^{1-\mathbb{I}_{\text{MRC}}}}{M_1 - K^{1-\mathbb{I}_{\text{MRC}}}}. \quad (37)$$

We further define the antenna scaling factor

$$\rho_M = M_2/M_1. \quad (38)$$

Given a value of  $\rho_M$  and initial bit resolution  $b_1$  and using (12) and (37), the post-scaling bit resolution  $b_2$  can be obtained, which through  $P_{\text{ADC}}$  can reveal how much the power consumption of individual ADCs changed after the scaling. Further, by assuming that  $p_n$  subsumes the impact on the performance of all the blocks excluding the ADCs, antenna scaling  $\rho_M$  should be connected with the change in power consumption of the other blocks,  $P_{\text{rest}}$ . For this we need an explicit relation between  $p_n$  and  $P_{\text{rest}}$  through the use of some intermediate per-block parameter(s), in analogy with how  $b$  connects  $\rho_M$  with  $P_{\text{ADC}}$ . This connection is typically difficult to find. We therefore directly assume that the constant-performance constraint induces a scaling of architecture factor  $\alpha$  that follows a power law

$$\alpha_2/\alpha_1 = \rho_M^\xi, \quad (39)$$

where  $\xi$  is a free scaling parameter. By referring to (16), we see that the law (39) results in power consumption of “other” blocks scaling as  $P_{\text{rest}2}/P_{\text{rest}1} = \rho_M^{\xi+1}$ . A power-law scaling of power consumption under the constant-performance constraint proves to be a valid behavioral model for receiver blocks that introduce additive distortion whose variance also scales with  $M$  following a power-law, and which are designed such that their figure of merit stays constant regardless of their quality [6].

In the analysis that follows, we observe the ratio of energy efficiencies  $\eta_2/\eta_1$  after and before the antenna/hardware scaling. Total power consumption model is given by (17). By taking all the relevant assumptions into account (with an additional assumption that  $b_{\text{ref}} = b_1$ ), the ratio of energy efficiencies before and after the scaling is given by the general expression

$$\frac{\eta_2^{\text{MRC, ZF}}}{\eta_1} = \frac{K_2 \log_2 \left[ 1 + \text{SINQR}_{\text{pCSI}}^{\text{MRC, ZF}}(M_2, K_2, p_u, p_{n2}, b_2) \right]}{K_1 \log_2 \left[ 1 + \text{SINQR}_{\text{pCSI}}^{\text{MRC, ZF}}(M_1, K_1, p_u, p_{n1}, b_1) \right]} \frac{1}{\rho_M} \frac{1 + \alpha_1}{\frac{P_{\text{ADC}2}}{P_{\text{ADC}1}} + \alpha_1 \rho_M^\xi}. \quad (40)$$

In cases where this ratio is larger than 1, it can be concluded that antenna/hardware scaling is beneficial from energy efficiency point of view. Similarly to the analysis in Section 5.1, we consider two fundamentally different cases:

- Spatial loading is kept constant as antennas scale,  $\lambda_s = \text{const.}$ ,
- Number of users is kept constant as antennas scale,  $K = \text{const.}$

**Case 1** ( $\lambda_s = \text{const.}$ ): we start with an important observation, valid irrespective of use of MRC or ZF.

**Observation 6:** Assume that  $M$  and  $K$  are large and constraint (36) is satisfied. It can be shown that, in order to satisfy the said constraint, ADC resolution must remain unchanged, i.e.  $b_1 = b_2$ .

**Proof of Observation 6:** For the case of MRC, we write out the first equation in (37) as

$$\frac{M_1 - 1}{M_2 - 1} \frac{p_u \lambda_s M_2 + p_{n2}}{p_u \lambda_s M_1 + p_{n1}} \frac{(\theta_0 + \theta_1 b_2) 2^{-2b_2}}{(\theta_0 + \theta_1 b_1) 2^{-2b_1}} = 1, \quad (41)$$

while the second equation in (37) gives

$$\frac{M_1 - 1}{M_2 - 1} \frac{p_{n2}}{p_{n1}} = 1. \quad (42)$$

By assuming a large number of antennas, we have  $M_1 - 1 \approx M_1$  and  $M_2 - 1 \approx M_2$ . Combined with the previous two equations, this allows us to write the entire constraint for postprocessing quantization noise when MRC is used as

$$1 = \frac{\mathbb{E}_{\mathbf{H}} \{ \mathbb{E} \{ |w_q|^2 \}_{\text{eff}} \} (M_2)}{\mathbb{E}_{\mathbf{H}} \{ \mathbb{E} \{ |w_q|^2 \}_{\text{eff}} \} (M_1)} \approx \quad (43)$$

$$\frac{1}{\rho_M} \frac{\text{SNR}_1 \lambda_s \rho_M M_1 + \rho_M}{\text{SNR}_1 \lambda_s M_1 + 1} \frac{(\theta_0 + \theta_1 b_2) 2^{-2b_2}}{(\theta_0 + \theta_1 b_1) 2^{-2b_1}} =$$

$$\frac{(\theta_0 + \theta_1 b_2) 2^{-2b_2}}{(\theta_0 + \theta_1 b_1) 2^{-2b_1}},$$

which is satisfied by  $b_1 = b_2$ . As for the case of ZF, using the property  $M - K = M(1 - \lambda_s)$ , the approximation in (43) becomes an equality, and therefore the conclusion is the same as for the case of MRC. ■

**Remark:** Observation 6 tells us that when  $K$  scales linearly with  $M$ , ADCs remain the same irrespective of how the antennas scale. This behavior is explained by the effects of the AGC: equivalent quantization noise  $\tilde{p}_q$  is proportional to received signal power and thermal noise, and both these quantities

grow as the number of antennas grows. Therefore, linear growth of  $\tilde{p}_q$  with  $M$  and the  $1/M$  factor in  $\mathbb{E}\{|w_q|^2\}_{\text{eff}}$  will cancel each other.

Next step in determining  $\eta_2/\eta_1$  entails looking into what happens with the sumrate. When MRC is used and constraints (37) are active, effective interuser interference term  $\frac{1}{M-1}p_u(\lambda_s M - 1)$  determines the ratio of sumrates before and after the scaling. By assuming that  $M$  and  $K$  are large, this term can be taken to be approximately equal before and after the scaling, so the sumrates for MRC can also be taken to be approximately equal. On the other hand, in the case of ZF and perfect CSI, interuser interference is 0, so the sumrates before and after the scaling are equal.

In summary, for the case when  $K$  scales linearly with  $M$ , the ratio of post- and pre-scaling energy efficiencies is given from (40) as

$$\frac{\eta_2^{\text{MRC, ZF}}}{\eta_1} = \frac{1 + \alpha_1}{1 + \alpha_1 \rho_M^\xi}. \quad (44)$$

This indicates that the behavior of energy efficiency with antenna scaling in this particular case does not depend on the ADC resolution. Moreover, it is easy to show from (44) that  $\eta_2/\eta_1 > 1$  for  $\rho_M > 1$  if and only if  $\xi < 0$ .

**Case 2** ( $K = \text{const.}$ ): in contrast to what was observed when  $K$  scaled linearly with  $M$ , when  $K$  is kept constant during the scaling, it is possible to degrade the ADC resolution and compensate for the degradation by increasing  $M$ .

**Observation 7:** Assume that  $K = \text{const.}$ , that  $M$  is large and that ADC resolutions before and after the antenna scaling are chosen such that the constraint (36) is satisfied. Given an initial ADC resolution  $b_1$  and a degradation of ADC resolution  $\Delta b < 0$  such that the ADC resolution after scaling is  $b_2 = b_1 + \Delta b$ , antenna scaling needed to keep the pre- and post-scaling thermal and quantization noise levels the same can be calculated as

$$\rho_M = \frac{\zeta_1}{\zeta_2 + \zeta_1 \zeta_2 - 1}, \quad (45)$$

where  $\zeta_1 = SNR_1 K$  and  $\zeta_2 = \frac{\theta_0 + \theta_1(b_1 + \Delta b)}{\theta_0 + \theta_1 b_1} 2^{2\Delta b}$ . Additionally, only the solutions of (45) where  $\rho_M > 1$  are taken into account, otherwise scaling is not performed.

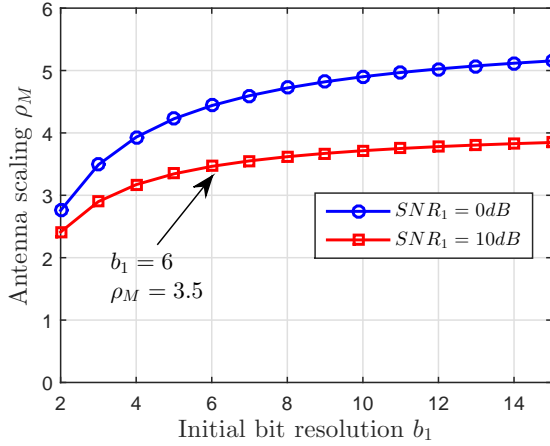
**Proof of Observation 7:** By employing all the listed assumptions together with (12) and (37), the following fixed-point equation connecting  $\Delta b$  and  $\rho_M$  is obtained, valid for both MRC and ZF cases <sup>8</sup>:

$$\Delta b = \frac{1}{2} \log_2 \left[ \frac{SNR_1 K + \rho_M}{\rho_M SNR_1 K + \rho_M} \frac{\theta_0 + \theta_1(b_1 + \Delta b)}{\theta_0 + \theta_1 b_1} \right]. \quad (46)$$

<sup>8</sup> The relation is only approximate but here it is represented as equality for convenience. This approximation is tight for large  $M$ .



Using elementary algebra, this relation can be reformulated as a closed-form equation for  $\rho_M$  as given by (45). ■



**Figure 10:** Antenna scaling vs. initial bit resolution, for  $\Delta b = -1$ . No. of users  $K = 10$ .

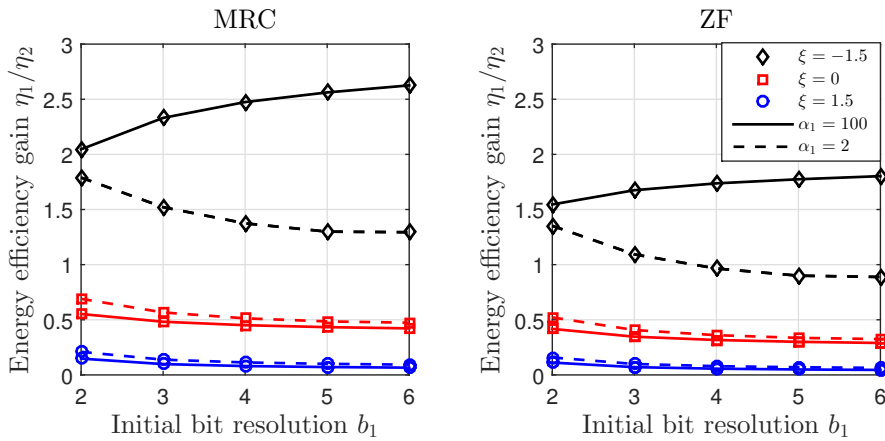
Antenna scaling as a function of initial bit resolution  $b_1$  is illustrated in Fig. 10, for two different values of  $SNR_1$ . The singled out data point serves as an example how the plot is interpreted: for  $SNR_1 = 10$  dB, degrading the ADC resolution from  $b_1 = 6$  bits to  $b_2 = b_1 + \Delta b = 5$  bits asks for an increase of number of antennas of 3.5 times, if we want to maintain the same level of postprocessing thermal and quantization noise. It is easily seen that the elementary case  $\Delta b = -1$  can be used to describe any arbitrary degradation of  $b$ , since the overall antenna scaling for arbitrary  $\Delta b$  is the cumulative product of elementary, unitary-step scalings. Therefore, it is possible to degrade the ADC resolution from an arbitrary  $b_{\text{init}}$  to  $b_{\text{final}} = 1$  and still maintain the same performance - provided that the number of antennas is scaled accordingly - indicating that with MaMI, using 1-bit ADCs is feasible. This particular feature of MaMI systems is analogous to traditional temporal domain oversample-and-filter systems which also enable the use of very coarse quantization, but with an interesting distinction: in MaMI, oversampling and filtering is performed in *spatial domain*.

The process of calculating  $\eta_2/\eta_1$  for the case  $K = \text{const.}$  differs between systems using ZF and MRC. When ZF is used, pre- and post-scaling sumrates are the same; on the other hand, with MRC, the scaling affects the effective interuser interference, so sumrates before and after the scaling will not be the

same. To simplify the calculation somewhat, we assume that at the onset of scaling,  $\tilde{p}_q = p_n$ . From the system design perspective, this means that if one of the sources of noise (thermal or quantization) is dominant, there is no reason for the other source to have a smaller impact since performance will be limited anyway. With additional employment of the usual large- $M$  assumption, (40) for the case of constant  $K$  becomes

$$\frac{\eta_2^{\text{MRC, ZF}}}{\eta_1} = \left\{ \frac{\log_2 \left[ 1 + \frac{M_1 SNR_1}{\rho_M SNR_1 (K-1) + 2} \right]}{\log_2 \left[ 1 + \frac{M_1 SNR_1}{SNR_1 (K-1) + 2} \right]} \right\}^{\mathbb{I}_{\text{MRC}}} \frac{1}{\rho_M} \frac{1 + \alpha_1}{\frac{P_{\text{ADC}2}}{P_{\text{ADC}1}} + \alpha_1 \rho_M^\xi}. \quad (47)$$

The analysis of energy efficiency ratio in this case is challenging because it involves 6 and 7 free parameters for ZF and MRC, respectively. Parameters  $b_1$ ,  $\Delta b$ ,  $K$  and  $SNR_1$  are first used to determine the antenna scaling  $\rho_M$  from (45); thereafter,  $\alpha_1$  and  $\xi$  (and  $M_1$  in the case of MRC) are used to completely determine the behaviour of  $\eta$  as per (47). In order to enable successful illustration of the impact of different parameters on the behavior of  $\eta$ , the parameters have been divided into two groups: “secondary” parameters ( $K$ ,  $SNR_1$ ,  $\Delta b$  and  $M_1$ ) are given typical fixed values, while the primary ones ( $b_1$ ,  $\alpha_1$  and  $\xi$ ) are swept.



**Figure 11:** Ratio of energy efficiency values after and before the antenna scaling for MRC and ZF processing, in the case  $K = \text{const}$ . Secondary parameter values:  $K = 10$ ,  $SNR_1 = 10$  dB,  $\Delta b = -1$ ,  $M_1 = 100$ .

**Numerical results** are shown in Fig. 11, where only  $P_{\text{ADC}}^{\text{th}}$  has been used.

The results reveal a strong influence of how the power consumption of other blocks relates in magnitude to ADCs (reflected through  $\alpha_1$ ), as well as how this power consumption scales with  $M$  (represented by exponent  $\xi$ ). Some energy efficiency gains are possible only in the case when  $\xi = -1.5$ . In effect, this means that  $P_{\text{rest}}$  needs to scale as  $\rho_M^{-1/2}$  if we want to improve energy efficiency by degrading the ADC resolution and simultaneously scaling up the antennas. For other tested values of scaling exponent, no gains in energy efficiency are possible when the system scales. Furthermore, as the power consumption of the ADCs becomes more prominent (lower  $\alpha_1$ ), the gains in energy efficiency with scaling become smaller, or even disappear when ZF is used.

**Discussion and main takeaways:** overall, we can conclude that the up-scaling of  $M$  and simultaneous degradation of ADC resolutions yield gains in energy efficiency only if the power consumption of the rest of the hardware scales down fast enough. Another deciding factor is the relation of power consumption of other blocks to  $P_{\text{ADC}}$ . Interestingly, in the case when  $K$  grows linearly with  $M$ , changes of  $\eta$  with system scaling are independent of bit resolution.

## 6 Conclusion

A parameterized analysis of the relation between ADC resolution and the uplink energy efficiency in a MaMI system has been performed. In one characteristic use case, we assume that upscaling of antennas is directly coupled to the degradation of ADC resolution and analyze whether reducing the number of bits will improve overall energy efficiency. The answer is affirmative only in the case when the number of users is kept constant during the process of scaling, quality of other blocks is also degraded and their power consumption scales down (at a particular rate) with the number of antennas. In another characteristic use case, we decouple the increase in the number of antennas and degradation of ADCs and observe which bit resolutions maximize the overall energy efficiency, and how these optimal bit resolutions behave with the number of antennas. The results show that the condition for energy-optimal bit resolutions to decrease as antennas are scaled up is that the number of users remains constant. Moreover, in this use case and for practical values of the most important system parameters, intermediate ADC resolutions (4 - 8) maximize energy efficiency. On a practical hardware design note, these values will increase if out-of-band interference is present in the system, with approximately 1 bit added for every 10 dB increase of OOB interference. The overall conclusion of the work is that using ADCs with intermediate bit resolutions is an optimal strategy from the energy efficiency point of view, and, except in some special cases, this strategy

does not change when antennas are scaled up.

## Acknowledgement

This work has been funded by the Swedish Foundation for Strategic Research (SSF) under the Digitally Assisted Radio Evolution (DARE) project and by European Commission under the Massive MIMO for Efficient Transmission (MAMMOET) project, and the authors would like to thank the funders. We would additionally like to thank Ellen Turner at SOL, LU and Jennifer Löfgreen at Genombrottet, LU for their feedback on the structure and organization of the text.



# Bibliography

- [1] T. L. Marzetta, “Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas,” in *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590-3600, November 2010.
- [2] H. Q. Ngo, E. G. Larsson and T. L. Marzetta, “Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems,” in *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436-1449, April 2013.
- [3] F. Rusek et al., “Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays,” in *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [4] U. Gustavsson et al., “On the impact of hardware impairments on massive MIMO,” 2014 *IEEE Globecom Workshops*, Austin, TX, 2014, pp. 294-300.
- [5] E. Björnson et al., “Massive MIMO Systems With Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits,” in *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7112-7139, Nov. 2014.
- [6] E. Björnson, M. Matthaiou and M. Debbah, “Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design,” in *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353-4368, Aug. 2015.
- [7] C. Risi, D. Persson, and E. G. Larsson. “Massive MIMO With 1-Bit ADC,” Apr. 2014 [Online]. Available: <http://arxiv.org/abs/1404.7736>
- [8] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson and C. Studer, “One-bit massive MIMO: Channel estimation and high-order modulations,” 2015 *IEEE International Conference on Communication Workshop (ICCW)*, London, 2015, pp. 1304-1309.

- [9] S. Jacobsson et al., “Throughput analysis of massive MIMO uplink with low resolution ADCs,” [Online]. Available: <https://arxiv.org/abs/1602.01139>
- [10] C. Mollén et al., “Uplink Performance of Wideband Massive MIMO With One-Bit ADCs,” in *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 87-100, Jan. 2017.
- [11] C. Studer and G. Durisi, “Quantized Massive MU-MIMO-OFDM Uplink,” in *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2387-2399, June 2016.
- [12] N. Liang and W. Zhang, “Mixed-ADC Massive MIMO,” in *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 983-997, April 2016.
- [13] L. Fan, S. Jin, C. K. Wen and H. Zhang, “Uplink Achievable Rate for Massive MIMO Systems With Low-Resolution ADC,” in *IEEE Communications Letters*, vol. 19, no. 12, pp. 2186-2189, Dec. 2015.
- [14] Q. Bai and J. A. Nossek, “Energy efficiency maximization for 5G multi-antenna receivers,” *Transactions on Emerging Telecommunication Technologies*, vol. 26, no. 1, pp. 3–14, Jan. 2015.
- [15] D. Verenzuela, E. Björnson and M. Matthaiou, “Hardware design and optimal ADC resolution for uplink massive MIMO systems,” 2016 *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Rio de Janeiro, 2016, pp. 1-5.
- [16] M. Sarajlić, L. Liu and O. Edfors, “An Energy Efficiency Perspective on Massive MIMO Quantization,” 2016 *50th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, 2016, pp. 473-478.
- [17] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA, USA: Kluwer, 1992.
- [18] A. Sripad and D. Snyder, “A necessary and sufficient condition for quantization errors to be uniform and white,” in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 25, no. 5, pp. 442-448, Oct. 1977.
- [19] M. Biguesh and A. B. Gershman, “Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals,” in *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 884-893, March 2006.

- [20] C. Wang et al., “On the Performance of the MIMO Zero-Forcing Receiver in the Presence of Channel Estimation Error,” in *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 805-810, March 2007.
- [21] B. Le et al., “Analog-to-digital converters,” in *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 69-77, Nov. 2005.
- [22] C. Svensson, “Towards Power Centric Analog Design,” in *IEEE Circuits and Systems Magazine*, vol. 15, no. 3, pp. 44-51, third quarter 2015.
- [23] B. Murmann, “Trends in low-power, digitally assisted A/D conversion,” *IEICE Trans. Electron.*, vol. E93-C, no. 6, pp. 718–727, 2010.
- [24] T. Sundström, B. Murmann and C. Svensson, “Power Dissipation Bounds for High-Speed Nyquist Analog-to-Digital Converters,” in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, no. 3, pp. 509-518, March 2009.
- [25] Dai Zhang, C. Svensson and A. Alvandpour, “Power consumption bounds for SAR ADCs,” 20th European Conference on Circuit Theory and Design (ECCTD), Linköping, 2011., pp. 556-559.
- [26] B. Murmann. ADC performance survey 1997–2016. [Online]. Available: <http://www.stanford.edu/murmann/adcsurvey.html>
- [27] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.
- [28] C. Mollén et al., “Achievable Uplink Rates for Massive MIMO With Coarse Quantization,” Nov. 2016. [Online]. Available: <http://arxiv.org/abs/1611.05723>
- [29] C. Desset, B. Debaillie and F. Louagie, “Modeling the hardware power consumption of large scale antenna systems,” 2014 IEEE Online Conference on Green Communications (OnlineGreenComm), Tucson, AZ, 2014, pp. 1-6.





## *Paper III*



# Modified Forced Convergence Decoding of LDPC Codes with Optimized Decoder Parameters

Reducing the complexity of decoding algorithms for LDPC codes is an important prerequisite for their practical implementation. In this work we propose a reduction of computational complexity targeting the highly reliable codeword bits and show that this approach can be seamlessly merged with the forced convergence scheme. We also show how the minimum achievable complexity of the resulting scheme for given performance constraints can be found by solving a constrained optimization problem, and successfully apply a gradient-descent based stochastic approximation (SA) method for solving this problem. The proposed methods are tested on LDPC codes from the IEEE 802.11n standard. Computational complexity reduction of 55% and a 75% reduction of memory access have been observed.

---

©2015 IEEE. Reprinted, with permission, from  
Muris Sarajlić, Liang Liu and Ove Edfors,  
“Modified Forced Convergence Decoding of LDPC Codes with Optimized Decoder  
Parameters,”  
in *Proceedings of the 2015 IEEE 26th Annual International Symposium on Personal,  
Indoor, and Mobile Radio Communications (PIMRC)*, Hong Kong, 2015, pp. 440-  
445.



## 1 Introduction

Low Density Parity Check (LDPC) codes, introduced in [1] have recently become a part of a number of communication standards, like WiMAX (IEEE 802.16e), IEEE 802.11n, 802.11ad and other [2]. The primary benefit of LDPC codes lies in their excellent error correcting properties that allow the systems using them to approach the information transmission capacity of the communication channel.

High complexity of the original decoding algorithm for LDPC codes, the iterative belief-propagation/sum-product (BP/SP) algorithm [1] has driven continuous research efforts targeting the reduction of its complexity while keeping the ensuing performance degradation at a tolerable level. Basic complexity reduction schemes [3] use mathematical approximations of the functions of the original algorithm. In a number of *early termination* schemes (e.g. [4]), the iterations stop as soon as some *stopping criterion* is met. Finally, a per-bit early termination scheme, referred to as *forced convergence* [5], employs a per-bit stopping criterion for the individual codeword bits.

In [6] it has been shown how the tunability of certain parameters of reduced-complexity decoding can be exploited to find the values of the parameters that minimize the decoding complexity while maintaining satisfactory performance, for given channel conditions. This paper is the continuation of work presented in [6] with:

1. A modification of the original forced convergence (FC) algorithm that yields a larger complexity reduction with the same performance degradation as the original FC;
2. The use of an iterative gradient descent-like algorithm that tries to find the optimum values of the decoder parameters for given channel conditions.

## 2 Background

### 2.1 General considerations

LDPC codes [1] are linear block codes with codeword length  $N$  described by a sparse parity check matrix  $\mathbf{H}$  with dimensions  $M \times N$ . Their structure can be represented in the form of a bipartite graph [7] where codeword bits are represented by bit (variable) nodes and parity checks by parity (check) nodes, with the interconnections between variable and check nodes mapped from the parity check matrix.

The decoding process for LDPC codes can be viewed as an iterative exchange of messages between adjacent variable and check nodes. In this work, a low-complexity approximation of the BP/SP algorithm, the offset min-sum (OMS) algorithm [3] is used in the analysis.

The order in which messages are exchanged between variable nodes and check nodes (message passing scheduling) has a direct impact on both the performance of the decoding algorithm and on the complexity of its implementation. In the so-called *layered* scheduling scheme [4], [8], which is used in the decoding algorithm in this work, c-nodes and their adjacent v-nodes are grouped in *layers*, and the exchange of messages between v-nodes and c-nodes is done for each layer separately, in a sequential fashion.

## 2.2 Forced convergence: theoretic background

Each v-node has an associated a posteriori LLR value, commonly denoted as  $Q_v$ . The sign of  $Q_v$  maps to bit values 0 and 1. The magnitude of  $Q_v$  corresponds to the amount of “confidence” that the v-node has in its sign. As the iterations of the decoding algorithm progress it can be observed how (for SNRs after the “turbo cliff”) the magnitudes of  $Q_v$  evolve towards  $+\infty$  or  $-\infty$ . This indicates that the nodes become increasingly confident that they are a 0 or a 1 as iterations progress.

It can then be reasonable to stop the updating of  $Q_v$  for the very confident nodes, i.e. the nodes for which the magnitude of  $Q_v$  crosses some predefined threshold  $\theta$ . Value of  $Q_v$  is therefore held at some fixed  $Q_{v, frozen}$  for the remainder of the decoding process. This is referred to as *forced convergence*. Depending on the value of  $\theta$ , forced convergence will result in certain performance degradation, but will also introduce a complexity reduction. By tuning the value of  $\theta$ , the performance can therefore be finely traded for complexity.

## 3 Modified offset min-sum algorithm with layered scheduling and forced convergence

In this work, a modification of the original layered OMS algorithm with FC is introduced. To explain the nature of the modification, a concise overview of the message passing activities in the original algorithm will be given first.

One layer is assumed to be the group of v-nodes connected to one c-node. The set of v-nodes connected to c-node  $c$  is denoted by  $N(c)$ . Then, at iteration  $i$ , for each layer  $c$  and each v-node  $v$  belonging to this layer, the following three operations are performed:

1.  $Q_{temp,vc}$  calculation:  $Q_{temp,vc}^{(i)} = Q_v^{(i)} - R_{cv}^{(i-1)}$
2.  $R_{cv}$  calculation:

$$R_{cv}^{(i)} = \left[ \prod_{v' \in N(c) \setminus v} \text{sign}(Q_{temp,v'c}^{(i)}) \right] \times \max \left\{ \min_{v' \in N(c) \setminus v} |Q_{temp,v'c}^{(i)}| - \omega, 0 \right\}$$

3.  $Q_v$  update:  $Q_v^{(i)} = Q_{temp,vc}^{(i)} + R_{cv}^{(i)}$

The  $Q_{temp,vc}$  can be seen as the message in which  $v$  informs  $c$  about its own sign and how confident it is that this is the actual value of its sign. On the other hand,  $R_{cv}$  is the total knowledge that other v-nodes in the layer have about the sign of  $v$ . It can be seen from the above expressions that the value of  $R_{cv}$  is influenced by  $Q_{temp}$  values in the layer that have the smallest magnitudes. V-nodes that are strongly convinced about their sign will send “strong”  $Q_{temp}$  messages of large magnitude that will not influence the value of  $R_{cv}$ . Therefore, the  $Q_{temp}$  messages for these “confident” nodes can be approximated by some constant value which is large enough.

Since the forced convergence approach also targets the “confident” v-nodes, it is natural to combine it together with the  $Q_{temp}$  approximation that was just described. The modified  $Q_{temp}$  calculation rule can then be formulated as follows:

$$Q_{temp,vc}^{(i)} = \begin{cases} Q_{v,frozen} & \text{if } v \text{ is frozen} \\ Q_v^{(i)} - R_{cv}^{(i-1)} & \text{otherwise} \end{cases} \quad (1)$$

The complete pseudocode formulation of the modified algorithm is given in Algorithm I. The completed notation is as follows:  $P_v$  are the a priori v-node LLRs (obtained from symbols received from the channel),  $\omega$  is the offset value, a standard part of the offset min-sum algorithm.  $X(\cdot)$  is a hard bit decision operator (converting  $Q_v$  to 0 or 1),  $\mathbf{Q} = [Q_1 \ Q_2 \ \dots \ Q_N]$  is the codeword LLR vector and  $\text{GF2}\{\cdot\}$  denotes operations in Galois field over  $\{0, 1\}$ .

## 4 Optimizing the LDPC decoder parameters

### 4.1 Problem formulation

It has been shown in [6] that, for a general vector  $\boldsymbol{\rho}$  of environment settings (such as SNR or fading properties) and minimum performance requirement  $FER_c$  (expressed in terms of the frame error rate FER), the value of the



**Algorithm 1** Layered OMS with FC and extrinsic message simplification

---

```

1: for all v-nodes  $v$  and c-nodes  $c$  do                                ▷ Initialization
2:    $R_{cv} \leftarrow 0$ 
3:    $Q_v \leftarrow P_v$ 
4: end for
5:  $Inact = \emptyset$ 
6: for iterations  $i$  to  $I_{max}$  do
7:   for all  $c$  do
8:     for all  $v \in N(c)$  do                                          ▷  $Q_{temp}$  calculation
9:       if  $v \in Inact$  then
10:         $Q_{temp} \leftarrow Q_v$ 
11:       else
12:         $Q_{temp} \leftarrow Q_v - R_{cv}$ 
13:       end if
14:     end for
15:      $Q_{min1} \leftarrow \min_{v \in N(c)} \{|Q_{temp,vc}|\}$                     ▷  $R_{cv}$  calculation
16:      $Q_{min2} \leftarrow \min_{\substack{v \in N(c), \\ v \neq v_{min1}}} \{|Q_{temp,vc}|\}$ 
17:      $Q_{min1} \leftarrow \max\{Q_{min1} - \omega, 0\}$ 
18:      $Q_{min2} \leftarrow \max\{Q_{min2} - \omega, 0\}$ 
19:      $S = \prod_{v \in N(c)} \text{sign}(Q_{temp,vc})$ 
20:     for all  $v \in N(c)$  AND  $v \notin Inact$  do
21:       if  $v = v_{min1}$  then
22:         $R_{cv} \leftarrow \text{sign}(Q_{vc}) \cdot S \cdot Q_{min2}$ 
23:       else
24:         $R_{cv} \leftarrow \text{sign}(Q_{vc}) \cdot S \cdot Q_{min1}$ 
25:       end if
26:        $Q_v \leftarrow Q_{temp,vc} + R_{cv}$                                 ▷  $Q_v$  update and thresholding
27:       if  $|Q_v| > \theta$  then
28:         $|Q_v| \leftarrow \theta$ 
29:         $v \in Inact$ 
30:       end if
31:     end for
32:   end for
33:   if  $\text{GF2}\{H \cdot X(Q^T)\} = \mathbf{0}$  then
34:     stop iterations
35:   end if
36: end for

```

---

threshold  $\theta$  that minimizes the computational complexity  $C(\theta, \boldsymbol{\rho})$  of the decoding algorithm is found by solving the general optimization problem

$$\begin{aligned} & \underset{\theta}{\text{minimize}} && C(\theta, \boldsymbol{\rho}) \\ & \text{subject to} && \text{FER}(\theta, \boldsymbol{\rho}) \leq \text{FER}_c \end{aligned} \quad (2)$$

In order to achieve optimum complexity reduction, the optimum value of  $\theta$  should be found and applied for any value of the current environment settings. Put in simple terms, it should be adapted to the channel.

This work uses an analytical model of the computational complexity that is drawn from the algorithm structure. In the derivation of the model, tilde ( $\sim$ ) will be used to denote random terms. Layers will be indexed by  $l$ , decoding algorithm iterations by  $i$  and individual decoded blocks (different runs of the decoding algorithm) by  $b$ . The number of active v-nodes (nodes that have not yet been frozen) in layer  $l$  (and at iteration  $i$  and block  $b$ ) is denoted by  $\tilde{n}_a^{(b,i,l)}$ . The total number of v-nodes adjacent to the c-node in layer  $l$  and the total number of c-nodes in the code (both deterministic and following from the code construction) are denoted by  $n^{(l)}$  and  $|c|$ , respectively. Number of iterations performed in the decoding of block  $b$  is denoted by  $\tilde{I}^{(b)}$ .

It should be pointed out that the number of active nodes  $\tilde{n}_a^{(b,i,l)}$  and the number of iterations  $\tilde{I}^{(b)}$  are *discrete random variables*; the randomness of  $\tilde{n}_a^{(b,i,l)}$  is the consequence of applying FC, and  $\tilde{I}^{(b)}$  is random due to the early termination (parity check at the end of each iteration). Probability mass functions  $f_N(\tilde{n}_a^{(b,i,l)}; \theta, \boldsymbol{\rho})$  and  $f_I(\tilde{I}^{(b)}; \theta, \boldsymbol{\rho})$  are parameterized by  $\theta$  and  $\boldsymbol{\rho}$ . Owing to the inherent complexity of the LDPC code structure and the nonlinearity of the decoding algorithm, these pmfs are in general case extremely hard (if not impossible) to obtain in closed form.

As in [6], the complexity is given in the number of additions (assumed equivalent in complexity as comparisons) performed per decoded block. Complexity analysis of the decoding algorithm is based on the complexity analysis for a single layer  $l$ :

- Complexity of the  $Q_{temp}$  calculation section in one layer is

$$\tilde{C}_{Q_{temp}}^{(b,i,l)} = \tilde{n}_a^{(b,i,l)} \quad (3)$$

- Complexity of the  $R_{cv}$  calculation section depends on the number of different  $|Q_{temp,vc}|$  values among which the two minimum elements are chosen. This number is denoted by  $\tilde{n}_x^{(b,i,l)}$ . The set of values of  $|Q_{temp,vc}|$  from which the two minimum elements are picked is formed by  $|Q_{temp,vc}|$

from active nodes, and a single  $\theta$  value representing all the frozen nodes. Therefore

$$\tilde{n}_x^{(b,i,l)} = \min\{\tilde{n}_a^{(b,i,l)} + 1, n^{(l)}\} \quad (4)$$

and the complexity of this section is

$$\tilde{C}_{R_{cv}}^{(b,i,l)} = \tilde{n}_x^{(b,i,l)} + \lceil \log_2 \tilde{n}_x^{(b,i,l)} \rceil + 2, \quad (5)$$

derived from the optimum complexity of finding two minimum elements in an unsorted array [9] and the four additions in lines 17 and 18.

- Finally, the complexity of the  $Q_v$  update and thresholding section is

$$\tilde{C}_{Q_v}^{(b,i,l)} = 2\tilde{n}_a^{(b,i,l)} \quad (6)$$

Total complexity of decoding one layer is

$$\tilde{C}^{(b,i,l)} = 3\tilde{n}_a^{(b,i,l)} + \tilde{n}_x^{(b,i,l)} + \lceil \log_2 \tilde{n}_x^{(b,i,l)} \rceil + 2, \quad (7)$$

and the complexity of decoding one block is then

$$\tilde{C}^{(b)} = \sum_{i=1}^{\tilde{I}^{(b)}} \sum_{l=1}^{|c|} \tilde{C}^{(b,i,l)} \quad (8)$$

Finally, a sample mean of  $\tilde{C}^{(b)}$  over a window of  $W$  blocks is taken as an estimate of the complexity  $C(\theta, \boldsymbol{\rho})$ :

$$\hat{C}(\theta, \boldsymbol{\rho}) = \tilde{C} = \frac{1}{W} \sum_{b=1}^W \tilde{C}^{(b)} = \frac{1}{W} \sum_{b=1}^W \sum_{i=1}^{\tilde{I}^{(b)}} \sum_{l=1}^{|c|} \tilde{C}^{(b,i,l)} \quad (9)$$

If environment conditions  $\boldsymbol{\rho}$  are assumed constant over  $W$  blocks and if additionally there is no dependence between noise or decoded data between different blocks,  $\tilde{C}^{(b)}$  can be assumed to be an i.i.d. random variable coming from an unknown discrete pmf. Then, from the central limit theorem it follows that the distribution of  $\hat{C}(\theta, \boldsymbol{\rho})$  is approximately

$$\mathcal{N}(C(\theta, \boldsymbol{\rho}), \sigma_C^2(\theta, \boldsymbol{\rho})), \quad (10)$$

with  $C(\theta, \boldsymbol{\rho}) = \mathbb{E}[\hat{C}(\theta, \boldsymbol{\rho})]$ . Note that  $C(\theta, \boldsymbol{\rho})$  is not available in closed form; it is only possible to obtain its (noisy) estimate  $\hat{C}(\theta, \boldsymbol{\rho})$ .

In order to solve the optimization problem (2),  $\text{FER}(\theta, \boldsymbol{\rho})$  needs to be obtained as well. Similar to  $C(\theta, \boldsymbol{\rho})$ ,  $\text{FER}(\theta, \boldsymbol{\rho})$  is not known in closed form and

has to be estimated. This can be done in the usual way of counting block errors over a window of  $W$  blocks and then dividing by  $W$ . Formally,

$$\widehat{\text{FER}}(\theta, \boldsymbol{\rho}) = \frac{1}{W} \sum_{b=1}^W \mathbb{1}_{err}^{(b)}, \quad (11)$$

where  $\mathbb{1}_{err}^{(b)}$  is an indicator function equal to 1 when block  $b$  is in error, and 0 otherwise. Values of the indicator function are Bernoulli distributed, and it is well known [10] that for a large enough  $W$ ,  $\widehat{\text{FER}}(\theta, \boldsymbol{\rho})$  is approximately distributed as

$$\mathcal{N}\left(\text{FER}(\theta, \boldsymbol{\rho}), \frac{\text{FER}(\theta, \boldsymbol{\rho})(1 - \text{FER}(\theta, \boldsymbol{\rho}))}{W}\right) \quad (12)$$

We can therefore conclude that, instead of the cost and constraint functions from (2), in practice we can only obtain their noisy estimates:

$$\hat{C}(\theta, \boldsymbol{\rho}) = C(\theta, \boldsymbol{\rho}) + \eta, \quad (13)$$

$$\widehat{\text{FER}}(\theta, \boldsymbol{\rho}) = \text{FER}(\theta, \boldsymbol{\rho}) + \epsilon, \quad (14)$$

where, following from (10) and (12),  $\eta$  and  $\epsilon$  are approximately zero-mean Gaussian with pdf parameterized by  $\boldsymbol{\rho}$  and  $\theta$ .

Optimization of  $\theta$  is then performed using the noisy function estimates and is formulated as

$$\begin{aligned} & \underset{\theta}{\text{minimize}} && \hat{C}(\theta, \boldsymbol{\rho}) \\ & \text{subject to} && \widehat{\text{FER}}(\theta, \boldsymbol{\rho}) \leq \phi_c, \end{aligned} \quad (15)$$

with  $\phi_c$  being the new value of the constraint that accounts for the random nature of  $\widehat{\text{FER}}$  and introduces a “safety margin”.

## 4.2 Problem solution

A family of optimization methods, known collectively as *stochastic approximation methods* is known to be applicable to optimization problems in which the cost function is not known and can only be observed through its noisy estimates (measurements), like in (15). The first stochastic approximation (SA) method was proposed by Kiefer and Wolfowitz in [11] and has been followed by a host of similar methods (e.g. simultaneous perturbation SA by Spall, [12]).

SA methods are based on the classic gradient descent algorithm, in which a starting point is chosen and the optimum is approached iteratively by following the direction of the negative gradient. The difference between the deterministic

gradient descent and SA is that SA uses a noisy estimate of the gradient instead of its actual value.

For the constrained problem (15), the iterates have to be confined to the set of feasible points; this is modeled by a projection operator  $\Pi_{\Theta}$  that projects the iterates back onto the feasible set  $\Theta$ . The recursive expression for SA with feasible set projection, applied to the optimization problem (15) is given by

$$\theta_{k+1} = \Pi_{\Theta} \left\{ \theta_k - a_k \frac{\hat{\partial}}{\partial \theta} C(\theta_k, \boldsymbol{\rho}) \right\}, \quad (16)$$

with the “gradient estimate”  $\frac{\hat{\partial}}{\partial \theta} C(\theta_k, \boldsymbol{\rho})$  calculated as

$$\frac{\hat{\partial}}{\partial \theta} C(\theta_k, \boldsymbol{\rho}) = \frac{\hat{C}(\theta_k + c_k, \boldsymbol{\rho}) - \hat{C}(\theta_k - c_k, \boldsymbol{\rho})}{2c_k} \quad (17)$$

The SA-based iterative algorithm with feasible set projection for estimating  $\theta^*$  that solves (15) at a given environment setting  $\boldsymbol{\rho}$  is given by Algorithm II.

---

**Algorithm 2** Stochastic approximation with feasible set projection

---

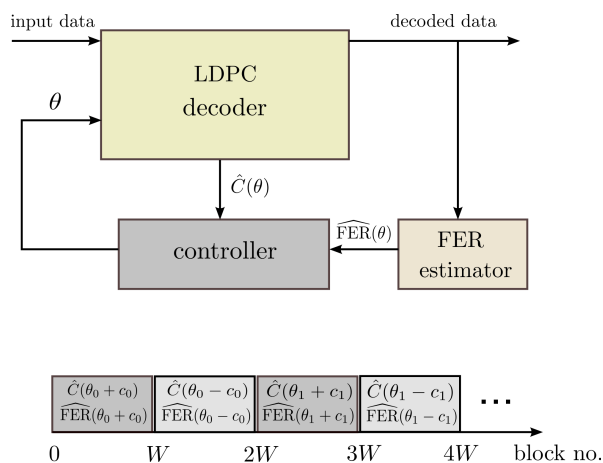
- 1: Initialize  $\theta_0$
  - 2: **for**  $k$  from 0 to  $I_{max} - 1$  **do**
  - 3:    $a_k = \frac{a}{(k+1)^\alpha}$ ,    $c_k = \frac{c}{(k+1)^\gamma}$
  - 4:   **if**  $(\theta_k - c_k) < 0$  OR  $\widehat{\text{FER}}(\theta_k - c_k, \boldsymbol{\rho}) > \phi_c$  **then**
  - 5:     stop iterations
  - 6:   **end if**
  - 7:    $\frac{\hat{\partial}}{\partial \theta} C(\theta_k, \boldsymbol{\rho}) = \frac{\hat{C}(\theta_k + c_k, \boldsymbol{\rho}) - \hat{C}(\theta_k - c_k, \boldsymbol{\rho})}{2c_k}$
  - 8:    $\theta_{k+1} = \theta_k - a_k \frac{\hat{\partial}}{\partial \theta} C(\theta_k, \boldsymbol{\rho})$
  - 9: **end for**
  - 10: **if**  $\widehat{\text{FER}}(\theta_k, \boldsymbol{\rho}) \leq \phi_c$  AND  $\theta_k \geq 0$  **then**
  - 11:    $\hat{\theta}^* = \theta_k$
  - 12: **else**
  - 13:    $\hat{\theta}^* = \theta_{k-1}$
  - 14: **end if**
- 

Some practical information regarding the optimization algorithm:

- Feasible set  $\Theta$  is defined as

$$\Theta = \left\{ \theta \geq 0 \mid \widehat{\text{FER}}(\theta) \leq \phi_c \right\} \quad (18)$$

Negative values of  $\theta$  produce undefined behaviour of the OMS-FC algorithm, hence the nonnegative constraint imposed on  $\theta$ .



**Figure 1:** System diagram and output timeline

- Projection  $\Pi_\theta$  is implemented in lines 4-6 and 10-14 of the algorithm.
- Initial point  $\theta_0$  and the finite difference step  $\theta_0 - c_0$  are considered to be in  $\Theta$ .
- Following the practical advice given in [13], the parameters  $\alpha$ ,  $\gamma$ ,  $a$  and  $c$  of the sequences  $a_k$  and  $c_k$  are chosen as follows:
  - $\alpha = 0.602$ ,  $\gamma = 0.101$
  - At the first iteration,  $a$  is set to  $\frac{\Delta\theta_0}{\left| \frac{\partial}{\partial \theta} C(\theta_0, \rho) \right|}$  where  $\Delta\theta_0$  is the desired step in the first iteration.
  - Value of  $c$  is set to the estimated value of the standard deviation of  $\hat{C}(\theta, \rho)$ .
- The new value of the constraint  $\phi_c$  is determined from confidence intervals for  $\widehat{\text{FER}}(\theta, \rho)$  from property (12). It is the value that, when chosen as the constraint, guarantees with a certain probability that the *actual*  $\text{FER}(\theta, \rho)$  will be smaller or equal than the original constraint  $\text{FER}_c$ .

Fig. 1 shows the block diagram of the complete system for the optimization of LDPC decoder parameters, together with the timeline of the decoder outputs (naturally, averaged over the entire duration of one slot). From this diagram it is evident that this is a “black box” optimization method in which a controller unit, implementing Algorithm II, chooses the inputs to the system, estimates

the gradient of the cost function from the observed system outputs and decides on the new input values based on the gradient estimate.

## 5 Simulation and results

The described reduced complexity decoding algorithm and the optimization algorithm were tested in three different setups, based on three different LDPC codes from the IEEE 802.11n standard [2], with code rates and blocklengths given in Table 1.

**Table 1:** LDPC codes used in the simulations

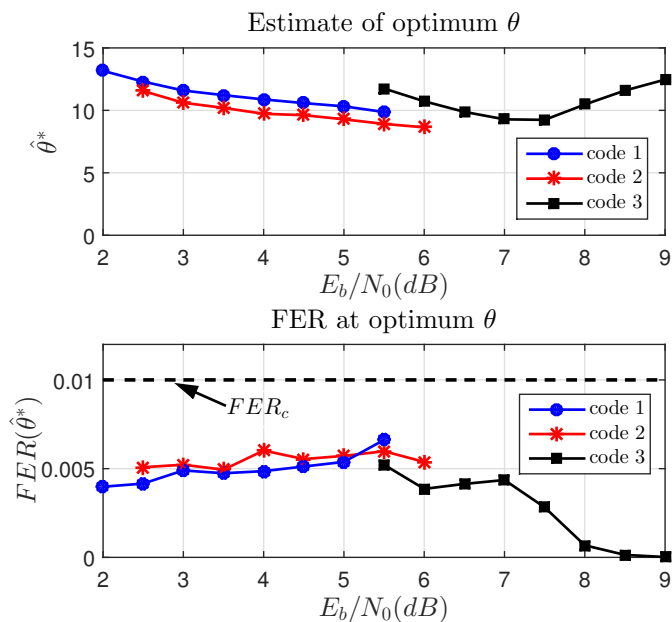
Code 1	R = 1/2, N = 1944
Code 2	R = 1/2, N = 648
Code 3	R = 3/4, N = 648

The selected channel is AWGN and the modulation for all three setups is QPSK. The rates of the codes determine their operational SNR ranges. Codes 1 and 2 are suitable for use at low SNRs, whereas code 3 is better suited for use in the mid-SNR range. Different values of SNR are chosen as different states of the environment setting  $\rho$ . Three decoders are compared in terms of complexity:

1. The plain OMS decoder, without FC
2. A “lazy” (that is, non-adaptive) OMS FC decoder with extrinsic message simplification that uses one value of  $\theta$  over the entire tested range of SNR values. This value of  $\theta$  is selected as minimum  $\theta$  that satisfies the performance requirement for all SNRs, while yielding a complexity reduction compared to plain OMS.
3. An “optimized” OMS FC decoder with extrinsic message simplification using the estimated optimum value of the threshold,  $\hat{\theta}^*$  (provided by the SA algorithm) at each SNR point.

Averaging window length  $W$  in the optimization is set to 1000. At each SNR point, optimization is run for 10 different random seeds, and the final value is the sample mean of the results obtained from these different runs. Performance constraint  $\text{FER}_c$  is set to  $10^{-2}$ , and for  $W = 1000$  and  $Pr(\widehat{\text{FER}} \leq \text{FER}_c) = 0.95$  this translates to  $\phi_c = 5.4 \cdot 10^{-3}$ .

In Fig. 2 the values of  $\hat{\theta}^*$  at each SNR are given together with  $\text{FER}(\hat{\theta}^*)$  (averaged over 50 000 blocks and therefore considered the “true” value). The

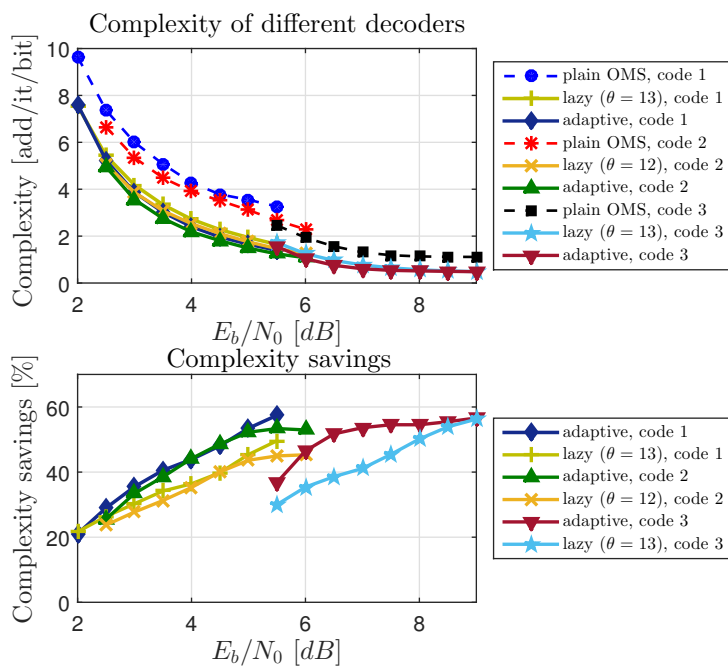


**Figure 2:** Estimate of optimum  $\theta$  and FER at  $\hat{\theta}^*$

obtained values of FER confirm that the performance constraint  $\text{FER} \leq 10^{-2}$  is satisfied at every  $\hat{\theta}^*$ . It was observed that, at high SNRs,  $\hat{\theta}^*$  for code 3 do not follow the same trend of decrease with SNR as in the two other codes due to a limited number of optimization iterations (set to 100). With a larger number of iterations it is possible to attain smaller values of  $\hat{\theta}^*$  at these SNRs.

The complexity, normalized by the maximum number of iterations and the number of information bits in the block, and the savings of lazy and adaptive schemes compared with the plain OMS scheme are shown in Fig. 3. Complexity  $C(\theta)$  for all three decoders is averaged over 50 000 blocks and is therefore considered the “true” value. The results lend themselves to a comparison with the results in [6], since the same code (IEEE 802.11n, R=1/2, N=648) is analyzed in both works. In [6], optimum  $\theta$  (found by a grid search) yielded maximum complexity savings of 35% compared to plain OMS for this particular code and the original OMS FC algorithm. In this work, the modified OMS FC algorithm achieves a 53% complexity reduction at the same SNR point, thereby confirming that the simple extrinsic message modification in (1) can result in significant savings of computational complexity. It can also be observed that the channel-adaptive decoder can bring in additional 5 - 12% of complexity re-

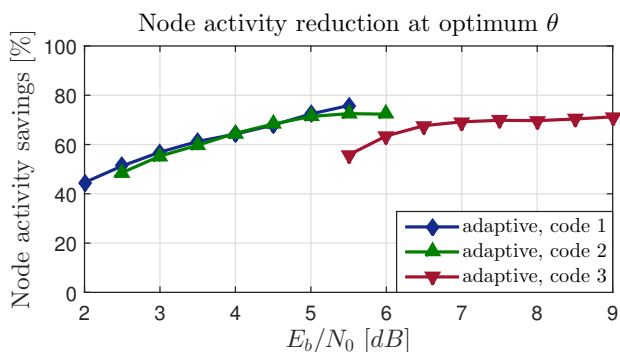




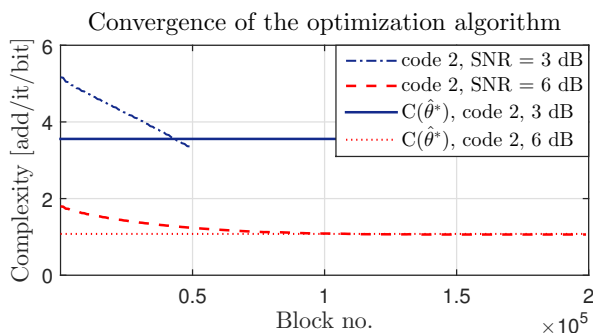
**Figure 3:** Comparison of complexities for different decoders and codes

duction compared to the “lazy” decoder, emphasizing the general notion that adapting the system to its environment is beneficial for system performance. It should be noted that the controller in Fig. 1 is of negligible complexity compared to the decoder.

In actual hardware implementations of decoders, a large part of total energy consumption is due to memory access activities [14]. It is therefore beneficial to estimate the reduction in memory access when analyzing decoding schemes with reduced complexity. Although this heavily depends on the actual implementation and memory design, some conclusions can be drawn from the structure of the algorithm. It can be identified that most memory access activity (reading/writing) will occur in “ $Q_{temp}$  calculation” and “ $Q_v$  update” sections of the decoding algorithm. Since both of these sections are performed when the  $v$ -node is active, it can then be safely assumed that the reduction in memory access will be proportional to the reduction of the number of active nodes. The node activity of the channel-adaptive OMS FC algorithm (at  $\hat{\theta}^*$ ) is therefore compared to the plain OMS and the corresponding reduction is presented in



**Figure 4:** Comparison of node activity savings for different codes with adaptive decoding



**Figure 5:** Convergence behaviour of the SA optimization algorithm at different SNRs

Fig. 4. A very high (up to 75%) decrease of v-node activity suggests that the presented complexity reduction scheme can be expected to yield a highly energy efficient hardware implementation, both in terms of computations as well as memory access.

Finally, we shortly turn to practical implications of using the SA algorithm to find the optimum  $\theta$ . Fig. 5 shows the rate of convergence of the optimization algorithm for two single runs (i.e. without averaging over different seeds) at different values of SNR. At the lower SNR value, the iterations stop after approximately 50 000 decoded blocks because the performance constraint is violated; at the higher SNR, they continue until the maximum number of iterations is exhausted, but the process can be seen to converge after around 100

000 decoded blocks. To put this into time perspective, we assume information bitrate of 100 Mbps. Given that one block of code 2 has 324 information bits, the convergence times for the two described cases are then  $\approx 0.16s$  and  $\approx 0.32s$ , respectively. That indicates that SA can be used to tune the decoder to the optimum value of  $\theta$  in real-time, provided that the channel is static or with very low mobility. The benefit of this approach lies in the fact that the optimization algorithm is of negligible complexity compared to the decoding algorithm and also in the fact that it does not need any channel information (conversely,  $\hat{\theta}^*$  produced by the algorithm implicitly contains a channel estimate).

## 6 Conclusion

This work proposes a modified rule for calculating the extrinsic messages in the LDPC decoding algorithm, in which the extrinsic messages corresponding to highly reliable bits can be simply approximated with the a posteriori LLRs, thereby reducing the computational complexity. It is additionally proposed that this modification is merged with the forced convergence scheme. It has been shown how the computational complexity of the resulting decoding algorithm can be modeled analytically, and how a gradient-descent based optimization scheme can be successfully applied to this model to find the maximum complexity reduction that the algorithm can achieve under some predefined performance constraints. The overall results show significant reduction of computational as well as memory access complexity, indicating high energy efficiency of a possible hardware implementation of the algorithm. Finally, it is shown that maximum complexity reduction is achieved if the parameters of the decoder are adapted to the environment.

## Acknowledgment

The work presented is a part of the Digitally Assisted Radio Evolution (DARE) project, and the authors would like to thank the Swedish Foundation for Strategic Research (SSF) for providing the funds for the project.

# Bibliography

- [1] R. G. Gallager, *Low-Density Parity-Check Codes*. Cambridge, MA: MIT Press, 1963.
- [2] IEEE 802.11n-2012, IEEE Standard for Information Technology-Telecommunications and information exchange between systems, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications [Online]. Available: <http://standards.ieee.org/getieee802/download/802.11-2012.pdf>
- [3] J. Chen, A. Dholakia, E. Eleftheriou, M.P.C. Fossorier, and X.Y. Hu, “Reduced-Complexity Decoding of LDPC Codes,” *IEEE Tran. Comm.*, vol.53, no.8, pp.1288-1299, Aug. 2005.
- [4] D.E. Hocevar, “A reduced complexity decoder architecture via layered decoding of LDPC codes,” *IEEE Workshop on Signal Process. Syst., 2004.*, pp.107-112, Oct. 2004.
- [5] E. Zimmermann, P. Pattisapu, P. K. Bora, and G. Fettweis, “Reduced Complexity LDPC Decoding Using Forced Convergence,” *Proc. 7th Int. Symp. on Wireless Personal Multimedia Commun.*, Padova, Italy, Sep. 2004, vol. 3, pp. 243–246, WA2-2.
- [6] M. Sarajlić, L. Liu, O. Edfors, “Reducing the Complexity of LDPC Decoding Algorithms: An Optimization-Oriented Approach”, *Personal Indoor and Mobile Radio Communications (PIMRC), 2014 IEEE 25th International Symposium on*, Sep. 2014.
- [7] R.M. Tanner, “A recursive approach to low complexity codes,” *IEEE Trans. Inf. Theory*, vol.27, no.5, pp.533-547, Sep. 1981.
- [8] E. Sharon, S. Litsyn, J. Goldberger, “An efficient message-passing schedule for LDPC decoding,” *Proc. 23rd IEEE Conv. of Elec. and Electron. Engineers in Israel, 2004.*, pp.223-226, Sept. 2004.

- 
- [9] P. V. Ramanan, L. Hyafil, “New algorithms for selection”, *J. of Algorithms*, vol. 5, no. 4, pp. 557-578, Dec. 1984.
  - [10] M. C. Jeruchim, P. Balaban, K. S. Shanmugan, *Simulation of Communication Systems*, 2nd ed. New York: Kluwer Academic Publishers, 2002.
  - [11] J. Kiefer and J. Wolfowitz, “Stochastic estimation of the maximum of a regression function”, *Ann. Math. Stat.*, vol. 23, pp. 462-466, 1952.
  - [12] J. C. Spall, “Multivariate stochastic approximation using a simultaneous perturbation gradient approximation”, *IEEE Trans. Autom. Control*, vol. 37, no. 3, pp. 332-341, 1992.
  - [13] J. C. Spall, “Implementation of the simultaneous perturbation algorithm for stochastic optimization,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 3, pp. 817-823, Jul. 1998.
  - [14] C. Studer, N. Preyss, C. Roth, and A. Burg, “Configurable high-throughput decoder architecture for quasi-cyclic LDPC codes,” in *Proc. 42nd Asilomar Conf. on Signals, Systems and Computers*, pp. 1137-1142, Oct. 2008.

*Paper IV*



# Fully Decentralized Approximate Zero-Forcing Precoding for Massive MIMO Systems

We analyze the downlink of a massive multiuser multiple input, multiple output (MIMO) system where antenna units at the base station are connected in a daisy chain without a central processing unit and only possess local channel knowledge. For this setup, we develop and analyze a linear precoding algorithm for suppressing interuser interference. It is demonstrated that the algorithm is close to zero-forcing precoding in terms of performance for a large number of antennas. Moreover, we show that with careful scheduling of processing across antennas, requirements for interconnection throughput are reduced compared with the fully centralized solution. Favorable tradeoff between performance and interconnection throughput makes the daisy chain a viable candidate topology for real-life implementations of base stations in MIMO systems where the number of antennas is very large.





## 1 Introduction

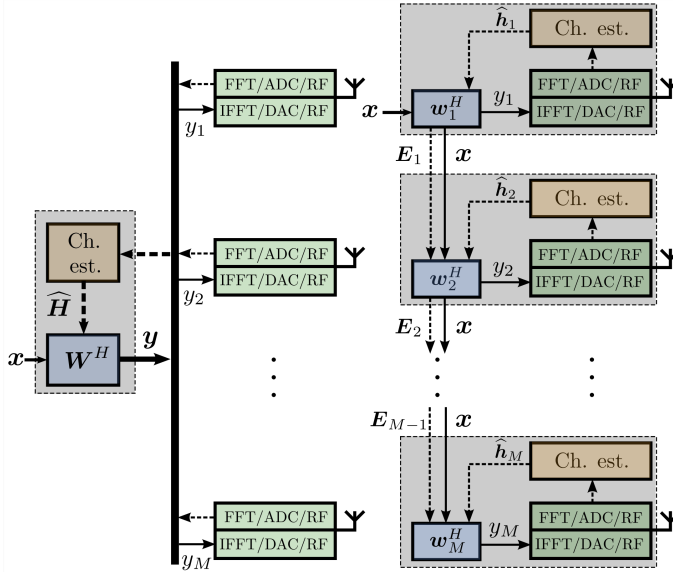
Massive multiple input, multiple output (MIMO) [1] is a technique whose benefits in multiuser mobile systems, such as superb spectral efficiency, have been confirmed by practical implementations ([2] and [3], among others). These works have, however, revealed an inherent weakness of centralized massive MIMO (MaMI) base station (BS) architectures: a prohibitively high throughput requirement for the links between the central processing unit (CU) and the remote antenna units (AUs) [4–6]. For the purpose of illustration, we assume a LTE-like multicarrier MaMI system identical to the one described in [3], with  $N_{\text{used}}$  data-carrying subcarriers and  $w$  bits representing one complex sample. In a centralized MaMI BS with  $M$  antennas, shown in the left-hand side of Fig. 1, the required throughput on the bus connecting AUs and CU is  $R_{\text{central}} = MwN_{\text{used}}/T_{\text{OFDM}}$ , where  $T_{\text{OFDM}}$  is the OFDM symbol duration. Linear scaling of  $R_{\text{central}}$  with  $M$  severely limits the scalability of the centralized system, prohibiting it from becoming truly “massive”.

A proposed remedy to this challenge is a partial or full decentralization of baseband processing in the BS. Partially decentralized structures rely on a careful division of processing tasks between the CU and AUs [5, 6], which still entails substantial shuffling of overhead data between units. On the other hand, in fully decentralized structures [2, 5, 6], CU is eliminated and AUs are joined into smaller groups that perform baseband processing independently of each other, in a parallel fashion. Finally, the structure described in [4] dispenses with a CU and has semi-independent groups of AUs exchange consensus information, with a negligible performance penalty compared to a fully centralized implementation but with the latency of information exchange limiting the throughput.

In this work, we take a fresh look on fully decentralized structures for MaMI BSs. Namely, we investigate a decentralized antenna array topology that has not been analyzed in this context before: a daisy chain of single-antenna AUs without a CU. We focus on developing an algorithm for suppressing interuser interference (IUI) that takes into account the limitations of the daisy chain topology. Additionally, we theoretically analyze the mechanisms of operation of this algorithm. The new algorithm is shown to get close in performance to ZF for a very large number of BS antennas. Moreover, with proper scheduling of calculations, the interconnect throughput between system parts can be substantially reduced compared with the centralized solution.

## 2 Problem setup and system model

We analyze the downlink (DL) of a single-cell system with  $M$  collocated BS antennas and  $K$  single-antenna users, where  $M \gg K \gg 1$  and antenna-user channels are narrowband and flat-faded. The system employs TDD transmission where channel reciprocity is assumed. Based on the uplink channel state information (CSI) estimates, the BS formulates a linear precoder which is applied to the data intended for the users and transmitted in the downlink. For simplicity of analysis and exposition, we assume the ideal case of perfect CSI and perfectly reciprocal radio channels.



**Figure 1:** MaMI array implementation diagrams. Left: centralized processing with a shared bus. Right: fully decentralized processing, daisy chain topology. Dashed links: channel estimation/precoder formulation phase, full links: precoding phase.  $\widehat{H}$  denotes channel estimate. Other variables of relevance defined in (1) and (3). Uplink transmission phase not illustrated.

In contrast to conventional BS architectures, we assume there is no central unit aggregating the CSI and formulating the precoded signal. Instead, precoding is formulated and executed *locally* at BS antennas, in a *fully decentralized* manner. More specifically, each antenna of the BS array, together with associated analog and digital hardware, is considered to form a low complexity,

cheap antenna processing block (APB), and adjacent APBs are connected using unidirectional links, forming a daisy chain. The  $m$ th APB in the chain only possesses knowledge of the *local* baseband channel between the corresponding array antenna and all the users, represented by  $\mathbf{h}_m \in \mathbb{C}^{K \times 1}$ . The local channel vector, together with side information passed from the preceding APB, is used by the  $m$ th APB to formulate the local precoder vector  $\mathbf{w}_m \in \mathbb{C}^{K \times 1}$ . Antenna units thus formulate their precoders in a sequential fashion. The data intended for users, represented by  $\mathbf{x} \in \mathbb{C}^{K \times 1}$ , is broadcast to APBs, which subsequently apply their precoders in parallel. A conceptual diagram of the described setup is given on the left hand side of Fig. 1.

The received complex baseband signal for all users, is a  $K \times 1$  vector

$$\mathbf{r} = \alpha \mathbf{H} \underbrace{\mathbf{W}^H \mathbf{x}}_{\mathbf{y}} + \mathbf{n}, \quad (1)$$

where  $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_M]$  and  $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_M]$ . It is assumed that  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_K$  and  $\mathbb{E}\{\mathbf{n}\mathbf{n}^H\} = N_0 \mathbf{I}_K$ . Factor  $\alpha$  is used for adjusting the total transmit power.

The design of the precoder  $\mathbf{W}^H$  is governed by various performance criteria. In this work, we investigate  $\mathbf{W}^H$  that will suppress IUI. Such a precoder can be found as a solution of the optimization problem

$$\underset{\mathbf{W}^H}{\text{minimize}} \quad \|\mathbf{H}\mathbf{W}^H - \mathbf{I}_K\|_F^2. \quad (2)$$

It should be noted that the pseudoinverse  $\mathbf{H}^\dagger = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1}$  is the solution of (2) with smallest Frobenius norm when  $\mathbf{H}$  has full row rank, yielding the well-known ZF solution that completely eliminates IUI. However, knowledge of the entire matrix  $\mathbf{H}$  is required for finding  $\mathbf{H}^\dagger$ , which in practical BS implementations means that  $\mathbf{H}$  must be made available to a CU. This option is ruled out in the setup we consider and approximate solutions of (2) tailored to fit the daisy chain topology need to be sought.

### 3 Fully decentralized approximate zero-forcing

In this section, we derive a greedy algorithm for finding an approximate solution of (2). As means of guaranteeing that total and per-antenna power constraints are met, we additionally impose norm constraints on the local solutions. The operation of the algorithm is analyzed by performing a closed-form statistical analysis in the regime where  $M, K \gg 1$ .

### 3.1 Algorithm development

To start with, we define the partial channel and precoder matrices  $\mathbf{H}_m \triangleq [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_m]$  and  $\mathbf{W}_m \triangleq [\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_m]$ . The partial equivalent channel matrix at  $m$ th APB, defined as

$$\mathbf{E}_m \triangleq \mathbf{H}_m \mathbf{W}_m^H = \sum_{j=1}^m \mathbf{h}_j \mathbf{w}_j^H \quad (3)$$

will contain complete information about the IUI created by precoding at antennas 1 through  $m$ . This matrix is central to the proposed algorithm, which represents a greedy approach to solving problem (2). The algorithm is in form of a sequence of  $M$  steps, where at the  $m$ th step, the  $m$ th APB in the chain

1. receives  $\mathbf{E}_{m-1}$  from the preceding APB;
2. uses  $\mathbf{E}_{m-1}$  and local CSI  $\mathbf{h}_m$  to find a local precoder  $\mathbf{w}_m^H$ . The goal is to “force” the equivalent channel matrix after the current step,  $\mathbf{E}_m$ , to be as close as possible to an identity matrix in Frobenius norm sense by minimizing  $\|\mathbf{E}_m - \mathbf{I}_K\|_F^2 = \|\mathbf{E}_{m-1} + \mathbf{h}_m \mathbf{w}_m^H - \mathbf{I}_K\|_F^2$ ;
3. calculates  $\mathbf{E}_m$  using the newly formulated  $\mathbf{w}_m^H$  and passes it on to the next APB.

Due to practical considerations, the precoder formulation algorithm should include a mechanism for controlling the total expected transmit power

$$\mathbb{E}_{\mathbf{x}} \{ \|\mathbf{W}^H \mathbf{x}\|_2^2 \} = \|\mathbf{W}^H\|_F^2.$$

The precoder  $\mathbf{W}^H$  is built up row by row, with every APB having only one “go” at determining its contribution  $\mathbf{w}_m^H$ . A reasonable strategy for establishing power control in such a setting is to prescribe that all antennas have the same expected transmit power, equal to  $\epsilon^2$ . This is equivalent to imposing the constraint  $\|\mathbf{w}_m\|_2^2 = \epsilon^2, \forall m$ . Additional tuning of transmit power can be achieved by multiplying all per-antenna precoders with factor  $\alpha$  *after the process of formulating the precoder is finished*<sup>9</sup>. Total expected transmit power is thus equal to  $\alpha^2 M \epsilon^2$ . The described approach simultaneously solves the total and per-antenna power control problems, both being of high practical relevance, in a decentralized way.

<sup>9</sup> The need for a two-step adjustment of transmit power is clarified in Section 3.2.

**Algorithm 3** Fully decentralized calculation of an approximate zero-forcing precoder with per-antenna power constraints

- 1: Input:  $\mathbf{H} = [\mathbf{h}_1 \ \mathbf{h}_2 \ \dots \ \mathbf{h}_M]$
- 2:  $\mathbf{E}_0 \leftarrow \mathbf{0}_{K \times K}$
- 3: **for** antennas  $m = 1$  to  $M$  **do**
- 4:      $\mathbf{w}_m \leftarrow \epsilon \frac{(\mathbf{I}_K - \mathbf{E}_{m-1})^H \mathbf{h}_m}{\|(\mathbf{I}_K - \mathbf{E}_{m-1})^H \mathbf{h}_m\|_2}$
- 5:      $\mathbf{E}_m \leftarrow \mathbf{E}_{m-1} + \mathbf{h}_m \mathbf{w}_m^H$
- 6: **end for**
- 7: Output:  $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_M]$

Taking into account the aforementioned considerations, the central operation at each step of the algorithm consists of solving the constrained optimization problem

$$\begin{aligned} & \underset{\mathbf{w}_m}{\text{minimize}} && \|(\mathbf{E}_{m-1} - \mathbf{I}_K) + \mathbf{h}_m \mathbf{w}_m^H\|_F^2 \\ & \text{subject to} && \|\mathbf{w}_m\|_2^2 = \epsilon^2. \end{aligned} \quad (4)$$

The solution to (4) can conveniently be found in a closed form:

$$\mathbf{w}_{m,\text{opt}} = \epsilon \frac{(\mathbf{I}_K - \mathbf{E}_{m-1})^H \mathbf{h}_m}{\|(\mathbf{I}_K - \mathbf{E}_{m-1})^H \mathbf{h}_m\|_2}. \quad (5)$$

This can be shown in standard fashion by formulating a Lagrangian  $L(\mathbf{w}_m, \lambda) = \|(\mathbf{E}_{m-1} - \mathbf{I}_K) + \mathbf{h}_m \mathbf{w}_m^H\|_F^2 + \lambda (\|\mathbf{w}_m\|_2^2 - \epsilon^2)$ , finding its stationary point  $\mathbf{w}_{m,\text{opt}}$  which is plugged back to the constraint and finally solving a quadratic equation for  $\lambda$  with only one admissible solution, which altogether results in (5). The complete pseudocode of the proposed scheme is given in Algorithm 3.

### 3.2 Algorithm analysis

Interference-suppressing performance of the proposed method manifests dependence on  $\epsilon$ , and in this section we aim at providing insight into this dependence. Equivalent channel matrix  $\mathbf{E} = [e_{kl}] = \mathbf{H}\mathbf{W}^H$  lies in the focal point of the analysis since the signal-to-interference ratio (SIR) for the  $k$ th user is found as

$$SIR_k = \frac{|e_{kk}|^2}{\sum_{j=1, j \neq k}^K |e_{kj}|^2}. \quad (6)$$

For clarity of exposition, we introduce the shorthand  $\Phi_i \triangleq \mathbf{h}_i \mathbf{h}_i^H$  and further define

$$\tilde{\mathbf{w}}_i \triangleq (\mathbf{I}_K - \mathbf{E}_{i-1})^H \mathbf{h}_i, \quad (7)$$

$$\tilde{\epsilon}_i \triangleq \epsilon / \|\tilde{\mathbf{w}}_i\|_2. \quad (8)$$

From the recursions of Algorithm 3 we get

$$\begin{aligned} \mathbf{E} &= \sum_{m=1}^M \tilde{\epsilon}_m \Phi_m - \sum_{n=2}^M \sum_{m=1}^{n-1} \tilde{\epsilon}_n \tilde{\epsilon}_m \Phi_n \Phi_m \\ &+ \sum_{p=3}^M \sum_{n=2}^{p-1} \sum_{m=1}^{n-1} \tilde{\epsilon}_p \tilde{\epsilon}_n \tilde{\epsilon}_m \Phi_p \Phi_n \Phi_m - \dots (-1)^{M+1} \prod_{m=M}^1 \tilde{\epsilon}_m \Phi_m. \end{aligned} \quad (9)$$

Statistical analysis of  $SIR_k$  is intractable due to complicated dependencies between the terms in (9). We now introduce a series of simplifications of (9) that will help us gain insight in the behavior of  $SIR_k$ . Firstly, we limit the analysis to independent and identically distributed (iid) channel coefficients  $h_{km} \sim \mathcal{CN}(0, 1)$ . We also make use of

**Definition 3.1 (channel hardening, [7])** *Given a  $K \times 1$  random vector  $\boldsymbol{\psi}$ , we say that  $\boldsymbol{\psi}$  experiences “hardening” when  $\|\boldsymbol{\psi}\|_2^2 / \mathbb{E}\{\|\boldsymbol{\psi}\|_2^2\} \xrightarrow{P} 1$  as  $K \rightarrow \infty$ , where  $\xrightarrow{P}$  denotes convergence in probability. Moreover,  $\boldsymbol{\psi}$  hardens iff*

$$\lim_{K \rightarrow \infty} \frac{\text{Var}\{\|\boldsymbol{\psi}\|_2^2\}}{(\mathbb{E}\{\|\boldsymbol{\psi}\|_2^2\})^2} = 0.$$

As a consequence of channel hardening, the approximation

$$\|\boldsymbol{\psi}\|_2^2 \approx \mathbb{E}\{\|\boldsymbol{\psi}\|_2^2\} \quad (10)$$

can be used for  $K \gg 1$ .

In the simplified analysis, we consider values of  $\epsilon$  such that

$$\epsilon \ll \sqrt{\mathbb{E}_{\mathbf{H}}\{|h_{km}|^2\}} = 1, \quad (11)$$

and recall that  $K \gg 1$ . The analysis proceeds as follows.

1. **Expression (9) is truncated by neglecting the impact of higher order terms when  $\epsilon \ll 1$ :**

- (a) Random  $\tilde{\epsilon}_i$  (8) is substituted with a constant deterministic value. This is done by discarding from  $\|\tilde{\mathbf{w}}_i\|_2^2 = \|\mathbf{h}_i\|_2^2 - \mathbf{h}_i^H (\mathbf{E}_{i-1} + \mathbf{E}_{i-1}^H - \mathbf{E}_{i-1} \mathbf{E}_{i-1}^H) \mathbf{h}_i$  all terms depending on  $\epsilon$ , where we take into account (11). Subsequently, the hardening argument is invoked, which by employing (10) yields  $\|\tilde{\mathbf{w}}_i\|_2^2 \approx \|\mathbf{h}_i\|_2^2 \approx \mathbb{E} \{ \|\mathbf{h}_i\|_2^2 \} = K, \forall i$ . Now, in conjunction with (8) we get

$$\tilde{\epsilon}_i \approx \tilde{\epsilon} \triangleq \epsilon / \sqrt{K}. \quad (12)$$

- (b) By using (12) and again invoking (11), we discard from (9) all terms depending on  $\epsilon^n, n \geq 3$ .

These manipulations yield the truncated equivalent channel matrix

$$\hat{\mathbf{E}} = [\hat{\epsilon}_{kl}] = \tilde{\epsilon} \sum_{m=1}^M \Phi_m - \tilde{\epsilon}^2 \sum_{n=2}^M \sum_{m=1}^{n-1} \Phi_n \Phi_m. \quad (13)$$

**2. The SIR metric (6), averaged over channel realizations, is calculated for  $\hat{\mathbf{E}}$ :**

- Hardening applies to  $\hat{\epsilon}_{kk}$ . Namely, by use of the identity

$$[\Phi_n \Phi_m]_{ij} = h_{in} \left( \sum_{l=1}^K h_{ln}^* h_{lm} \right) h_{jm}^*,$$

we note that  $\hat{\epsilon}_{kk}$  can be decomposed as  $\hat{\epsilon}_{kk} = \hat{\epsilon}_{kk,s} + \hat{\epsilon}_{kk,\text{IUI}}$ , where

$$\hat{\epsilon}_{kk,s} = \tilde{\epsilon} \sum_{m=1}^M |h_{km}|^2 - \tilde{\epsilon}^2 \sum_{n=2}^M \sum_{m=1}^{n-1} |h_{kn}|^2 |h_{km}|^2$$

and

$$\hat{\epsilon}_{kk,\text{IUI}} = -\tilde{\epsilon}^2 \sum_{n=2}^M \sum_{m=1}^{n-1} h_{kn} \left( \sum_{l=1, l \neq k}^K h_{ln}^* h_{lm} \right) h_{km}^*.$$

Importantly,  $\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk,\text{IUI}} \} = 0$  and  $\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk,s} \hat{\epsilon}_{kk,\text{IUI}} \} = 0$ . Now we test whether  $\lim_{M \rightarrow \infty} \frac{\mathbb{E}_{\mathbf{H}} \{ |\hat{\epsilon}_{kk}|^2 \}}{(\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk} \})^2} = 1$ , which is analogous to the hardening test in Definition 3.1. The decomposition of  $\hat{\epsilon}_{kk}$  allows for writing  $\frac{\mathbb{E}_{\mathbf{H}} \{ |\hat{\epsilon}_{kk}|^2 \}}{(\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk} \})^2} = \frac{\mathbb{E}_{\mathbf{H}} \{ |\hat{\epsilon}_{kk,s}|^2 \}}{(\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk,s} \})^2} + \frac{\mathbb{E}_{\mathbf{H}} \{ |\hat{\epsilon}_{kk,\text{IUI}}|^2 \}}{(\mathbb{E}_{\mathbf{H}} \{ \hat{\epsilon}_{kk,s} \})^2}$ , and with some straightforward calculations it can be shown that the first fraction on the right-hand side of the equality sign tends to 1 and the second to 0 as  $M \rightarrow \infty$ .



The presented considerations enable us to write  $\mathbb{E}_{\mathbf{H}}\{|\hat{e}_{kk}|^2\} \approx (\mathbb{E}_{\mathbf{H}}\{\hat{e}_{kk}\})^2$ , resulting in

$$\begin{aligned} \mathbb{E}_{\mathbf{H}} \left\{ \frac{|\hat{e}_{kk}|^2}{\sum_{j=1, j \neq k}^K |\hat{e}_{kj}|^2} \right\} &\approx \frac{\mathbb{E}_{\mathbf{H}} \{|\hat{e}_{kk}|^2\}}{\sum_{j=1, j \neq k}^K \mathbb{E}_{\mathbf{H}} \{|\hat{e}_{kj}|^2\}} \\ &\approx \frac{(\mathbb{E}_{\mathbf{H}} \{\hat{e}_{kk}\})^2}{\sum_{j=1, j \neq k}^K \mathbb{E}_{\mathbf{H}} \{|\hat{e}_{kj}|^2\}} = \frac{\Delta_k}{\Omega_k} \triangleq \varsigma. \end{aligned} \quad (14)$$

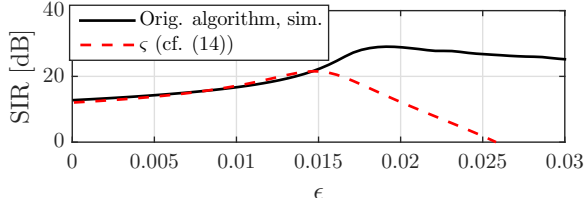
3. **The dependence of “average received power”  $\Delta_k$  and “average IUI power”  $\Omega_k$  on  $M$  and  $K$  is analyzed.** In the process, we use straightforwardly provable identities ( $i \neq j \neq k$ ):

$$\begin{aligned} \mathbb{E}_{\mathbf{H}}\{\Phi_i\} &= \mathbb{E}_{\mathbf{H}}\{\Phi_i\Phi_j\} = \mathbf{I}_K, \\ \mathbb{E}_{\mathbf{H}}\{\text{Tr}(\Phi_i^2)\} &= \mathbb{E}_{\mathbf{H}}\{\text{Tr}(\Phi_j\Phi_i^2)\} = K(K+1), \\ \mathbb{E}_{\mathbf{H}}\{\text{Tr}(\Phi_i\Phi_j\Phi_i\Phi_j)\} &= 2K(K+1), \\ \mathbb{E}_{\mathbf{H}}\{\text{Tr}(\Phi_i\Phi_j\Phi_k)\} &= K. \end{aligned}$$

By noting that all  $\hat{e}_{kl}$  are iid and recalling that  $\hat{e}_{kk}$  hardens, after some calculations we get

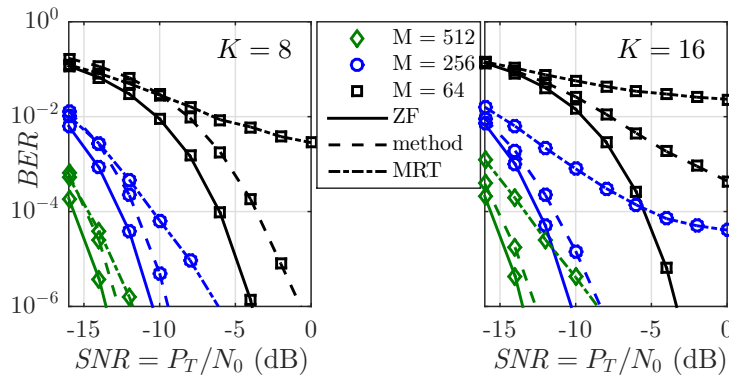
$$\begin{aligned} \Delta_k &= \tilde{\epsilon}^2 M^2 [1 - \tilde{\epsilon}(M-1)/2]^2, \\ \Omega_k &\approx \frac{1}{K} \mathbb{E}_{\mathbf{H}} \left\{ \text{Tr}(\widehat{\mathbf{E}}\widehat{\mathbf{E}}^H) \right\} - \Delta_k = \kappa_2 \tilde{\epsilon}^2 - \kappa_3 \tilde{\epsilon}^3 + \kappa_4 \tilde{\epsilon}^4, \end{aligned} \quad (15)$$

where  $\kappa_2 = MK$ ,  $\kappa_3 = 2MK(M-1)$  and  $\kappa_4 = MK(M-1)(2M+K-2)/2$ .



**Figure 2:** Comparison of simulated  $\mathbb{E}_{\mathbf{H}}\{SIR_k\}$  for the original algorithm and metric  $\varsigma$  from (14).  $M = 256$ ,  $K = 16$ .

Finally,  $\varsigma$  from (14) is compared with the simulated  $\mathbb{E}_{\mathbf{H}}\{SIR_k\}$ , as shown in Fig. 2. Notwithstanding the discrepancy between  $\varsigma$  and  $\mathbb{E}_{\mathbf{H}}\{SIR_k\}$  at larger  $\epsilon$  which is due to truncation of (9) and approximations in (14), both



**Figure 3:** Comparison of channel-averaged  $BER$  for the proposed scheme (Algorithm 1), ZF and MRT. Modulation: QPSK.

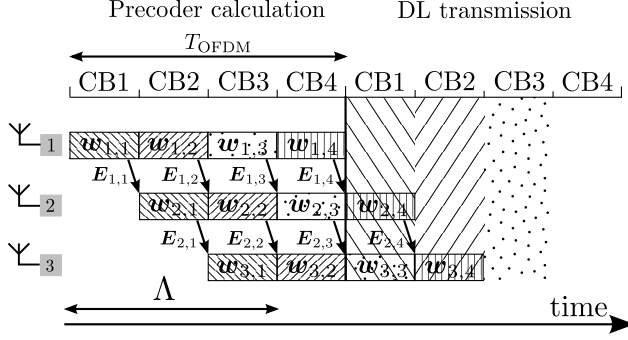
metrics are observed to follow a similar trend, where average  $SIR$  is maximized by some value of  $\tilde{\epsilon}$  (or  $\epsilon$ , cf. (12)). For the approximation (14) it can be shown that with  $M \gg 1$ , this value is asymptotically equal to  $\tilde{\epsilon}$  that minimizes ergodic IUI  $\Omega_k$ , with both being asymptotically equal to  $\tilde{\epsilon}^* = 1/M$ . Using this fact, we can conjecture why and how the performance of the proposed algorithm depends on  $\epsilon$  by observing the behavior of  $\Omega_k$ . At large  $M$ ,  $\Omega_k \approx MK\tilde{\epsilon}^2 - 2M^2K\tilde{\epsilon}^3 + M^3K\tilde{\epsilon}^4$  and for  $\tilde{\epsilon} = 1/M$  the contributions of even- and odd-order terms in  $\Omega_k$  cancel out, resulting in zero ergodic IUI. Due to related forms of (9) and (13), it can be extrapolated that similar canceling of interference terms at optimal  $\epsilon$  occurs in the original algorithm.

The choice of  $\epsilon$  therefore has a decisive impact on the performance of the proposed algorithm. A properly chosen  $\epsilon$  maximizes IUI suppression. On the other hand, if we simply let  $\epsilon \rightarrow 0^+$ , all higher order terms in (9) can be neglected, yielding  $\mathbf{E} \approx \tilde{\epsilon}\mathbf{H}\mathbf{H}^H$ . In this case, the algorithm is asymptotically equivalent to maximum ratio transmission (MRT) ( $\mathbf{W}^H = \mathbf{H}^H$ ) where performance is limited by IUI. Dependence of performance on  $\epsilon$  makes clear why a two-step adjustment of transmit power is needed: the choice of  $\epsilon$  determines the performance, and  $\alpha$  tunes the transmit power to the desired value.

## 4 Performance of the proposed algorithm. Implementation considerations

Uncoded bit error rate ( $BER$ ), averaged over channel realizations, has been simulated for the proposed algorithm, ZF and MRT for iid  $h_{km}$ . Total average

transmit power for all schemes is set to  $P_T K$ , and different values of  $M$ ,  $K$  and  $SNR \triangleq P_T/N_0$  are tested. The  $SIR$ -maximizing value of  $\epsilon$ , found numerically for each  $(M, K)$  is applied. The results, shown in Fig. 3, demonstrate that the  $SNR$  gap between the method and ZF for a very large  $M = 512$  stays essentially constant ( $\approx 1$  dB at  $BER = 10^{-6}$ ) as  $K$  doubles. Moreover, the smaller the  $M$ , the more drastic the increase of this  $SNR$  gap with the doubling of  $K$ . Hence, the proposed method is capable of removing essentially all of the IUI at extremely large  $M$ , with the capability decreasing as  $M$  decreases.



**Figure 4:** Timing diagram of precoder formulation. Toy example system with 3 antennas and 4 coherence blocks (CB).

In order to completely assess the viability of the daisy chain topology in MaMI BS, two implementation aspects need to be investigated. One is the processing latency of the precoder formulation, the other - required throughput of intra-chain links. To this end, we assume a multicarrier setting as in Section 1, with the channel coherence bandwidth spanning  $N_{\text{coh}}$  subcarriers and with  $N_{\text{bl}} = \lceil N_{\text{used}}/N_{\text{coh}} \rceil$  coherence blocks (CB). By exploiting channel coherence, the same precoder  $\mathbf{w}_{m,b}$  is used at antenna  $m$  for all the subcarriers in the CB  $b$ . Computational tasks of formulating  $\mathbf{w}_{m,b}$  and precoding are scheduled based on dividing each OFDM symbol into  $N_{\text{bl}}$  time slots, one for each CB. The formulation of  $\mathbf{w}_{m,b}$  is then pipelined over time and antennas as illustrated in Fig. 4, with precoding done in an analogous pipelined fashion. Total latency of calculating all  $\mathbf{w}_{m,b}$  for one CB is  $\Lambda = MT_{\text{OFDM}}/N_{\text{bl}}$ . Taking into account that  $\mathbf{E}_{m,b}$  contains  $K^2$  complex samples and by neglecting the time needed to perform the actual computation of  $\mathbf{w}_{m,b}$ , the required throughput of intra-chain links is  $R_{\text{daisy}} = N_{\text{bl}}K^2w/T_{\text{OFDM}}$ , which yields (cf. Section 1)  $R_{\text{daisy}}/R_{\text{central}} \approx K^2/(N_{\text{coh}}M)$ . Presented results indicate that using the daisy chain is beneficial (reduced interconnect throughput with small performance

degradation compared to ZF) when  $M/K^2$  is large. Moreover, when  $M < N_{\text{bl}}$ , latency  $A$  is less than the duration of one OFDM symbol.

## 5 Conclusion

We have developed and analyzed an algorithm for suppressing interuser interference in the downlink of a massive MIMO system where the BS lacks a central processing unit and antennas are connected in a daisy chain topology, having knowledge of the local channel only. For a very large number of BS antennas, the algorithm performs close to ZF, while having reduced requirements on interconnection throughput. These insights indicate that daisy chain is the topology of choice for the implementation of MIMO BSs with a very large number of antennas.



# Bibliography

- [1] F. Rusek *et al.*, “Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays,” *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40-60, Jan. 2013.
- [2] C. Shepard *et al.*, “Argos: Practical Many-Antenna Base Stations,” in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, Istanbul, Turkey, Aug. 2012, pp. 53-64.
- [3] S. Malkowsky *et al.*, “The World’s First Real-Time Testbed for Massive MIMO: Design, Implementation, and Validation,” *IEEE Access*, vol. 5, pp. 9073-9088, 2017.
- [4] K. Li *et al.*, “Decentralized Baseband Processing for Massive MU-MIMO Systems,” *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 4, pp. 491-507, Dec. 2017.
- [5] K. Li *et al.*, “Feedforward Architectures for Decentralized Precoding in Massive MU-MIMO Systems.” [Online]. Available: <http://arxiv.org/abs/1804.10987>
- [6] C. Jeon *et al.*, “Decentralized Equalization with Feedforward Architectures for Massive MU-MIMO.” [Online]. Available: <http://arxiv.org/abs/1808.04473>.
- [7] H. Q. Ngo and E. G. Larsson, “No Downlink Pilots Are Needed in TDD Massive MIMO,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921-2935, May 2017.



*Paper V*





# Impact of Relay Cooperation on the Performance of Large-scale Multipair Two-way Relay Networks

We consider a multipair two-way relay communication network, where pairs of user devices exchange information via a relay system. The communication between users employs time division duplex, with all users transmitting simultaneously to relays in one time slot and relays sending the processed information to all users in the next time slot. The relay system consists of a large number of single antenna units that can form groups. Within each group, relays exchange channel state information (CSI), signals received in the uplink and signals intended for downlink transmission. On the other hand, per-group CSI and uplink/downlink signals (data) are not exchanged between groups, which perform the data processing completely independently. Assuming that the groups perform zero-forcing in both uplink and downlink, we derive a lower bound for the ergodic sumrate of the described system as a function of the relay group size. By close observation of this lower bound, it is concluded that the sumrate is essentially independent of group size when the group size is much larger than the number of user pairs. This indicates that a very large group of cooperating relays can be substituted by a number of smaller groups, without incurring any significant performance reduction. Moreover, this result implies that relay cooperation is more efficient (in terms of resources spent on cooperation) when several smaller relay groups are used in contrast to a single, large group.

---

©2018 IEEE. Based on

Muris Sarajlić, Liang Liu, Fredrik Rusek, Farhana Sheikh and Ove Edfors,

“Impact of Relay Cooperation on the Performance of Large-scale Multipair Two-way Relay Networks,”

to appear in *Proceedings of the IEEE Global Communications Conference (GLOBECOM) 2018*, Abu Dhabi, UAE, 2018.



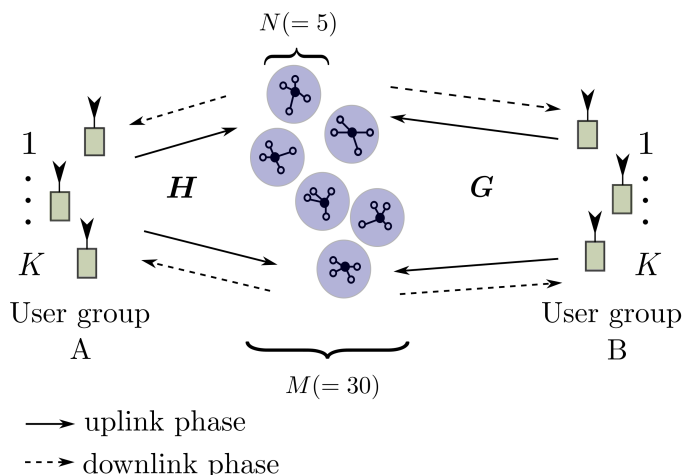
## 1 Introduction

Multipair two-way relay systems have attracted significant attention in the research community, due to their inherent ability to overcome the halving of system sumrate (stemming from half-duplex operation), with essential doubling of sumrate compared to ordinary one-way relaying [1]. Much of the research efforts regarding these systems were focused on developing signal processing algorithms at the relays that are tailored to fit a certain target objective, e.g. interference cancellation or sumrate maximization [1], [2]. Recently, multipair two-way relaying systems with a large number of relays/relay antennas were considered [3–8]. By employing a large-scale relay network, system performance is boosted via the channel hardening effect, thus either improving system sumrate or increasing coverage compared to relay systems that operate on a smaller scale.

Previous work on large-scale multipair two-way (LS-MTW) relay systems is limited to two extreme scenarios. In the first scenario, a large number of non-cooperating single-antenna relays processes the data in a fully decentralized fashion. Such a setup is described in [3], where individual single-antenna relays perform amplify-and-forward processing of the data on the bidirectional links, based only on their local CSI. No data or CSI is exchanged between the relays, and their sheer large number is relied upon to deliver satisfying performance. On the other end of the spectrum is the scenario where a single relay with a large number of antennas performs the uplink and downlink processing. Theoretical performance characterization for this setup was analyzed in [4–8], where the relay is assumed to employ simple linear processing (maximum ratio combining/transmission and zero forcing).

To the best of our knowledge, there is no prior analysis of LS-MTW relay networks with an arbitrary degree of cooperation among relays. By “degree of cooperation” we here refer to the number of relays that will exchange data and channel state information inside a closed group, with no exchange occurring between groups. The number of closely cooperating relays directly trades system performance against data exchange cost, and is therefore an important design parameter for practical implementations of LS-MTW networks. An illustration of a MTW network with grouped relays is shown in Fig. 1.

This work provides a comprehensive analysis of the effects of relay cooperation in an LS-MTW relay system where zero-forcing processing is used. We derive a lower bound for ergodic system sumrate that is tight at high SNR. Furthermore, we make use of this bound to analyze the behavior of system performance as the number of closely cooperating relays changes. Finally, we reflect on the choice of the degree of relay cooperation that maximizes the cost-effectiveness of cooperation.



**Figure 1:** Multipair bidirectional relay network with an arbitrary degree of relay cooperation

## 2 System setup

In this paper, we analyze the multi-pair bidirectional symmetric relay network, where  $M$  relays serve to connect two separate groups of  $K$  user units, as illustrated in Fig. 1. Each user from group A is connected in a pairwise fashion with a corresponding user in group B, with  $K$  pairs formed in total. It is assumed that there is no direct link between the users in a pair, so the pairwise connections are established solely via the relays. Moreover, the information flow between the two users in a pair is assumed symmetric.

The exchange of information is split in two phases, uplink and downlink, observed from the perspective of the relays, which occur in alternating time slots. In the uplink phase, all  $2K$  user units simultaneously transmit to the relays. The relays process the received signal and send the processed information to the users in the downlink phase, with the direction of information flow swapped compared to the uplink phase (processed information that came on the uplink of channel  $\mathbf{H}$  is transmitted on the downlink of channel  $\mathbf{G}$ , and vice versa).

Furthermore, we assume that it is only the relays that have knowledge of the channels  $\mathbf{H}$  and  $\mathbf{G}$ . In practice, the channels can be estimated in the uplink by transmission of orthogonal pilot sequences from the user units, and downlink channels are then automatically obtained assuming that radio channel

reciprocity holds and that reciprocity calibration is performed at the relays. In this sense, the analyzed relay system is equivalent to collocated or distributed massive MIMO (MaMI) systems working in time division duplex (TDD), and channel estimation and reciprocity calibration methods developed for TDD MaMI readily apply [9–11]. However, for clarity of analysis, in this work we assume perfect channel knowledge and channel reciprocity.

The focus of investigation in this work is the impact of cooperation between the relays on the overall system performance. The level of inter-relay cooperation is the parameter that trades off network performance with cost of backhaul information exchange. To this end, we assume the following hierarchical structure of the relay system:

- The relays are assumed to be divided in equally-sized groups, each containing  $N$  relays. Inside the group, channel state information (CSI), symbols received in the uplink phase and symbols to be transmitted in the downlink phase are shared mutually among all relays. Moreover, the relays inside the group are assumed to be time and frequency synchronized. The data is congregated and processed for uplink and downlink in a central group processor (CGP). One of the relays can take on the role of the CGP, and is referred to as the group master (represented by  $\bullet$  in Fig. 1). Most importantly, no data or CSI information is exchanged between the groups, and each group performs data processing independently of others.
- Groups (or equivalently, group masters) are assumed to be synchronized in time and frequency, and this is the only form of inter-group cooperation.

The two-tier hierarchy of cooperation enables us to cover the entire space of cooperative networks that lies in between the two extreme cases:

- For  $N = 1$  we have the fully decentralized multipair relaying scenario, where single-antenna relays use their local CSI to process and relay the received data without exchanging any CSI or received data information with other relays. Such a scenario was analyzed in [3].
- The case of  $N = M$  represents the perfectly centralized relaying scenario where all CSI and received data is available at a central point that performs data processing. Usually this setup is cast in the form of one (massive) MIMO relay, as in [4].

In general, there are no constraints on the geographical distribution of relays in a group, which can be collocated or distributed. Likewise, the type of

connections between the relays in a group is arbitrary and can be wireless or wired. We note, however, that a group of  $N$  collocated users can be observed as a single MIMO relay with  $N$  antennas. We also note that the stratification of relay cooperation enables the design of a layered and scalable synchronization protocol. Instead of synchronizing all the relays to a common beacon, synchronization can be done first on the group level and then among the group masters, resulting in a completely decentralized synchronization scheme. For the sake of simplicity, we assume that the intra- and inter-group synchronization is perfect, and leave the analysis of the impact of synchronization errors as a subject for future work.

### 3 System model

We start the description of the system model by denoting with

$$L = \frac{M}{N} \quad (1)$$

the total number of relay groups. In each channel use, transmitted user symbols are represented by the  $2K \times 1$  complex vector  $\mathbf{x}$  with covariance matrix  $\mathbb{E}\{\mathbf{x}\mathbf{x}^H\} = \mathbf{I}_{2K}$ . The user symbol vector can be represented as  $\mathbf{x} = [\mathbf{x}_A^T \mathbf{x}_B^T]^T$ , where  $\mathbf{x}_A$  and  $\mathbf{x}_B$  are symbols transmitted from the left-hand-side and right-hand-side groups of users, illustrated in Figure 1, respectively.

Focusing on the  $i$ th relay group, we build the system model step by step, following the uplink - downlink flow of information. First, we denote by

$$\mathbf{\Xi}_{u,i} = [\mathbf{H}_i \mathbf{G}_i]_{N \times 2K} \quad (2)$$

the composite uplink channel between the  $i$ th relay group and all the users. Received signal vector at the  $i$ th relay group is then

$$\mathbf{y}_i = \sqrt{P_U} \mathbf{\Xi}_{u,i} \mathbf{x} + \mathbf{n}_{R,i}, \quad (3)$$

where  $\mathbf{n}_{R,i}$  is an  $N \times 1$  zero-mean circularly symmetric complex Gaussian (ZMCSCG) vector of thermal noise with covariance matrix  $\mathbb{E}\{\mathbf{n}_{R,i} \mathbf{n}_{R,i}^H\} = N_{0,R} \mathbf{I}_N$  and  $P_U$  is the uplink transmit power per user, assumed to be same for all users.

The received signal in the uplink is linearly filtered with  $\mathbf{W}_{u,i}$  to yield the estimates of user data symbols:

$$\hat{\mathbf{x}}_i = \mathbf{W}_{u,i} \mathbf{y}_i = \sqrt{P_U} \mathbf{W}_{u,i} \mathbf{\Xi}_{u,i} \mathbf{x} + \mathbf{W}_{u,i} \mathbf{n}_{R,i}. \quad (4)$$

After uplink filtering, the downlink precoder  $\mathbf{W}_{d,i}$  is applied to symbol estimates, together with a scaling factor  $\alpha_i$  ensuring proper transmitted power. We assume that the user power allocation in the downlink is uniform.

The uplink/downlink linear processing can be compactly represented by a general complex gain matrix

$$\mathbf{W}_i = \mathbf{W}_{d,i} \mathbf{W}_{u,i}. \quad (5)$$

Altogether, the transmit signal vector from the  $i$ th relay group is

$$\begin{aligned} \mathbf{t}_i &= \alpha_i \mathbf{W}_{d,i} \hat{\mathbf{x}}_i \\ &= \alpha_i \sqrt{P_U} \mathbf{W}_i \boldsymbol{\Xi}_{u,i} \mathbf{x} + \alpha_i \mathbf{W}_i \mathbf{n}_{R,i}. \end{aligned} \quad (6)$$

The power scaling coefficient  $\alpha_i$  is determined so that the transmitted power per group averaged over data and noise realizations equals  $P_{R,i}$ . For the heavily restricted decentralized setup considered here, a practically implementable strategy of power allocation between relay groups is that all groups have the same transmit power, so  $P_{R,i} = P_R, \forall i$ . Overall, we have

$$\mathbb{E}_{\mathbf{x}, \mathbf{n}} \left\{ \|\mathbf{t}_i\|^2 \right\} = P_R, \quad (7)$$

which readily yields

$$\alpha_i = \sqrt{\frac{P_R}{P_U \|\mathbf{W}_i \boldsymbol{\Xi}_{u,i}\|_F^2 + N_{0,R} \|\mathbf{W}_i\|_F^2}}. \quad (8)$$

We define the composite downlink channel between the  $i$ th relay group and all the users as

$$\boldsymbol{\Xi}_{d,i}^T = [\mathbf{G}_i \ \mathbf{H}_i]_{N \times 2K}^T \quad (9)$$

The contribution of the  $i$ th relay group to the received signal at the users,  $\mathbf{z}_i$ , is thus

$$\begin{aligned} \mathbf{z}_i &= \begin{bmatrix} z_{B,i} \\ z_{A,i} \end{bmatrix} = \boldsymbol{\Xi}_{d,i}^T \mathbf{t}_i \\ &= \alpha_i \sqrt{P_U} \boldsymbol{\Xi}_{d,i}^T \mathbf{W}_i \boldsymbol{\Xi}_{u,i} \mathbf{x} + \alpha_i \boldsymbol{\Xi}_{d,i}^T \mathbf{W}_i \mathbf{n}_{R,i}. \end{aligned} \quad (10)$$



The total received signal vector at the users is hence

$$\begin{aligned} \mathbf{z} &= \sum_{i=1}^L \mathbf{z}_i + \mathbf{n}_U \\ &= \sqrt{P_U} \left( \sum_{i=1}^L \alpha_i \boldsymbol{\Xi}_{d,i}^T \mathbf{W}_i \boldsymbol{\Xi}_{u,i} \right) \mathbf{x} \\ &\quad + \sum_{i=1}^L \alpha_i \boldsymbol{\Xi}_{d,i}^T \mathbf{W}_i \mathbf{n}_{R,i} + \mathbf{n}_U, \end{aligned} \quad (11)$$

where  $\mathbf{n}_U$  is the  $N \times 1$  ZMCSCG vector of thermal noise at the users, with covariance  $\mathbb{E}\{\mathbf{n}_U \mathbf{n}_U^H\} = N_{0,U} \mathbf{I}_{2K}$ .

For the benefit of further analysis, we define the uplink and downlink SNRs as

$$SNR_u = \frac{P_U}{N_{0,R}}, \text{ and } SNR_d = \frac{P_R}{N_{0,U}}. \quad (12)$$

The overall system model (11) can be expanded for the received symbol at a particular user, say  $k$ th user from group A. This reveals that the performance in the general case is limited by four distinct impairments: self-interference, interuser interference, precoded thermal noise at the relays and thermal noise at the users:

$$\begin{aligned} z_{A,k} &= \underbrace{\sqrt{P_U} \left( \sum_{i=1}^L \alpha_i \mathbf{h}_{k,i}^T \mathbf{W}_i \mathbf{g}_{k,i} \right)}_{\text{wanted information, } x_{W,k}} x_{B,k} \\ &\quad + \underbrace{\sqrt{P_U} \left( \sum_{i=1}^L \alpha_i \mathbf{h}_{k,i}^T \mathbf{W}_i \mathbf{h}_{k,i} \right)}_{\text{self-interference, } \nu_{SI,k}} x_{A,k} \\ &\quad + \underbrace{\sqrt{P_U} \sum_{i=1}^L \alpha_i \sum_{\substack{j=1, \\ j \neq k}}^K \mathbf{h}_{k,i}^T \mathbf{W}_i (\mathbf{h}_{j,i} x_{A,j} + \mathbf{g}_{j,i} x_{B,j})}_{\text{interuser interference, } \nu_{UI,k}} \\ &\quad + \underbrace{\sum_{i=1}^L \alpha_i \mathbf{h}_{k,i}^T \mathbf{W}_i \mathbf{n}_{R,i}}_{\text{precoded noise from relays, } \nu_{PN,k}} + \underbrace{n_{A,k}}_{\text{thermal noise at users}}. \end{aligned} \quad (13)$$

In the follow-up, we consider the case when zero-forcing is chosen as the linear processing scheme at individual relay groups and analyze system performance, averaged over channel realizations. The goal of the analysis is to determine the closed-form dependence of system performance (quantified by ergodic system sumrate) on relay group size  $N$ .

## 4 Ergodic system sumrate calculation with per-group zero-forcing

If zero-forcing (ZF) is chosen for linear processing, the uplink and downlink processing matrices at each relay group are calculated as

$$\begin{aligned} \mathbf{W}_{u,i} &= (\mathbf{\Xi}_{u,i}^H \mathbf{\Xi}_{u,i})^{-1} \mathbf{\Xi}_{u,i}^H, \text{ and} \\ \mathbf{W}_{d,i} &= \mathbf{\Xi}_{d,i}^* (\mathbf{\Xi}_{d,i}^T \mathbf{\Xi}_{d,i}^*)^{-1}. \end{aligned} \quad (14)$$

Back-to-back ZF processing will completely eliminate self- and interuser interference, leaving the precoded noise from relays and noise at the user terminals as sources of impairment. A strong requirement for total interference elimination is that  $N > 2K$ .

In order to gain some insight in the connection between system performance (quantified by system sumrate) and system parameters  $M$  and  $N$ , we assume that  $SNR_u$  is high. Typically, a high SNR would mean that the geographical distances between users and relays are small, and that the relays are used to boost the system sumrate (in contrast to e.g. a range extension scenario).

Under the high-SNR assumption and additionally assuming that  $\mathbf{H}_i$  and  $\mathbf{G}_i$  are well-conditioned, the influence of precoded thermal noise at the relays can be neglected, so the system model (11) simplifies to

$$\mathbf{z} = \sqrt{P_U} \sum_{i=1}^L \alpha_i \mathbf{x} + \mathbf{n}_U. \quad (15)$$

A basis for performance evaluation is the per-user  $SNR$ , defined for the  $k$ th user in group A as the ratio of powers of the information signal part and impairments from (13), which, due to all interference being eliminated and the

high-SNR assumption, becomes

$$\begin{aligned}
SNR_{A,k} &= \frac{\mathbb{E}_x \{ |x_{W,k}|^2 \}}{N_{0,U}} = \frac{P_U}{N_{0,U}} \left( \sum_{i=1}^L \alpha_i \right)^2 \\
&\stackrel{a)}{=} \frac{P_R}{N_{0,U}} \left( \sum_{i=1}^L \frac{1}{\sqrt{\|\mathbf{W}_i \boldsymbol{\Xi}_{u,i}\|_F^2}} \right)^2 \\
&\stackrel{b)}{=} \frac{P_R}{N_{0,U}} \left( \sum_{i=1}^L \frac{1}{\sqrt{\text{Tr} \left[ \left( \boldsymbol{\Xi}_{d,i}^T \boldsymbol{\Xi}_{d,i}^* \right)^{-1} \right]}} \right)^2,
\end{aligned} \tag{16}$$

where *a*) follows from (8), with the assumption of high-SNR at the relays, and *b*) from (5) and (14).

Instantaneous per-user performance is characterized by user rate:

$$R_{A,k} = \log_2 (1 + SNR_{A,k}) \quad [\text{bps/Hz}], \tag{17}$$

and overall system performance by ergodic system sumrate, calculated as

$$R = \mathbb{E}_{\mathbf{H}, \mathbf{G}} \left\{ \frac{1}{2} \left( 2 \sum_{k=1}^K R_{A,k} \right) \right\} = \sum_{k=1}^K \mathbb{E}_{\mathbf{H}, \mathbf{G}} \{ R_{A,k} \}. \tag{18}$$

The factor of 2 accounts for the fact that the information flow in the system is symmetric, so members of the *k*th user pair have the same information transmission capacity. The factor of 1/2, on the other hand, stems from half-duplex operation. From here, the benefit of the symmetric multi-pair setup compared with ordinary relaying schemes becomes clear: simultaneous and symmetric transmission from both user groups manages to (approximately) compensate for the halving of the capacity due to TDD splitting of uplink and downlink.

We proceed with calculating (18), and in the process, we make use of

**Lemma 1** *Let  $\boldsymbol{\Psi} = [\Psi_1 \ \Psi_2 \ \dots \ \Psi_N]$  be a vector of nonnegative random variables  $\Psi_i$ , with  $\psi_i$  denoting realizations of  $\Psi_i$ . Then*

$$\mathbb{E}_{\boldsymbol{\Psi}} \log_2 \left[ 1 + \left( \sum_i \frac{1}{\sqrt{\psi_i}} \right)^2 \right] > \log_2 \left[ \left( \sum_i \frac{1}{\sqrt{\mathbb{E}_{\boldsymbol{\Psi}} \psi_i}} \right)^2 \right].$$

*Proof:* The proof is given in Appendix A.

We employ Lemma 1, assuming in the process that  $P_R/N_{0,U} = 1$  without loss of generality, to obtain a lower bound on the ergodic information rate of the  $k$ th user as

$$\mathbb{E}_{\mathbf{H},\mathbf{G}} \{R_{A,k}\} > \log_2 \left[ \frac{P_R}{N_{0,U}} \left( \sum_{i=1}^L \frac{1}{\sqrt{\mathbb{E}_{\mathbf{H},\mathbf{G}} \{\zeta_i\}}} \right)^2 \right], \quad (19)$$

where, for sake of clarity, we introduce the ZF precoding scaling factor

$$\zeta_i = \text{Tr} \left[ (\mathbf{\Xi}_{d,i}^T \mathbf{\Xi}_{d,i}^*)^{-1} \right]. \quad (20)$$

Now we assume that  $\mathbf{H}_i$  and  $\mathbf{G}_i$  are iid Rayleigh fading channels with pathloss and shadowing, modeled as

$$\mathbf{H}_i = \widetilde{\mathbf{H}}_i \mathbf{D}_{A,i}^{1/2} \text{ and } \mathbf{G}_i = \widetilde{\mathbf{G}}_i \mathbf{D}_{B,i}^{1/2}. \quad (21)$$

The entries of  $N \times K$  matrices  $\widetilde{\mathbf{H}}$  and  $\widetilde{\mathbf{G}}$  are iid ZMCSG with unit variance, and the diagonal matrices

$$\begin{aligned} \mathbf{D}_{A,i}^{1/2} &= \text{diag} \left( \sqrt{\beta_{A,1,i}}, \sqrt{\beta_{A,2,i}}, \dots, \sqrt{\beta_{A,K,i}} \right) \text{ and} \\ \mathbf{D}_{B,i}^{1/2} &= \text{diag} \left( \sqrt{\beta_{B,1,i}}, \sqrt{\beta_{B,2,i}}, \dots, \sqrt{\beta_{B,K,i}} \right) \end{aligned} \quad (22)$$

are used to model propagation losses and large-scale fading. In order for channel matrices to be decomposable as in (21), the propagation amplitude gain  $\sqrt{\beta_{(A,B),k,i}} > 0$  needs to be the same from user  $k$  to all relays in group  $i$ , which implies that the relays of that group are assumed to be collocated and experience the same large scale fading in relation to user  $k$ . These conditions are readily satisfied if a relay group is implemented in form of a single MIMO relay with a compact form factor. Otherwise, they can be met by applying an appropriate relay grouping scheme, which is an interesting research problem in itself, but falls outside of the scope of this paper.

Using well-known results from random matrix theory [12] and the identity  $\text{Tr}(\mathbf{AB}) = \text{Tr}(\mathbf{BA})$ , it can be shown that

$$\mathbb{E}_{\mathbf{H},\mathbf{G}} \{\zeta_i\} = \frac{\gamma_i}{N - 2K}, \quad (23)$$

where

$$\gamma_i = \sum_{k=1}^K \left( \frac{1}{\beta_{A,k,i}} + \frac{1}{\beta_{B,k,i}} \right). \quad (24)$$

Combining (23) with (19) yields the lower bound on per-user rate

$$\mathbb{E}_{\mathbf{H}, \mathbf{G}} \{R_{A,k}\} > \log_2 \left[ \frac{P_R}{N_{0,U}} (N - 2K) \delta \right], \quad (25)$$

with

$$\delta = \left( \sum_{i=1}^L \frac{1}{\sqrt{\gamma_i}} \right)^2. \quad (26)$$

For convenience of exposition, in the follow-up we will refer to  $\gamma_i$  and  $\delta$  as power imbalance factor and array gain degradation factor, respectively.

Overall, the lower bound on system sumrate for the multipair two-way relay system with relay grouping and ZF processing at high SNR is given from (18) and (25) by

$$R > \max \left\{ 0, K \log_2 \left[ \frac{P_R}{N_{0,U}} (N - 2K) \delta \right] \right\}. \quad (27)$$

## 5 Analysis and discussion

In order to have a fair comparison between systems, we assume that the pathloss and shadowing power gains are normalized so that

$$\text{Tr}(\mathbf{D}_{A,i}) = \text{Tr}(\mathbf{D}_{B,i}) = K, \quad \forall i, \quad (28)$$

which implies  $\mathbb{E} \{ \|\mathbf{H}_i\|_F^2 \} = \mathbb{E} \{ \|\mathbf{G}_i\|_F^2 \} = NK, \forall i$ . Given the constraint (28), it is easy to show that power imbalance and array gain degradation factors are lower (respectively, upper) bounded as

$$\gamma_i \geq 2K \text{ and } \delta \leq \frac{L^2}{2K}, \quad (29)$$

where equality holds in the case  $\mathbf{D}_{A,i} = \mathbf{D}_{B,i} = \mathbf{I}_K$ . In other words, the lower bound on system sumrate from (27) is maximized when there are no pathloss/shadowing power imbalances between users. In practical system deployments, such imbalances will invariably exist and the resulting degradation of sumrate can be combatted by either performing waterfilling-based user power weighting in the downlink, or by employing advanced user scheduling techniques. Analysis of the effects of these approaches is beyond the scope of this work.

For a fair comparison between relay systems with differing  $M$  and  $N$ , we need to assume that the amount of transmit power allocated to the entire relay

system is fixed, and we denote this power by  $P_T$ . As mentioned previously, due to limited coordination, it is reasonable to assume that the total power allocated to relays is distributed equally among relay groups, so  $P_R = P_T/L$ . By taking into account (29), we can write  $\delta = \frac{L^2}{2K}\epsilon$ ,  $\epsilon \leq 1$ , which yields the lower bound on sumrate that allows for a fair comparison between different relay systems:

$$R > \max \left\{ 0, K \log_2 \left( \frac{P_T}{N_{0,U}} \frac{M}{N} \frac{N - 2K}{2K} \epsilon \right) \right\}. \quad (30)$$

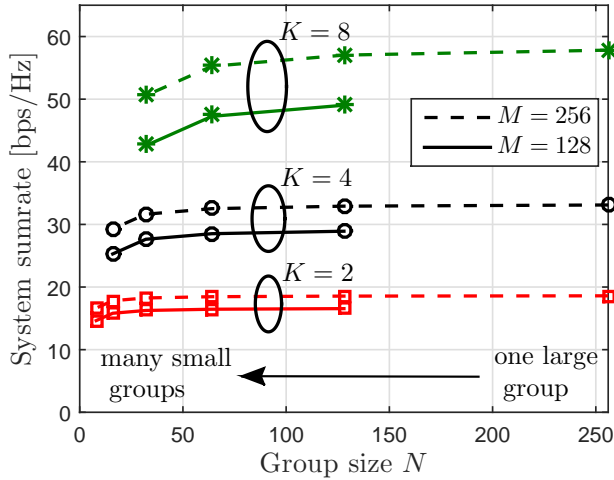
Now we can consider the case when  $N \gg 2K$ . Even for a large number of user pairs, this case is feasible due to the fundamental assumption of a large number of relay units,  $M \gg 1$ . The array gain term from (30) then becomes

$$\frac{M}{N} \frac{N - 2K}{2K} \epsilon \approx \frac{M}{2K} \epsilon, \quad (31)$$

that is, the array gain becomes independent of  $N$ . This insight is of fundamental importance for practical deployments of LS-MTW systems. What it implies is that, in the regime with a large number of relays  $M$ , tight cooperation in information processing between all  $M$  relays is not necessary. Instead, small, independent groups of tightly cooperating relays can be formed, and such a setup experiences only a marginal degradation of system sumrate compared to the case when all relays are cooperating. If we substitute the notion of a tightly cooperating relay group with a more specific notion of a MIMO relay, we can conclude that a single massive MIMO relay performing ZF can be substituted with several simpler and cheaper MIMO relays with smaller numbers of antennas, with a negligible reduction in system performance.

The presented observations are corroborated by simulations, results of which are presented in Fig. 2, where the lower bound (30) is compared to simulated system sumrate, averaged over channel realizations. It is assumed that the total transmit power in the system, which we denote by  $P_{tot}$ , is split between users and relays in two equal parts, so  $P_U = P_{tot}/4K$  and  $P_R = P_{tot}/2L$ . The results show an excellent match between the theoretical lower bound (30) and simulations. Moreover, it is clearly demonstrated how substituting one large group ( $N = M$ ) with several smaller and independent groups of relays introduces only a slight degradation of sumrate (in the most extreme cases, sumrate degrades by 10.5% to 12.5% for the setups considered).

In order to gain deeper understanding of tradeoffs encountered in the design of LS-MTW systems, in addition to sumrate, we also need to take into account the cost of enabling cooperation between the relays. This cost, which we denote by  $C$ , quantifies the resources spent (e.g. energy, bandwidth) or



**Figure 2:** Theoretical and simulated ergodic sumrate of a LS-MTW system with ZF in uplink and downlink. Markers: simulation results, full and dashed lines: theoretical lower bounds. Minimum theoretical relay group size =  $2K + 1$ , maximum =  $M$ .  $SNR_u = SNR_d = 10$  dB and no power imbalance assumed.

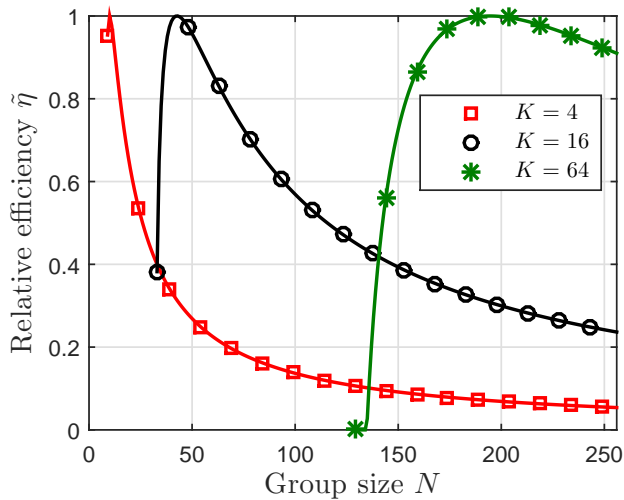
penalties in system performance incurred (e.g. latency) when CSI and up-link/downlink data are exchanged inside a relay group. In particular, we focus on resources that can be *reused* between groups. An example system setup would feature relays inside a group exchanging CSI and data with the CGP over dedicated short-range wireless links and employing frequency division multiplexing (FDM). With enough physical separation between individual groups, the short range of intra-group backhaul links would mean that the bandwidth dedicated for cooperation can be reused between groups. Moreover, the use of FDM implies that this bandwidth is proportional to the group size:

$$C = c_{\text{BW}}N \text{ [Hz]}, \quad (32)$$

where  $c_{\text{BW}}$  is the bandwidth of the frequency slot allocated for one user-CGP link. As discussed previously, for  $N \gg 2K$ , sumrate is independent of  $N$ . Therefore, the cooperation efficiency of the described system,

$$\eta = \frac{R}{C}, \quad (33)$$

increases for decreasing  $N$  when  $N$  is large.



**Figure 3:** Relative cooperation efficiency with reusable cooperation resources.  $SNR_u = SNR_d = 10$  dB,  $M = 256$ .

Relative cooperation efficiency function  $\tilde{\eta} = \eta / \max\{\eta\}$  is shown in Fig. 3, where  $N \in \mathbb{N}$  but the constraint  $L \in \mathbb{N}$  is relaxed. These results support the notion that using one large cooperating group is a suboptimal strategy from the point of view of cooperation efficiency, especially for low values of  $K$ .

## 6 Conclusion

We have analyzed a multipair two-way relay system with a large number of relays. The relays are assumed to form groups inside which data and channel state information is exchanged, and processing is done independently from other groups. Assuming that the groups perform zero-forcing and that the SNR is large, we derive a closed-form expression for a tight lower bound on the system sumrate. An asymptotic analysis of the bound shows that the sumrate is essentially independent from group size  $N$  when  $N \gg 2K$ . This implies that one large group of cooperating users can be substituted with several smaller groups, with no significant impact on performance. We extend this result to take into account the efficiency of information exchange that supports relay cooperation. It is shown that using several smaller relay groups is more efficient than the use of a single large group, if the resource used for intra-group



information exchange is reusable between groups.

## **Acknowledgment**

The authors would like to thank the Intel Corporation for providing the funds for the research, which was conducted as a part of the SRC MSR-Intel Research Project P28832, “Coordination in Distributed Multi-User High-Performance Dense Networks”.

## Appendix A: Proof of Lemma 1

By taking into account Jensen's inequality

$$\mathbb{E}f(X) \geq (\leq) f(\mathbb{E}X),$$

$f(X)$  convex (concave), we can form a chain of (in)equalities

$$\begin{aligned} \mathbb{E} \log_2 \left[ 1 + \left( \sum_i \frac{1}{\sqrt{\psi_i}} \right)^2 \right] &> \frac{1}{\ln 2} \mathbb{E} \ln \left[ \left( \sum_i \frac{1}{\sqrt{\psi_i}} \right)^2 \right] \\ &= \frac{2}{\ln 2} \mathbb{E} \ln \left( \sum_i e^{-\frac{1}{2} \ln \psi_i} \right) \stackrel{a)}{\geq} \frac{2}{\ln 2} \ln \left( \sum_i e^{-\frac{1}{2} \mathbb{E} \ln \psi_i} \right) \\ &\stackrel{b)}{\geq} \frac{2}{\ln 2} \ln \left( \sum_i e^{-\frac{1}{2} \ln \mathbb{E} \psi_i} \right) = \log_2 \left[ \left( \sum_i \frac{1}{\sqrt{\mathbb{E} \psi_i}} \right)^2 \right], \end{aligned}$$

where inequality  $a)$  follows from convexity of  $\ln \sum_i e^{y_i}$  on  $\mathbb{R}^n$ , so

$$\mathbb{E} \ln \sum_i e^{y_i} \geq \ln \sum_i e^{\mathbb{E} y_i}.$$

Inequality  $b)$  follows from concavity of  $\ln(\cdot)$  and from the fact that  $e^{-z}$  is monotonically decreasing, which yields

$$e^{-\frac{1}{2} \mathbb{E} \ln \psi_i} \geq e^{-\frac{1}{2} \ln \mathbb{E} \psi_i}.$$



# Bibliography

- [1] C. Wang, H. Chen, Q. Yin, A. Feng and A. F. Molisch, "Multi-User Two-Way Relay Networks with Distributed Beamforming," in *IEEE Transactions on Wireless Communications*, vol. 10, no. 10, pp. 3460-3471, October 2011.
- [2] J. Zhang, F. Roemer and M. Haardt, "Distributed beamforming for two-way relaying networks with individual power constraints," 2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), Pacific Grove, CA, 2012, pp. 542-546.
- [3] H. Q. Ngo and E. G. Larsson, "Large-Scale Multipair Two-Way Relay Networks with Distributed AF Beamforming," in *IEEE Communications Letters*, vol. 17, no. 12, pp. 1-4, December 2013.
- [4] S. Jin, X. Liang, K. Wong, X. Gao and Q. Zhu, "Ergodic Rate Analysis for Multipair Massive MIMO Two-Way Relay Networks," in *IEEE Transactions on Wireless Communications*, vol. 14, no. 3, pp. 1480-1491, March 2015.
- [5] T. V. T. Le and Y. H. Kim, "Power and Spectral Efficiency of Multi-Pair Massive Antenna Relaying Systems With Zero-Forcing Relay Beamforming," in *IEEE Communications Letters*, vol. 19, no. 2, pp. 243-246, Feb. 2015.
- [6] Y. Dai and X. Dong, "Power Allocation for Multi-Pair Massive MIMO Two-Way AF Relaying With Linear Processing," in *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 5932-5946, Sept. 2016.
- [7] M. Liu, J. Zhang and P. Zhang, "Multipair Two-Way Relay Networks with Very Large Antenna Arrays," 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall), Vancouver, BC, 2014, pp. 1-5.

- 
- [8] C. Kong, C. Zhong, M. Matthaiou, E. Björnson and Z. Zhang, “Multi-pair two-way AF relaying systems with massive arrays and imperfect CSI,” 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016, pp. 3651-3655.
  - [9] E. G. Larsson, O. Edfors, F. Tufvesson and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” in IEEE Communications Magazine, vol. 52, no. 2, pp. 186-195, February 2014.
  - [10] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu and F. Tufvesson, ”Reciprocity Calibration for Massive MIMO: Proposal, Modeling, and Validation,” in IEEE Transactions on Wireless Communications, vol. 16, no. 5, pp. 3042-3056, May 2017.
  - [11] R. Rogalin et al., “Scalable Synchronization and Reciprocity Calibration for Distributed Multiuser MIMO,” in IEEE Transactions on Wireless Communications, vol. 13, no. 4, pp. 1815-1831, April 2014.
  - [12] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Now Publishers, 2004.