LUND UNIVERSITY

Handling nonlinear power amplifiers in massive MIMO system-level tradeoffs

Muneer, Sidra

Handling Nonlinear Power Amplifiers in Massive MIMO: System-level Tradeoffs

Handling Nonlinear Power Amplifiers in Massive MIMO: System-level Tradeoffs

Licentiate Thesis

Sidra Muneer

LUND
UNIVERSITY

Thesis for the degree of Licentiate
Thesis advisors: Prof. Liang Liu
Opponent: Dr. Bo Göransson

To be presented, with the permission of the Faculty of Engineering LTH of Lund University, for public criticism at E Building) on Monday, the 15th of December 2025 at 09:15.

Organization
LUND UNIVERSITY
Electrical and Information Technology
Box 118
SE-221 00 LUND
Sweden

Document name
LICENTIATE THESIS

Date of disputation
2025-12-15

Sponsoring organization
Ericsson AB
Excellence Center at Linköping – Lund in Information Technology (ELLIIT)

Author(s)
Sidra Muneer

Title and subtitle
Handling Nonlinear Power Amplifiers in Massive MIMO: System-level Tradeoffs

Abstract
With the rapid growth of mobile communication systems, massive multiple-input multiple-output (MIMO) technology plays a key role in enhancing communication systems to meet growing data traffic demands. However, scaling the technology further challenges the system designers to balance design tradeoffs in order to provide efficient and low-cost communication systems. This requires efficient hardware as well as low-complexity digital signal processing techniques at the physical layer. Power amplifiers are most critical components in radio transmitters as they account for most of the power consumption and contribute significantly to system cost and complexity. Although massive MIMO reduces the per-antenna transmit power requirement, achieving high efficiency in power amplifiers remains a challenge due to the trade-off between efficiency and linearity and signal attributes such as high peak-to-average power ratio.

This work investigates the behavior of PAs in this regard particularly the nonlinear distortion emissions from base stations and explores the digital signal processing methods that improve PA efficiency in OFDM based massive MIMO systems. Linear precoding techniques have been considered in frequency-selective channels scenario with a particular emphasis on low complexity signal peak reduction methods that exploit the large number of antennas. By combining distortion-based peak reduction methods with compensation strategies and regularized precoding, the proposed approach mitigates the performance loss associated with such methods in the presence of nonlinear PAs. The results demonstrate improved performance while maintaining low complexity, enabling the use of cost-effective and low-power analog components.

The findings contribute towards the development of more sustainable and energy-efficient massive MIMO networks, supporting future wireless communication systems that improve both performance and efficiency.

Key words
Massive MIMO, OFDM, Base Station, System Performance, Nonlinear Power Amplifiers, Peak-to-average Power Ratio Reduction, Clipping Distortion, Low Complexity

Classification system and/or index terms (if any)

Language
English

Supplementary bibliographical information

ISBN (print)
978-91-8104-778-3

ISSN and key title
Series of Licentiate and Doctoral Thesis; ISSN 1654-790X, No. 191

ISBN (pdf)
978-91-8104-779-0

Recipient's notes

Number of pages
100

Price

Security classification

Signature _____ Date 2025-11-20

Handling Nonlinear Power Amplifiers in Massive MIMO: System-level Tradeoffs

Licentiate Thesis

Sidra Muneer

LUND
UNIVERSITY

Cover illustration front: Difference between out-of-band power distribution in space for different precodings

Funding information: The thesis work was financially supported by the Ericsson AB through a bilateral research grant and Excellence Center at Linköping-Lund in Information Technology (ELLIIT).

Faculty of Engineering LTH, Electrical and Information Technology

ISBN: 978-91-8104-778-3 (print)
ISBN: 978-91-8104-779-0 (pdf)

This thesis is produced by using the LATEX Documentation System
Printed in Sweden by Tryckeriet i E-huset, Lund. Department of Electrical and Information Technology Lund University, Box 118, SE-221 00 Lund, Sweden 2025

Dedicated to my kids,
Eshaal, Sualeh
and Danial

Popular Science

The mobile phones we use today, with their unlimited apps, entertainment and uninterrupted connectivity are a result of decades of technological breakthroughs. These advances range from faster computing capabilities to smarter, more user-friendly devices. While we benefit from these systems today, we must also design for tomorrow. Sustainable, energy-efficient technologies are essential to ensure that these advancements remain viable for generations to come. Whereas, reliable and time critical connectivity enables additional use cases such as industrial automation, extended reality, self-driving vehicles and artificial intelligence.

Behind every mobile connection are base stations, the towers your phone connects to. These stations must be reliable, low-cost and energy-efficient to reduce their carbon footprint. At the heart of each base-station transmitter is a key component, the power amplifier. This part is responsible for sending radio waves through antenna over long distances, but it has also been one of the biggest consumers of energy. If not carefully handled, it can also generate unwanted emissions that interfere with other signals and also degrade the user experience.

In modern Fifth Generation (5G) systems, massive Multiple-Input Multiple-Output (MIMO) technology is deployed which uses many antennas to improve wireless system performance. It has changed how amplifiers are deployed in the base stations as it requires many smaller components instead of bulky ones used in previous generation cellular communication systems. This can provide similar coverage while reducing amplifier's output power requirements. However, to operate efficiently, amplifiers must handle a wide range of signals which can increase the unwanted emissions from these amplifiers. The study of the behavior of these systems, considering practical system constraints, in realistic environments forms the core of current research. It also helps understand the system behavior in order to define standards and regulatory requirements.

This work therefore explores different system design choices in massive MIMO, to understand the behavior of unwanted emissions from amplifiers in complex environments. New methods have also been investigated, which benefit from large number of antennas in the system to help power amplifiers operate more efficiently, while reducing these emissions. The result is more reliable wireless communication and reduced costs, helping to build greener and more sustainable mobile networks for the future.

Acknowledgments

I would like to express my heartfelt gratitude to everyone who has been part of my research journey. In particular, I am deeply thankful to my supervisor Prof. Liang Liu for his constant support and guidance throughout this journey. His expertise, perseverance and commitment have taught me a great deal, from reshaping my analytical capabilities to overcoming my fears of thinking out of box.

I am also sincerely grateful to Prof. Ove Edfors for generously sharing his profound wireless expertise and his remarkable jack-of-all-trades insights throughout this journey. My special thanks go to Prof. Henrik Sjöland for his kind and invaluable guidance during the course of this research and for enriching us with his exceptional analog knowledge. I would also like to thank Prof. Liesbet Van der Perre for her kind and knowledgeable guidance during the early years of my research and Prof. Fredrik Tufvesson for his support. I am incredibly thankful to Jesus for all his help, valuable time, and the many fruitful discussions we've had. I am also very grateful to Sara for sharing her code during the early days of my PhD and for her help with channels. Thank you, Bakhtash, Iman and Ashkan for always being ready to help, whenever needed. My heartfelt thanks also goes to Rikard, Masoud, Muhammad Attari, Mojtaba, Hamid, Steffen, Lucas, Arturo and Ilayda for filling our time at LTH with laughter and good memories. Thank you, Haorui, Zahra, Fatemeh, Hadieh and Hanieh for all the wonderful moments we shared, memories I will always cherish. Also, my best wishes to all the new faces at IES, I wish you all the very best in your journey ahead. Thanks to Josep for corrections. Finally, I would like to pay gratitude to exceptional administrative and technical staff at EIT for their support.

Most importantly, I would like to thanks my husband Khurram for being a source of constant support and love all these years and to my kids who are always passionate about my endeavors and have always been source of my strength. My family like friends here in Sweden who have been keeping my sanity in check all these years. My parents for their prayers and encouragement and my siblings for their constant support.

Finally, I would also like to thank everyone at Ericsson who has been involved, for their help and for generously sharing their expertise throughout my research journey.

Contents

List of Publications

This thesis is based on the following publications, referred to by their Roman numerals:

I **Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption**
Sidra Muneer, Liang Liu, Ove Edfors, Henrik Sjoland, Liesbet Van Der Petre
2020 54th Asilomar Conference on Signals, Systems, and Computers Year: 2020

II **System Design and Performance for Antenna Reservation in Massive MIMO**
Sidra Muneer, Jesus Rodriguez Sanchez, Liesbet Van der Perre, Ove Edfors, Henrik Sjöland, Liang Liu
2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall) Year: 2022

III **Low Complexity Clipping Distortion Compensation for PAPR Reduction in Massive MIMO-OFDM for Frequency Selective Channels**
Sidra Muneer, Henrik Sjöland, Ove Edfors, Liang Liu
IEEE Transactions on Green Communications and Networking Year: 2025

List of Acronyms and Abbreviations

1G First Generation

4G Fourth Generation

3GPP Third Generation Partnership Project

5G Fifth Generation

6G Sixth Generation

ACLR Adjacent Channel Leakage Ratio

AI Artificial Intelligence

BER Bit Error Rate

BS Base Station

CCDF complementary cumulative distribution Function

CSI Channel State Information

DAC Digital-to-Analog Converter

DC Direct Current

DPD Digital Pre-Distortion

EVM Error Vector Magnitude

FFT Fast Fourier Transforms

IFFT Inverse Fast Fourier Transforms

IID Independent Identically Distributed

IoT Internet of Things

IUI Inter-user Interference

LIS Large Intelligent Surface

LoS Line-of-Sight

LTE Long Term Evolution

MIMO Multiple-Input Multiple-Output

MMSE Minimum Mean Square Error

MRT Maximum Ratio Transmission

NR New Radio

OOB Out-Of-Band

OFDM Orthogonal Frequency Division Multiplexing

PA Power Amplifier

PAPR Peak-to-Average Power Ratio

RAN Radio Access Network

RF Radio Frequency

SNR Signal-to-Noise Ratio

TDD Time-Division Duplex

Tx Transmitter

ULA Uniform Linear Array

V2X Vehicle-to-Everything

ZF Zero Forcing

Notation

The following notation has been adopted throughout this thesis:

- All upper-case bold-face letters represent matrices.
- All lower-case bold-face letters represent vectors.
- $(.)^{\mathtt{H}}$ and $(.)^{\mathtt{T}}$ represents Hermitian conjugate and transpose respectively.
- $\mathbb{E}\{.\}$ represents expectation operator.
- $\mathbb{F}\{.\}$ represents Fourier transform.
- $\mathbb{C}$ represents complex domain.
- $\circ$ represents element-wise matrix product.
- $\boldsymbol{I}_K$ represents an identity matrix of dimension $K \times K$.
- $||.||_{\mathtt{F}}$ represents Frobenius norm.

Handling Nonlinear Power Amplifiers in Massive MIMO: System-level Tradeoffs

Chapter 1

Introduction

With the advent of cellular communication systems, from the early First Generation (1G) networks to the more sophisticated 5G systems of today, significant advancements have been made to enable faster and more reliable communication. These systems have facilitated a plethora of applications that were previously impossible.

New 5G networks can deliver order of magnitudes faster data rates and energy efficiency than previous Fourth Generation (4G) systems [1]. They also provide up to five times lower latency, enabling new possibilities such as the Internet of Things (IoT), autonomous vehicles, and immersive applications that have the potential to revolutionize industries, education, entertainment and training. Additionally, the enhanced connectivity of 5G serves as a cornerstone for the development of smart cities, where interconnected devices and systems can collaborate to improve urban living through smarter infrastructure, efficient resource management, and enhanced public services.

As we stand on the brink of Sixth Generation (6G), the potential for future transformation is large. With the new emerging technologies such as Large Intelligent Surface (LIS) [2] and distributed multiple antenna systems [3], which can enhance system capacity and can pave ways to new possibilities. These advancements aim to meet the ever-increasing data traffic demands and support future applications.

With the global push towards reducing carbon footprints and achieving sustainability goals, the increased energy demands of telecommunication operators pose a challenge. They need to find ways to power expanded in-

frastructures with renewable energy sources and implement energy-saving technologies to mitigate environmental impacts. To address these challenges, several strategies at various level of system design can be adopted. These include:

1. Energy-efficient hardware designs that reduce the power consumption of network elements.
2. Adaptation of advanced algorithms to dynamically adjust power usage based on network demand.
3. Optimization of network architecture to balance performance and energy consumption effectively.
4. Deployment of renewable energy sources and energy storage solutions to power network infrastructure.
5. Research into new materials and technologies that improve the energy efficiency of antennas and transceivers.

By tackling the energy challenges, the telecoms industry can ensure that the benefits of new technologies are realized without compromising on its energy efficiency and sustainability.

Massive MIMO is a key technology in the evolution of core radio access networks for 5G, providing higher spectral efficiency for network operators with limited frequency spectrum licensing. By deploying large number of antennas at the base station, the capacity of a communication link can be increased significantly, enabling multi-user communication with high spatial multiplexing capabilities. This makes it essential to meet the growing connectivity demands of current and future wireless communication systems. As shown in Fig. 1.1 from Ericsson's June 2025 technical report, global 5G subscriptions are expected to reach 6.3 billion in 2030, and will make up two-thirds of all mobile subscriptions. With Artificial Intelligence (AI) on rise, the requirements for higher data rates as well as more sophisticated data usage scenarios are expected. Consequently, it is crucial to advance massive antenna systems to optimize cost, performance, energy efficiency, and size of these systems to meet these increased demands. Besides the core network infrastructure of telecom operators and 5G stand alone implementations, the massive MIMO technology extends its applications into various other domains including extremely large aperture arrays, holographic massive MIMO, localization, sensing and tracking applications [4].

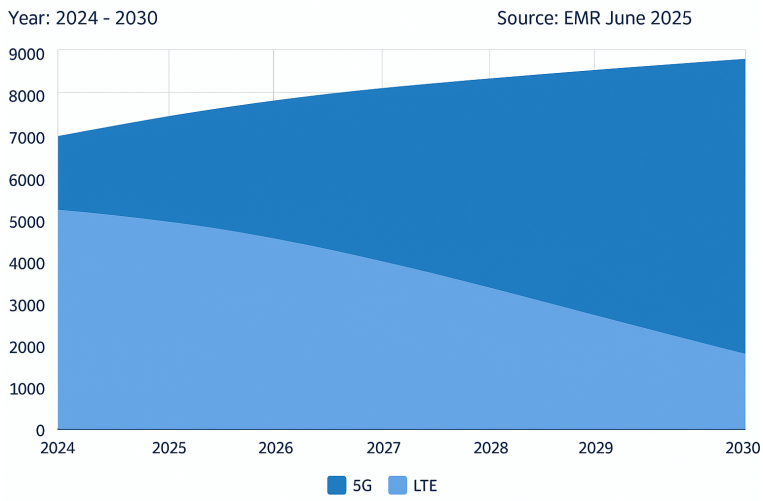


Figure 1.1: Expected Growth in 5G Mobile Subscriptions including all devices Worldwide (Unit: Million) Source: Ericsson Mobility Report June 2025 [1]

Although the current 5G technology provides support for IoT and Vehicle-to-Everything (V2X) communications, these applications coupled with AI capabilities in future, will require more efficient and robust communication systems to open up scalable opportunities combined with cloud computing infrastructures [1].

Spectral Efficiency

Spectral efficiency refers to the effective use of bandwidth resource. It is measured in bits/s/Hz. The multi-user MIMO system where multiple users can be served simultaneously using all frequency resources is considered to achieve high spectral efficiency.

The Power Amplifier (PA) in radio base stations is crucial for amplifying low-power Radio Frequency (RF) signals to levels suitable for transmission over long distances, extending the coverage area. As PAs consume a significant portion of the base station's power, their efficiency directly impacts the overall energy consumption and operational costs [5]. The efficiency of a PA refers to its ability to convert the supplied Direct Current (DC) power into useful RF output power, ideally operating close to its maximum output region. Inefficient PA operation results in significant power losses in the form of heat dissipation, which not only reduces the overall energy efficiency but also increases the cooling and maintenance costs at base sta-

tion sites. According to [5] in 2010, PA accounts for $50 - 80\%$ of power consumption in a base station. With the 5G technology, this has been improved by introducing Transmitter (Tx) micro sleep modes [6], however, PA efficiency still remains crucial in reducing the carbon footprint of the base stations. High-efficiency PAs reduce power leakage and mitigate excessive heat generation, minimizing the need for extensive cooling systems. Efficient PAs also enhance network performance by providing higher average transmit power, hence improving signal quality (with linearization) and therefore data rates and cellular coverage. The reduced distortions also results in less bit-error rates and required retransmissions, hence contributing to improved reliability and lower latency, crucial for real-time applications. Economically, improved PA efficiency leads to cost savings for network operators, while environmentally, it supports sustainability by reducing energy consumption. Overall, efficient PAs are vital for the energy efficiency and sustainable operation of radio base stations.

Power Amplifier Efficiency

The efficiency of a PA refers to its ability to convert the supplied DC power into useful RF output power, ideally operating close to its maximum output region. Inefficient PA operation results in significant power losses in the form of heat dissipation, which not only reduces overall energy efficiency but also increases the cooling and maintenance costs at base station sites.

Besides efficiency, quality amplification without distortion is essential for maintaining signal integrity, as non-linear amplification can degrade signal quality and cause interference. Moreover, PAs must support wide frequency ranges and multiple bands to accommodate modern wireless standards, enabling seamless connectivity across various services. Their ability to handle advanced modulation schemes and wide dynamic ranges is critical for high data throughput, spectral efficiency, and network capacity.

In practice, PAs exhibit a trade-off between linearity and efficiency, where improving one typically compromises the other. Since some compromise on efficiency is inevitable, PA nonlinearity is usually addressed with compensation techniques. Traditionally, the nonlinear behavior of analog front-end components is characterized and compensated within the digital signal processing domain.

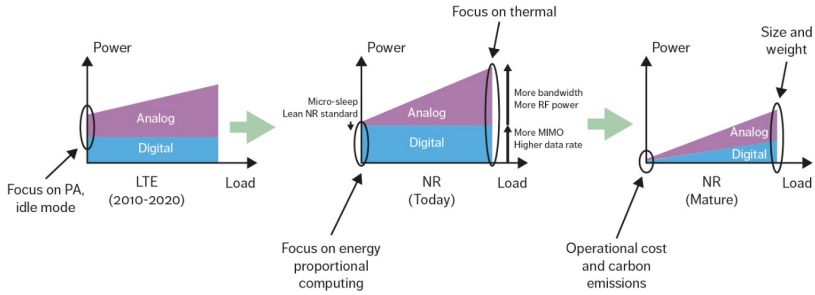


Figure 1.2: Power balance of Analog and Digital Part of Radios as 5G-NR evolves (Source:Ericsson Technology Review [6]).

1.1 Research Motivation

As discussed, the efficiency-linearity tradeoffs in power amplifiers require to tackle the challenge of PA nonlinearity in wireless communication systems. Nonlinearity occurs when a PA amplifies an input signal in a manner that distorts its original waveform, such that the output is not directly proportional to its input. This results in undesirable effects in the transmit signal such as amplitude and phase distortion. These distortions degrade the quality of the transmitted signal, causing errors in data transmission and increased adjacent channel interference hence compromising the overall performance and efficiency of the network. Addressing power amplifier nonlinearity is therefore essential for improved performance in modern communication systems along with optimizing PA efficiency. This can require complex digital signal processing techniques such as Digital Pre-Distortion (DPD) and crest factor reduction [7].

Power Amplifier Nonlineairty

> Nonlinearity occurs when a PA amplifies an input signal in a manner that distorts its original waveform, such that the output is not directly proportional to its input. This results in undesirable effects in the transmit signal such as amplitude and phase distortion.

With the development of massive MIMO technology, the number of antennas increase, leading to increased signal processing requirements in analog and digital front-ends (e.g. DPD, Digital-to-Analog Converter (DAC), PAs, Mixers etc.). Thanks to the significant array gain provided by a large number of antennas, the required output power per antenna can be reduced,

which relaxes the requirements on front-end components [6] e.g. allows for the use of lower-cost PAs. However, the digital signal processing workload increases significantly, for example, to manage the nonlinearity, each PA requires individual DPD. As shown in Fig.1.2, as the technology matures, the power consumption in the digital domain has become dominant, resulting in higher overall system power usage.

The current 5G New Radio (NR) technology enables dynamic power savings for PAs through base station sleep modes, where PAs are turned off when not in use [6]. Nonetheless, digital signal processing platforms generally remain powered on, and the increased complexity of advanced digital signal processing algorithms has led to higher digital power consumption. Therefore, there is a critical need to develop low-complexity digital processing techniques to reduce power consumption in this domain. These techniques can also leverage the benefits of large number of antennas in the system. In this work, we therefore focus on low-complexity methods designed to improve energy efficiency and system performance leveraging massive MIMO systems.

1.2 Thesis Contributions

This work advances the domain of PA handling techniques in massive MIMO including the system level tradeoffs and efficiency enhancement when using the digital signal processing techniques. With the technology scaling, the complexity and scalability of these techniques are of concern due to the risk of increase in digital power consumption [8]. The following list describes the focus of this work.

- Study and analyze the behavior of nonlinear distortion emissions from massive MIMO base stations for different user configurations in realistic channel environments.
- Study and analyze the waveform characteristics to study PA efficiency-linearity tradeoffs in massive MIMO. In particular, to quantify the system level impact of low-end hardware on performance for different system configurations.
- Study and analyze the system-level tradeoffs for using digital signal processing techniques to improve system efficiency and linearity, particularly in the context of massive MIMO.

- Development of low-complexity and improved Peak-to-Average Power Ratio (PAPR) reduction methods which leverage the large number of antennas in massive MIMO and caters to frequency selective channels.

1.3 Thesis Outline

Chapter 2 presents background knowledge in the domain of massive MIMO, PAs and associated digital signal processing techniques. Chapter 3 presents a deeper outlook on our research in this domain and justifies our contributions.
Paper I presents PA nonlinearity behavior and power consumption tradeoffs in massive MIMO systems. Paper II presents system design and performance for antenna reservation technique for Massive MIMO. Paper III presents a low complexity method for PAPR reduction in frequency selective massive MIMO.

Chapter 2

Background

2.1 Massive MIMO Systems

Massive MIMO refers to the technology in wireless communication constituting of base-station with very large number of transmission/reception antennas which allows muti-user spatial multiplexing and beamforming to achieve very high system capacity. Massive MIMO is in continuous evolvement for integration into the existing networks as well as for massive deployments for future 6G networks [1].

Beamforming in traditional sense refers to directing signal from multiple antennas to a specific direction by controlling its magnitude and phase so that the signals can add up constructively in that direction. There are three prevalent beamforming architectures in wireless communications namely analog, digital and hybrid beamforming as shown in Fig. 2.1.

Spatial multiplexing refers to a technique, where independent data streams are transmitted using same time and frequency resource, but over different spatial resources. It thus requires large number of antennas to provide unique spatial signatures for each signal path. The more prevalent and flexible architectures for this purpose are digital/hybrid beamforming [4].

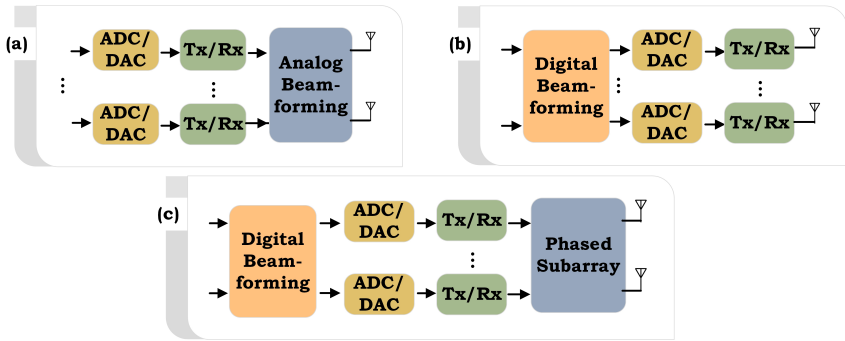


Figure 2.1: Illustrations of the different beamforming system architectures: (a) analog beamforming, (b) full-digital beamforming, and (c) hybrid beamforming.

Beamforming

Beamforming refers to the directing of signal from multiple antennas to a specific direction by controlling its magnitude and phase so that signals can add up constructively in that direction. There are three prevalent beamforming architectures namely analog, digital and hybrid beamforming.

2.1.1 Beamforming Architectures

Different beamforming architectures provide a different balance between spatial multiplexing capabilities, spectral efficiency, system complexity, cost and power consumption. These are shown in Fig. 2.1. In analog beamforming, a dedicated RF chain is typically used with many phase shifters for each antenna. This can direct a signal to a specific spatial direction to create a beam. Whereas, in digital beamforming, each beam phase and magnitude is formulated in digital signal processing domain which provides more flexibility and spatial multiplexing opportunities. However, this architecture requires a dedicated signal processing and RF chain for each antenna. Thus its complexity scales with the number of antennas, requiring very high interconnection data-rates and synchronization [3]. It is primarily used in cases where high spectral efficiency is prioritized such as 5G sub-6GHz multi-user MIMO systems.

Spatial Multiplexing

> Spatial multiplexing refers to a technique where independent streams are transmitted using same time and frequency resource but over different spatial resources. It thus requires large number of antennas to provide unique spatial signatures for each signal path.

This is one of the reasons why hybrid beamforming architectures have been more prevalent, where the digital and analog beamforming can be combined together to reduce the cost and provide implementation scalability, however it limits the spatial multiplexing capabilities and hence spectral efficiency in most cases [9].

Massive MIMO requires digital beamforming architecture with large number of antennas for transmission at base stations to provide multi-user spatial multiplexing and channel hardening. In its essence, it provides a beamforming gain by diversifying signal from multiple antennas while using all the space-time resource for all the users hence providing excellent spectral efficiency. Massive MIMO thrives on channel state information (CSI). This requires that an accurate channel information is acquired at the base station, which can be done via pilot based schemes in the uplink.

A generic physical layer architecture of a 5G Radio Access Network (RAN) is shown in Fig. 2.2. It consists of a baseband processing unit and a radio front-end unit integrated closer to antennas. With massive MIMO, these units are integrated into one single unit to perform functions such as modulation, coding, precoding, resource allocation, Fast Fourier Transforms (FFTs), framing, filtering, and digital signal processing techniques aimed at improving system linearity and efficiency [1]. Massive MIMO base station deployments often employ 128 or 256 antennas to balance cost and performance tradeoffs [1].

A direct-conversion transmitter architecture (or zero-IF) can be considered for the design of massive MIMO transceiver stage due to its simplicity, cost, and scalability over traditional RF heterodyne architecture. In such a Tx architecture, the digital signal is directly converted to analog using high speed RF DACs and filters. It is further modulated onto a carrier frequency via a mixer and a local oscillator. The signal is subsequently amplified before radiating through an antenna array. While the specific front-end components depend on the design specifications and the bandwidth requirements, the overall objective is to realize highly power-efficient, linear, and wide-band transmitters.

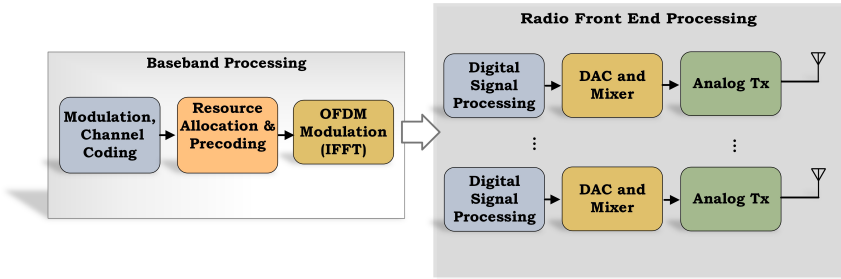


Figure 2.2: Simplified 5G RAN System Architecture

2.1.2 System Model

A simplified down-link baseband equivalent system model of massive MIMO base station with nonlinear PAs in M antennas serving K users with set of S subcarriers is shown in Fig. 2.3.

The received vector $\boldsymbol{r}_s$ for user K for each subcarrier s is given as,

$$\boldsymbol{r}_s = \boldsymbol{H}_s \boldsymbol{y}_s + \boldsymbol{n}_s \tag{2.1}$$

where $\boldsymbol{H}_s \in \mathbb{C}^{K \times M}$ is a complex propagation channel matrix in frequency domain for each subcarrier s, $\boldsymbol{y}_s \in \mathbb{C}^{M \times 1}$ is a frequency domain equivalent of transmitted signal $\boldsymbol{y}$, and $\boldsymbol{n}_s \in \mathbb{C}^{K \times 1}$ is a vector of Independent Identically Distributed (IID) circularly-symmetric zero-mean complex Gaussian noise at each subcarrier s.

2.1.3 Precoding

A process of removing the effect of channel on the signal is called equalization. Traditionally, it has been used in receivers to combat the effect of channel on the transmitted signal. Precoding is a pre-equalization technique used in transmitters so that the receiver does not require complex equalization and can therefore be simplified. However, this requires that the Channel State Information (CSI) exists at the transmitter. In Time-Division Duplex (TDD) supported systems such as Long Term Evolution (LTE), this is achieved through uplink pilot transmission from the user equipment to the Base Station (BS). The BS then estimates the downlink channel by exploiting the principle of channel reciprocity, which assumes that the uplink and downlink channels are reciprocal within the channel coherence time. Massive MIMO systems can deploy different precoding

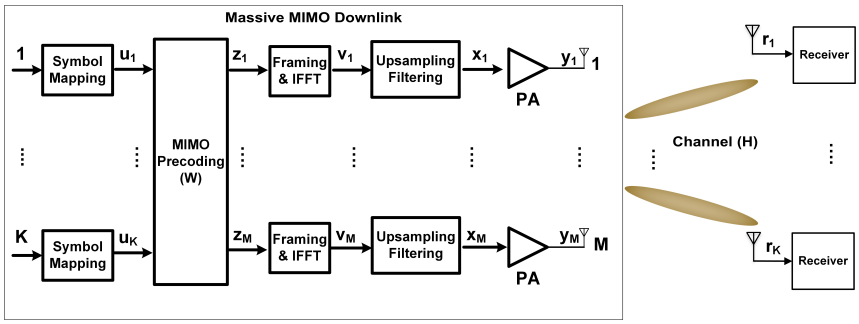


Figure 2.3: Simplified Functional Block Diagram of a Massive MIMO Downlink Baseband Equivalent System

techniques for digital beamforming. There have been plethora of precoding techniques proposed for massive MIMO, but only linear precoders have been considered in this work as these provide reasonable performance with low-complexity in massive MIMO [10]. These are described below.

Maximum Ratio Transmission

Maximum Ratio Transmission (MRT) is a simple technique which uses channel information at the transmitter to maximize the Signal-to-Noise Ratio (SNR) in the direction of users. It only requires channel conjugate transpose to do so which is simple to implement in hardware. Mathematically,

$$\boldsymbol{W}_{\mathtt{MRT}} = \boldsymbol{H}^{\mathtt{H}} \tag{2.2}$$

It is also considered close to optimal in very low interference scenarios and more robust against CSI errors as small fluctuation or noise does not result in huge variation in the computation of precoder [11]. Although simple, the major drawback of using MRT is that it does not minimize Inter-user Interference (IUI). IUI is an interference caused by simultaneous transmission of users data in the same time-frequency resource. More specifically, it is the presence of signal at the receiver besides the signal intended for that receiver in the absence of system noise and nonlinearities.

Zero Forcing

Zero Forcing (ZF) is an interference cancellation technique which minimizes IUI but requires a matrix inverse calculation for this purpose. Finding a matrix inverse for very large dimensional matrices is computationally in-

tensive. It also amplifies noise in case of ill-conditioned matrices. Mathematically,

$$\boldsymbol{W}_{\texttt{ZF}} = \boldsymbol{H}^{\texttt{H}}(\boldsymbol{H}\boldsymbol{H}^{\texttt{H}})^{-1} \tag{2.3}$$

where, right hand side is the right Moore-Penrose inverse of $\boldsymbol{H}$ such that $\boldsymbol{H}\boldsymbol{W} = \boldsymbol{I}$. ZF completely eliminates IUI when the channel is full rank. It is considered near-optimal in high SNR scenarios, and is significantly simpler compared to using maximum likelihood detection receivers.

MMSE Precoder

A Minimum Mean Square Error (MMSE) precoder achieves a trade-off between interference cancellation and noise amplification by introducing regularization. In ZF, small singular values of $\boldsymbol{H}$ can lead to significantly large singular values in $\boldsymbol{W}$, which amplifies noise after precoding/detection. MMSE mitigates this effect by solving a regularized inverse, such that,

$$\boldsymbol{W}_{\texttt{MMSE}} = \boldsymbol{H}^{\texttt{H}}(\boldsymbol{H}\boldsymbol{H}^{\texttt{H}} + \sigma^2\boldsymbol{I}_K)^{-1} \tag{2.4}$$

where σ^2 denotes the noise variance. This regularization limits noise enhancement while still suppressing interference. At high SNR, the MMSE solution converges to ZF, whereas at low to moderate SNR it significantly outperforms ZF due to its robustness against noise amplification.

The users in a communication system can be located anywhere within the cell and may require different data rates and quality-of-service depending upon the application. Therefore, an appropriate power and resource allocation is usually required to maximize system throughput, employing strategies such as SNR maximization, proportional fairness, or max-min fairness [12].

In this thesis, we adopt an *equal power allocation*, where each user receives the same share of the total transmit power. To avoid variations in received SNR among users, we assume that all users are placed at the same distance from the base station in different directions in a far field. This user-selection scheme simplifies the evaluations, although in a more realistic scenario users would be distributed randomly across the cell, leading to unequal channel gains and requiring more sophisticated power-control.

A *precoder scaling factor* α_k is therefore adopted which ensures that the total transmit power satisfies the base station's power constraint. Let P_T denote the total available transmit power such that,

$$\mathbb{E}\{\|\boldsymbol{x}\|^2\} = P_T, \tag{2.5}$$

If $\mathbf{W}_k \in \mathbb{C}^{M \times S}$ is a precoding matrix for a user, it is scaled by factor α_k to satisfy the total power constraint.

The scaling factor is chosen as,

$$\alpha_k = \frac{1}{||\boldsymbol{W}_k||_{\mathrm{F}}} \sqrt{\frac{P_T}{K}}. \tag{2.6}$$

This implies that each user is allocated $\frac{P_T}{K}$ transmit power. In practical deployments, user scaling factors α_k may be used for transmit power allocation depending on the user location and data rate requirements, using more advanced power-control strategies.

2.2 OFDM based Massive MIMO Systems

The precoding in time-domain becomes highly complicated in massive MIMO for frequency-selective channels, because each channel tap corresponds to a convolution in time. By using Orthogonal Frequency Division Multiplexing (OFDM), in addition to its inherent robustness against multi-path fading, the use of narrow-band subcarriers converts the frequency-selective channel into many flat-fading subchannels. This enables simple per-subcarrier equalization (matrix multiplication) and efficient spatial multiplexing.

In Fig. 2.3, a block of symbols $\boldsymbol{U}$ can be defined as $\boldsymbol{U} \triangleq [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_S] \in \mathbb{C}^{K \times S}$. This block is precoded using a user interference cancellation precoding scheme such as one given in 2.3. A precoded block, $\boldsymbol{Z} \triangleq [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_S] \in \mathbb{C}^{M \times S}$ is a discrete domain signal for all subcarriers where each $\boldsymbol{z}_s \in \mathbb{C}^{M \times 1}$. The relationship can be described as,

$$\boldsymbol{Z} = \boldsymbol{W} \circ \boldsymbol{U} \tag{2.7}$$

where $\circ$ denotes element-wise matrix product and $\boldsymbol{W} \triangleq [\boldsymbol{W}_1, \ldots, \boldsymbol{W}_S] \in (\mathbb{C}^{M \times K})^S$ is a block of precoding matrices for each subcarrier where each $\boldsymbol{W}_s \in \mathbb{C}^{M \times K}$.

This above processing results in signal for each antenna which require further digital signal processing for conversion to a composite discrete time-domain signal using Inverse Fast Fourier Transforms (IFFT). A frame allocation for each antenna signal in $\boldsymbol{Z}$ includes addition of guard bands subcarriers prior to conversion.

The signal $\boldsymbol{V}$ after IFFT is given as,

$$\boldsymbol{V} = \mathbb{F}^{-1} \boldsymbol{Z}^{\mathrm{T}} \tag{2.8}$$

where, $\boldsymbol{V} \in \mathbb{C}^{S \times M}$. A cyclic prefix is then inserted to eliminate inter symbol interference at receiver and to turn linear convolution by channel, into a circular one for OFDM processing at receiver. An up-sampling and pulse shaping filter follows which converts a low pass complex discrete baseband signal to a higher sample-rate, for further conversion into analog domain for transmission.

The composite discrete time OFDM signal, $\boldsymbol{x}_m$, at each m^{th} antenna in Fig. 2.3 contains very high peaks, resulting in a very high PAPR signal. Transmitting this signal directly would require either large PAs to be able to deliver the power of the peaks or a large PA backoff which eventually results in a poor PA efficiency and high system operating costs. The peaks in the signal must therefore be reduced using PAPR reduction techniques.

Peak to Average Power Ratio

PAPR is defined as, the ratio of maximum instantaneous signal power to its mean power often expressed in decibels (dB). Mathematically,

$$\texttt{PAPR} = 10 * \log_{10} \frac{\texttt{max}\{|\boldsymbol{x}_m|^2\}}{\mathbb{E}\{|\boldsymbol{x}_m|^2\}} \tag{2.9}$$

In an OFDM system, the superposition of a large number of subcarriers generates a composite time-domain signal with potentially high PAPR. This necessitates front-end components capable of handling a wide dynamic range, thereby increasing their cost. A common way to characterize this behavior is through the complementary cumulative distribution Function (CCDF) of the PAPR. A CCDF of PAPR for OFDM system with various subcarriers is shown in Fig. 2.4. It is interpreted as a probability of PAPR greater than a certain PAPR threshold. It should be noticed that the peak occurrence in time domain signal is irregular and the probability of very high peak is very low, front-end components catering to full dynamic range of the signal often remain underutilized, which degrades overall system efficiency.

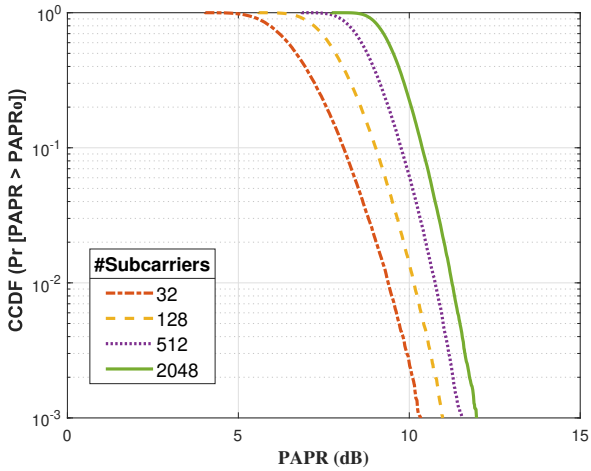


Figure 2.4: Complementary Cumulative Distribution Function for different subcarrier based OFDM signal with QPSK and 4-fold oversampling.

2.3 System Performance Measures

The information that can be transmitted over a communication channel is fundamentally limited by the channel's properties. These performance limitations are characterized by several key parameters, such as the signal-to-noise ratio (SNR) and the available bandwidth. The Shannon-Hartley capacity theorem defines the maximum achievable error-free data rate (in bits per second) for a single input single output communication link expressed as,

$$C = B \log_2(1 + \text{SNR}) \tag{2.10}$$

where C is the channel capacity and B is the bandwidth (in Hz). This theorem has two major implications: the capacity can be increased by either increasing the bandwidth or the signal power. A more practical measure for evaluating performance for modulation schemes is Spectral Efficiency (measured in bits/s/Hz), which describes how efficiently a given bandwidth is used. These high-level performance figures are directly affected by more concrete physical-layer measurements such as Error Vector Magnitude (EVM) and Bit Error Rate (BER). For example, for systems where distortion is primarily due to additive Gaussian noise, SNR is approximately inversely proportional to EVM, therefore, an increased EVM decreases system achievable capacity. Understanding the impact of system nonlinearity on these metrics under realistic channel conditions and user configurations is therefore crucial in order to determine the effective

Table 2.1: 3GPP Release 16 BS EVM Requirements [13].

Modulation scheme	Required EVM (%)
QPSK	17.5
16QAM	12.5
64QAM	8
256QAM	3.5

capacity that can be achieved in a system.

2.3.1 Error Vector Magnitude

EVM is a measure of the modulation quality or accuracy of a transmitter or receiver. It quantifies the deviation of a received symbol from its ideal reference point in the constellation diagram. For a set of symbols, the root-mean-square (RMS) EVM is commonly defined as:

$$\texttt{EVM}_{\texttt{RMS}}(\%) = \sqrt{\frac{\mathbb{E}\{|\boldsymbol{r}_k - \boldsymbol{u}_k|^2\}}{\mathbb{E}\{|\boldsymbol{u}_k|^2\}}} \times 100 \tag{2.11}$$

where $\boldsymbol{r}_k$ is the received constellation symbol and $\boldsymbol{u}_k$ is the ideal reference constellation symbols for the user. The Third Generation Partnership Project (3GPP) BS EVM requirements from [13] are given in Table 2.1. Notice that the higher-order modulation schemes have higher stringent requirements on EVM. A system supporting these modulation formats should ensure that the maximum tolerable limit is always met when using various techniques.

2.3.2 Adjacent Channel Leakage Ratio

Adjacent Channel Leakage Ratio (ACLR), also known as Adjacent Channel Power Ratio (ACPR), measures the amount of power that leaks from a transmitted signal into adjacent radio channels. It is a critical metric for regulatory compliance and network interference management. It is defined as:

$$\texttt{ACLR}(\text{dB}) = 10\log_{10}\left(\frac{P_{\texttt{adj}}}{P_{\texttt{main}}}\right) \tag{2.12}$$

where, P_{main} is the power integrated over the main allocated channel, and P_{adj} is the power integrated over a specified adjacent channel. Typically, both the left and right adjacent channels are measured, and the worst-case value is reported. A basic limit in 3GPP NR specification [13] is -45dBc.

2.3.3 Bit Error Rate

BER is the probability that a transmitted bit is received in error. It is a fundamental measure of the reliability of a digital communication system. The BER depends on factors such as the SNR, the modulation format, and the techniques used to mitigate interference and fading.

In multi-antenna systems like massive MIMO, the BER can improve as the number of base station antennas increases. This is due to the channel hardening effect, where the effective channel seen by the user becomes more deterministic and less prone to deep fades as the number of signal paths averages out the variations. However, the nonlinear distortion emissions from transmitter components (e.g., power amplifiers) require careful consideration in such systems.

2.4 Wireless Channels

Wireless channel is a medium which guides radio frequency waves from transmitter to the intended receiver. For multiple antenna systems, each antenna element presents a unique pathway for signal communication. Assuming each element is an isotropic antenna, a signal attenuation and phase shift between the base station and a receiver assuming an isotopic radiating element can be described in complex baseband form as,

$$H_s = \frac{\lambda_s}{4\pi d} e^{(-j\frac{2\pi d}{\lambda_s})} \tag{2.13}$$

referred to as free space Line-of-Sight (LoS) channel. Here d is the distance between a point in space and radiating element and λ is the wavelength.

A link budget analysis in such scenario can be done in order to determine required transmit power per antenna for a massive MIMO base station. In particular,

$$P_T = P_R + L - G_{\texttt{array}} \tag{2.14}$$

where P_T is the transmit power in dBm, P_R is the receive power in dBm, L is the path loss in dB and $G_{\texttt{array}}$ is the array gain of $10\log_{10}(M)$. For simplicity here, the antenna gains are assumed as 0dBi.

In case, where the user is not in direct line-of-sight to the base station (NLOS), the small-scale channel is often modeled as a sum of many multipath components. When the scatterers are rich and no single path dominates, a common simplifying assumption is that the channel coefficients

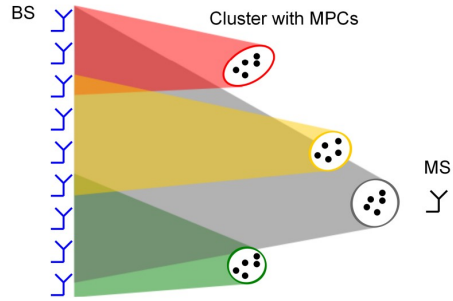


Figure 2.5: COST2100 Massive MIMO channel: A stochastic-geometrical channel model with random clusters with multipath components (MPCs) and their visibility regions [14]

are IID complex Gaussian (circularly symmetric) random variables, often written $h \sim CN(0, \sigma^2)$ where σ^2 denotes noise variance. The envelope $|h|$ is then Rayleigh distributed and the model is referred to as an IID Rayleigh-fading channel.

The realistic channels however have both direct LoS and non-LoS components and hence possess correlation and behave between the two extremes. The system behavior is typically assessed for these two channel conditions but the system capacity in massive MIMO is expected to be high in rich scattering environments (due to high rank channel matrix).

A more sophisticated channel model based on channel geometry can be used for more realistic system evaluations. One such channel model for MIMO is COST2100 [14], which is a stochastic-geometrical channel model. It models the geometry of channels by using clusters of scatterers with random positions as shown in Fig. 2.5 and stochastic channel properties to represent more realistic environment. The COST2100 has extensions for massive MIMO and supports for mobility and multiple frequencies [15]. It captures multi-link, multi-user, and multi-antenna correlation structures, which makes it particularly suited for MIMO and massive MIMO analysis.

2.5 Antenna Arrays

An antenna array is a system of multiple antenna elements arranged in a specific geometrical configuration as shown in Fig. 2.6 to achieve desired radiation characteristics that cannot be obtained with a single antenna ele-

ment. Arrays are fundamental to modern wireless communication systems, enabling advanced techniques like beamforming, spatial multiplexing, and interference mitigation.

Antenna arrays leverage the principle of wave interference by controlling the relative phases and amplitudes of signals fed to individual elements. These can be configured for various radiation patterns which can be directional or omnidirectional. The former gives a focused beams and reduced interference while latter gives 360-degree coverage for cellular base stations. Few configurations for the arrays are discussed below.

2.5.1 Uniform Linear Array

A Uniform Linear Array (ULA) consists of N identical antenna elements arranged along a straight line with equal spacing d between adjacent elements. This configuration is widely used due to its simplicity and analytical tractability.

An array factor (AF) is defined as,

$$AF(\theta) = \sum_{n=0}^{N-1} w_n e^{jnld\cos\theta} \tag{2.15}$$

where, w_n is a complex weight of nth element, $l = \frac{2\pi}{\lambda}$ is a wave number and θ is the angle from array axis. The inter-element spacing d is critical for array performance. Typically an optimal $\lambda/2$ spacing is used as it prevents grating lobes (undesired secondary maxima) and provides maximum spatial sampling without aliasing.

The radiation pattern of a ULA with $\lambda/2$ spacing is shown in Fig. 2.6. This spacing also balances mutual coupling and array size. A spacing $d < \lambda/2$ increases mutual coupling between elements, which can lead to impedance matching challenges and degraded array performance. However, it allows more elements to be placed within a given aperture, improving spatial resolution. Conversely, a spacing $d > \lambda/2$ can produce grating lobes and spatial aliasing in direction-finding applications. Nevertheless, since it reduces the number of elements required for a given aperture, it may be used strategically in certain specialized applications.

The ULA can electronically steer beams using phase shifters, a progressive

phase shift is given by,

$$\Delta\phi = ld\cos\theta_0 \tag{2.16}$$

where θ_0 is the desired steering angle.

2.5.2 Rectangular Planar Array

A rectangular planar array consists of antenna elements arranged in a rectangular grid with M elements along the x-axis and N elements along the y-axis as shown in Fig. 2.6. This two-dimensional configuration enables beam steering in both azimuth and elevation planes. Array factor in this case is,

$$AF(\theta,\phi) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{mn} e^{jmld_x \sin\theta\cos\phi} e^{jnld_y \sin\theta\sin\phi} \tag{2.17}$$

where, d_x and d_y represents horizontal and vertical element spacing and θ and ϕ are beam steering angles in elevation and azimuth. Rectangular array is particularly suitable for massive MIMO base stations due to their compact form and flexible beamforming (3D). It is also scalable for large number of elements. The radiation pattern in 2D with $\lambda/2$ spacing is shown in Fig. 2.6.

In practical massive MIMO implementations the horizontal spacing is typically 0.5λ to 0.7λ while vertical spacing is 0.5λ to 0.8λ with elements arranged in cross-polarized pairs [1]. Panel configurations can deploy multiple sub-arrays and are more useful in localization and tracking applications. Besides these linear and planar arrays, cylindrical, circular and concentric configurations are also prevalent in other applications [4].

2.6 Power Amplifiers

Power amplifiers are commonly used in transmitters in base stations and other wireless equipment. They amplify the radio frequency signal to an appropriate power level before transmission. These are typically the most power consuming parts of a base station due to the high power requirements of the transmit signals. A simple PA circuit as shown in figure 2.7 consists of transistors which can produce an output signal that has a level following the input waveform. A system is said to be efficient if a large part of the

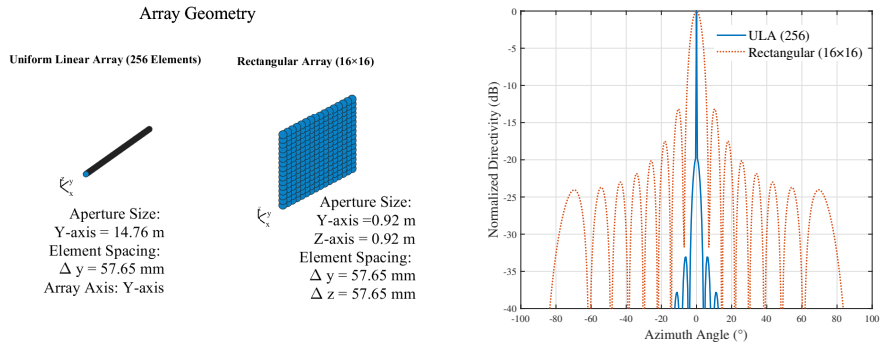


Figure 2.6: Array geometry and 2D Radiation Pattern comparison (Azimuth Cut, Elevation=0°) of a 256-element Uniform Linear Array and a Rectangular Array with $\frac{\lambda}{2}$ spacing.

energy supplied to the PA circuit is delivered to the antenna (by delivering power to the output load), rather than being dissipated as heat.

Efficiency of the PA is often measured as power added efficiency (PAE), defined as,

$$\texttt{PAE} = \frac{P_{\texttt{out}}}{P_{\texttt{in}} + P_{\texttt{dc}}} \tag{2.18}$$

where, $P_{\texttt{in}}$ is input signal power, which is usually very small compared to $P_{\texttt{dc}}$, which is the power supplied by the DC voltage supply, and $P_{\texttt{out}}$ is the PA output delivered to the load.

PAs typically have a signal level dependent power consumption, where more power is consumed when delivering large output power levels than at lower power levels. This is desirable as it enables power savings at low output levels. However, it should be noted that even when the signal level is zero, typically the PA has a non-zero power consumption. In modern BS, the sleep modes allow to turn off the PA power supply to reduce static power consumption [6]. Therefore, it is the power consumption during its dynamic operation which is of concern in this work, and the signal level dependent power consumption and associated efficiency must then be taken into account.

2.6.1 PA Nonlinearity

The PA behavior is usually characterized by a Pin-Pout curve as shown in Fig. 2.7. The PA output usually saturates at a maximum deliverable power for a given input level. When approaching the saturation level the

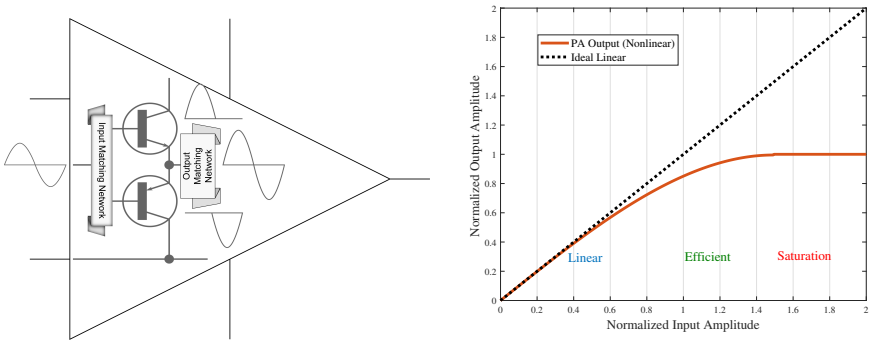


Figure 2.7: A class-B PA simplified circuit level architecture and a third order nonlinear PA input output curve showing PA operational regions (linear, high efficiency and saturation).

PA generates increasing amounts of nonlinear distortion, which can corrupt the transmit signals and disturb other transmissions in neighboring frequency channels. However, operating closer to the saturation region entails a more efficient PA operation and thus nonlinearity is unavoidable. Due to nonlinearity, additional frequency components populate the system, which degrades signal accuracy and effects system capacity.

Feeding a high PAPR signal into a PA requires the PA to be operated at a significant backoff to avoid excessive clipping distortion due to PA saturation at the signal peaks. This leads to inefficient PA operation, and even if special PA architectures are used for increased backoff efficiency, for instance Dohetry power amplifiers, the situation may still be problematic in terms of system cost. Peak reduction techniques are therefore usually deployed to reduce the peaks in the signal and to allow more efficient PA operation.

2.6.2 PA Classes

The efficiency also depends upon the circuit level architecture of the PA. With regard to this, PAs are categorized into different classes based on the operating conditions of their transistors.

The PA class determines what efficiency and linearity can be achieved. Some commonly used classes are described below:

1. In class A the transistors operate by conducting during the entire cycle of the input signal (360 degrees). It provides the highest linearity

and signal fidelity. However, it is also the least efficient class of operation, with a maximum theoretical efficiency of 50%. It should also be noted that the power consumption for small signals is as large as for the maximum signal, leading to extremely poor back-off efficiency, with an efficiency proportional to the output power.

2. In class B, the transistors operate by conducting for half of the input signal cycle (180 degrees). It thereby provides improved efficiency compared to Class A, with a maximum theoretical efficiency of 78.5% and thus generates less heat. It can ,however, introduce crossover distortion at the point where the transistors switch on and off. Practically, efficiency of class B is limited to 60% with efficiency following the $\sqrt{P_{out}}$. The efficiency in backoff is therefore superior compared to class A.

3. Class AB operates by conducting for slightly more than half of the input signal cycle (between 180 and 360 degrees). It provides a tradeoff between Class A and Class B by providing better efficiency than Class A and reduced crossover distortion compared to Class B. It is widely used in audio amplification and RF applications.

4. Class C operates by conducting for less than half of the input signal cycle (less than 180 degrees). It provides high efficiency, with a maximum theoretical efficiency often above 80% but introduces significant signal distortion, making it unsuitable for audio but useful for RF applications where signals can be filtered.

Each class of power amplifier has its own advantages and tradeoffs in terms of efficiency, linearity, complexity, and application suitability. The choice of amplifier class depends on the specific requirements of the application, such as signal fidelity, power efficiency, and operational frequency.

In 5G radio base stations, class AB based Dohetry PA architecture is prevalent due to its high efficiency and PAPR handling capability making them ideal for maintaining signal quality while optimizing power use [16]. The Dohetry amplifier also has an auxiliary amplifier that is only active during signal peaks. The auxiliary amplifier is typically biased in class C (with a conduction angle below 180 degrees), preventing it from conducting at low signal levels, which are then handled alone by the main amplifier biased in class AB.

Beside these linear PA classes, so called switched mode classes (D,E,F) are sometimes also employed for their high efficiency characteristics. These am-

plifiers, however, require more complex transmitter architectures to provide signals with amplitude modulation.

2.6.3 PA issues in Massive MIMO

In base stations, PAs face several challenges since nonlinear distortion is significant due to complex modulation schemes, affecting signal quality. Achieving high efficiency while handling diverse frequency ranges and bandwidth is also a major challenge. With massive MIMO, the number of required PAs scales with the number of antennas. Although the size of each PA can be reduced, the cost of using more sophisticated architectures such as Doherty becomes crucial. To simplify the hardware and reduce costs, simple class B PAs may then be deployed, but maintaining efficiency and signal linearity requires complex linearization techniques which can eventually lead to higher power consumption in digital part. Addressing these issues requires innovative PA circuit level designs and development of low complexity digital signal processing methods to ensure optimal system performance.

2.6.4 PA Modeling

The purpose of PA modeling is to characterize the PA behavior mathematically, for use in system analysis and to develop methods to combat the distortion. For this purpose several approaches have been adopted. These range from simple third order polynomial models to very complex Volterra series approximations. A popular model is generalized memory polynomial model [17], which can effectively model both the static nonlinearity and the memory effects of a PA. In this thesis, a few simple approaches have been used, such as a memory-less third order nonlinear model and a Modified Rapp model [7] for performance evaluations as they can capture the essential nonlinearity introduced by the amplifier. The memory effects have been ignored as these are more prevalent in high bandwidth scenarios. The simplicity of these models also allow for tractable expressions for RF parameter based modeling although more sophisticated and accurate modeling is required for nonlinearity combating techniques such as DPD. The baseband equivalent response of each of the m^{th} complex memoryless third order nonlinear amplifier [7], in the array is given by,

$$\boldsymbol{y}_m = \boldsymbol{x}_m(a_1 + \frac{3}{4}a_3|\boldsymbol{x}_m|^2) \tag{2.19}$$

where a_1 and a_3 are the coefficients representing the linear and third-order distortion terms, respectively. a_1 corresponds to the small-signal gain of the amplifier and is most significant when the amplifier operates in its linear region whereas a_3 term represents the third-order nonlinearity, which produces third-order inter-modulation distortion (IM3). This distortion has frequency components at combinations of the input frequencies f (such as $2f_1 - f_2$ and $2f_2 - f_1$). IM3 is particularly problematic because it falls within the bandwidth of the original signal, degrading the communication quality, as well as, in the adjacent frequency channels, disturbing neighboring communications. The compression behavior of a PA, particularly the 1dB compression point, is influenced by third order terms. As the input signal increases, the nonlinear term begins to dominate, reducing the gain (compression). The third-order coefficient a_3 typically plays a large role in determining the onset of gain compression.

RF Nonlinearity Metrics

These metrics are often used to describe non-linearity of RF building blocks as described below.

1. Third-Order Intercept Point (IP3): The third-order intercept point (IP3) is one of the most important RF parameters used to describe nonlinearity. It measures how strong the third-order inter-modulation distortion products are relative to the desired signal, typically in a two-tone test.

2. Output IP3 (OIP3): The extrapolated output power level at which the power of the third-order intermodulation products would equal the power of the fundamental signal if both increased asymptotically.

3. Input IP3 (IIP3): The input power level at which third-order intermodulation products would equal the fundamental tones at the output.

4. 1 dB Compression Point (P1dB): P1dB is the (input or) output power level at which the gain of the amplifier drops by 1dB due to nonlinearity. It is an important measure of how much output power the PA can deliver before significant gain compression occurs.

5. Output Saturation Power (OPSat): The output saturation power refers to the maximum output power that the PA can deliver. Beyond

this level the PA cannot further increase its output power even if the input power continues to increase. This occurs because the amplifier reaches its maximum capability and becomes fully saturated. After the saturation level the PA gain drops significantly, as additional input power does not result in any increase in output power.

Relationship between RF non-linearity metrics and PA model parameters

The metrics that are used to characterize the nonlinearity of the PA are analytically related to the coefficients of the PA's nonlinear model [7]. For a third order nonlinear model, given 1dB compression point (P1dB), the coefficient a_3 can be estimated. Assuming a normalized amplifier model where $a_1 = 1$, the relationship between a_3 and 1dB compression can be derived as,

$$a_3 = (1 - 10^{(-1/20)})\frac{4a_1}{3(A_{1\text{dB}})^2} \tag{2.20}$$

where $A_{1\text{dB}}$ is the input amplitude of the signal at which output power is 1dB lower than the input. The above expression indicates that a higher 1dB compression point, with less compression, corresponds to a smaller a_3 indicating less third-order nonlinearity. Similarly,

$$IA_{sat} = (2/3)\sqrt{\frac{a_1}{|a_3|}} \qquad\qquad OA_{sat} = (4a_1/9)\sqrt{\frac{a_1}{|a_3|}}; \tag{2.21}$$

where, IA_{sat} and OA_{sat} are input and output referred saturation voltages. In saturation, the amplifier is deep into its nonlinear region, and any further increase in input power contributes mostly to distortion rather than useful amplification.

Spectral Regrowth

Spectral regrowth refers to the broadening of the signal spectrum due to nonlinearity, which causes energy to spill into adjacent frequency channels. This is the result of odd order inter-modulation distortion. The nonlinearity can be the result of an amplifier operating close or into its saturation region. This can happen especially in communication systems using modulation schemes like OFDM, which have high PAPR. A direct consequence of nonlinearity on spectral regrowth is measured in terms of ACLR defined

in previous sections. The regulations for ACLR are usually specified by standardization authorities such as 3GPP [13].

*In this work, PA nonlinearity is modeled by characterizing its model coefficients using its RF metrics such as P*1dB*. Although low-fidelity PA models have been used, this provides a generalized framework to characterize a PA based on its RF metrics, enabling a direct analysis of the impact of nonlinear distortion on system performance and efficiency.*

2.7 Digital Signal Processing Techniques for Handling Nonlinear PAs in Massive MIMO Transmitters

Since PA is a major power-consuming component in a BS, the digital signal processing techniques required to maintain its linearity and efficiency, contribute substantially to the power consumption of the digital front end. Among these, DPD is very power-hungry, with its complexity scaling proportionally to the system bandwidth and the number of antenna elements [8]. Additional contributors include PAPR reduction, digital filtering, and up-conversion processes. The computational complexity of filtering operations increases with the desired bandwidth and spectral selectivity, whereas the complexity of PAPR reduction depends on the specific technique employed and the associated tradeoffs between signal quality and implementation cost. The following sections dive deeper into DPD and more particularly focus on PAPR reduction methods to emphasize the significance of understanding their tradeoffs under practical system constraints.

2.7.1 Digital Predistortion Techniques

Digital predistortion is a popular technique where digital signal processing is used to combat PA nonlinear behavior. PA designs are mainly adopted for high efficiency, while compromising linearity which then requires compensation in analog or digital [8]. In practice, DPD is more popular as it allows PA to operate closer to its saturation region. The signal saturation in PA is avoided as it is irrecoverable and incurs strong inband and out-of-band distortion. As shown in Fig. 2.8, an inverse of a PA nonlinearity is estimated so that the cascaded impact of DPD and PA system is linear.

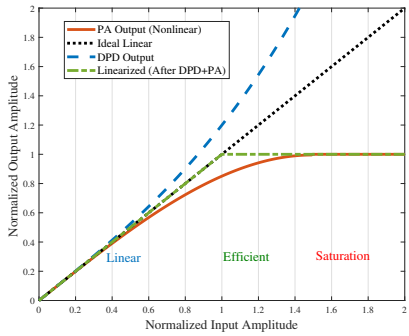


Figure 2.8: PA AM-AM response of normalized third Order nonlinear model, Digital Pre-distorter and Linearized output

This requires real time input-output characterization of amplifier through a feedback to the digital processing system.

Nonlinear system models are typically employed to represent the inverse response, offering tradeoffs between accuracy and implementation complexity [8]. However, even with highly accurate DPD models, systems operating with significant PA backoff due to high PAPR experience under-utilization of DPD resources. Consequently, PAPR reduction remains essential and is a standard part in OFDM based systems.

In massive MIMO systems, PA efficiency emerges as a critical bottleneck due to the large number of amplifiers required. Furthermore, linearization demands that DPD scale with the number of antennas, which has become a prominent research focus in recent years [18]. One promising approach for improving DPD scalability is the use of time-shared feedback paths from PA outputs for nonlinearity characterization in the digital domain, since PA output characterization is not required continuously [8].

Beyond scalability, the complexity of digital compensation for wide-band signals poses a major challenge, as the digital processing overhead can become a system bottleneck, as depicted in Fig. 1.2. Beam-domain DPD techniques have been proposed to alleviate this issue, although further advancements are still required [8]. Similar challenges of complexity and scalability extend to PAPR reduction methods as well. For this purpose, the PAPR methods are discussed in detail in the next section.

2.7.2 Peak to Average Power Ratio Reduction Techniques

As discussed above, PAPR reduction is critical for the efficiency of OFDM-based massive MIMO systems. Even when DPD is employed to mitigate PA nonlinearity, its effectiveness is significantly diminished if the PAPR is high and the system is forced to operate with large backoff. In such cases, PA efficiency is severely reduced, becoming a fundamental system bottleneck.

Research on PAPR reduction has spanned more than two decades, and comprehensive surveys are available in the literature. Existing approaches can be broadly categorized into three categorizes [19, 20].

1. *Distortion-based methods*: These techniques reduce peaks at the cost of introducing in-band distortion or out-of-band radiations. Common examples include clipping and filtering, peak windowing, and companding. While they offer low complexity and effective PAPR reduction, the introduced distortion can degrade EVM and increase adjacent channel interference.

2. *Distortionless methods*: These approaches achieve PAPR reduction without introducing distortion, but typically at the cost of high computational complexity or signaling overhead. Representative methods include selected mapping (SLM), partial transmit sequences (PTS), and tone reservation (TR). While they preserve signal fidelity, their practical implementation in large-scale systems is challenging due to complexity, side information requirements, or additional power consumption.

3. *Hybrid methods*: These techniques combine aspects of distortion-based and distortionless approaches to strike a balance between performance and complexity. Examples include, adaptive clipping with error correction, active constellation extension (ACE), and coding-based methods that jointly optimize PAPR reduction with error control. Such methods are often attractive in practice, as they offer a compromise between achievable PAPR reduction, spectral efficiency, and implementation cost.

Similar to DPD, the PAPR reduction techniques do not seem to fit well, when the system is scaled with massive number of antennas. The distortion free methods are not scalable and distortion based methods involve

a huge system performance loss. For example, the system complexity can significantly increase if the distortionless methods such as SLM and PTS are deployed. Usually distortion based methods are favored as these are also compatible with many legacy systems [21] and scalable for frequency selective channels. The distortion based approach such as clipping is particularly useful to provide scalable and low complexity solution for PAPR reduction. We therefore dive further deep into the technique in the following section.

Clipping

Clipping is a method of reducing signal amplitude by limiting it beyond a set threshold. The Clipping can be described mathematically as is given as,

$$\boldsymbol{v}_m = \begin{cases} \boldsymbol{q}_m, \text{ if } |\boldsymbol{q}_m| < T \\ Te^{j\angle \boldsymbol{q}_m}, \text{ otherwise} \end{cases} \tag{2.22}$$

where, $T{=}\sqrt{E[\|\boldsymbol{q}_m\|^2] \cdot 10^{0.1\alpha}})$ is a peak signal threshold determined based on the required PAPR (α) of the signal $\boldsymbol{v}_m$. The process of hard clipping produces inband and out of band distortion in the signal which can degrade EVM and ACLR.

Despite extensive research in the domain of PAPR reduction, practical deployment in massive antenna systems remains challenging, particularly due to scalability concerns, feedback overhead, performance degradation and the need for tight integration with DPD and precoding schemes. This motivates the exploration of approaches, which exploit the excess spatial degrees of freedom in massive MIMO to achieve PAPR reduction without introducing significant distortion or increasing system complexity.

Peak Reduction in Massive MIMO

In massive antenna systems, the issue of PAPR is not only limited to multicarrier modulation but is also inherent to certain precoding techniques. This arises because the transmitted signal power across different antennas may become highly imbalanced, leading to very large instantaneous peaks at some antenna ports. Such peaks necessitate the use of power amplifiers with a wide dynamic range, which significantly increases hardware cost and complexity. From a system design perspective, this makes PAPR re-

duction an important design goal in addition to the traditional objective of interference mitigation.

As a result, the problem is often treated as a multi-objective optimization task, and several optimization-based techniques have been investigated in the literature [10]. While these methods can provide effective tradeoffs between interference suppression and peak power control, their computational requirements grow rapidly, making them impractical for OFDM-based massive MIMO systems operating over frequency-selective channels.

In comparison, distortion-based approaches (such as clipping or companding) remain attractive due to their relative simplicity and ease of implementation in frequency-selective environments. However, these methods inevitably introduce signal distortion, which can degrade system performance. Therefore, an important line of research in this domain focuses on developing strategies that minimize the performance loss associated with distortion-based methods, while still preserving their implementation advantages [21].

2.7.3 Antenna Reservation for PAPR Reduction

Antenna reservation is a technique that exploits the large number of antennas in massive MIMO systems to reduce the PAPR [22]. It typically employs a low-complexity distortion-based method such as clipping to suppress signal peaks. The resulting clipping distortion is then compensated by utilizing the availability of multiple transmit antennas together with CSI.

In antenna reservation, as shown in Fig.2.9, a subset of transmit antennas is reserved for transmitting a compensation signal. This signal is designed such that its effect cancels the clipping distortion at the receiver through destructive interference in space. Assuming linear and unitary gain PAs, the received symbol at a particular subcarrier s in a massive MIMO system with antenna reservation can be expressed as,

$$\boldsymbol{r}_s = \boldsymbol{H}_s \boldsymbol{y}_s + \tilde{\boldsymbol{H}}_s \tilde{\boldsymbol{y}}_s, \tag{2.23}$$

where, $\boldsymbol{y}_s$ denotes the main signal transmitted from the primary (non-reserved) antennas and $\tilde{\boldsymbol{y}}_s$ represents the compensation signal transmitted from the reserved antennas in frequency domain. $\boldsymbol{H}_s$ and $\tilde{\boldsymbol{H}}_s$ represents their channels matrices respectively. The principle of this operation is illustrated in Fig.2.9. In practice, the effectiveness of compensation signal

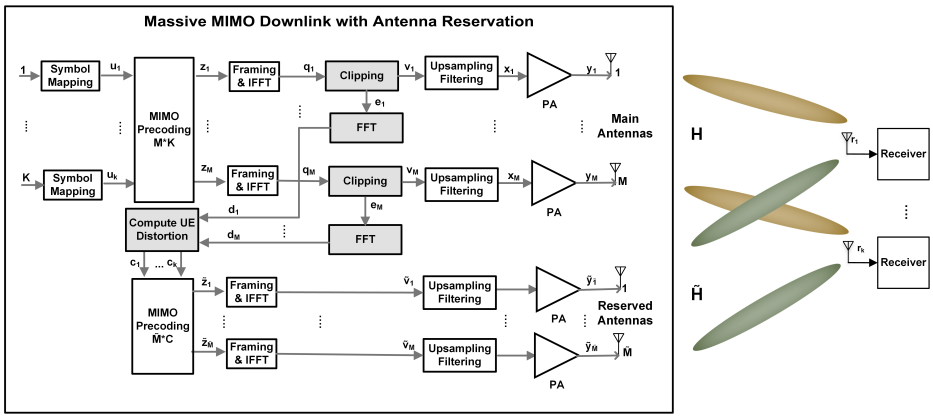


Figure 2.9: Antenna Reservation Technique in Massive MIMO

depends upon the number of reserved antennas and therefore this work focuses on the tradeoffs associated with using such methods in massive MIMO.

2.8 Summary

Energy efficiency and performance are two major aspect in design of communication systems. In this work, the focus is on efficiency enhancement for PAs in massive MIMO, which plays a major role in base station power consumption. In particular, the purpose is to develop and analyze low complexity digital signal processing techniques leveraging large number of antennas in the system to improve PAPR for frequency-selective channel scenarios. This entails that the techniques do not scale excessively in terms of computations, while simultaneously enabling the use of low-power analog components without sacrificing performance. Combined with other methods such as DPD, the low PAPR considerations define the requirements for PAs and can allow system design flexibility by deploying low-cost PAs pertaining to sustainable and energy-efficient massive MIMO architectures in future.

Chapter 3

Research and Outlook

Using massive number of antennas in a base station can allow to decrease the requirements of the analog components in the transceivers. This enables improved communication capabilities with low-cost hardware compared to traditional MIMO used in 4G systems. However, the digital signal processing requirements in massive MIMO demand scalability. Especially with pure digital beamforming architecture, increased processing requirements for each antenna lead to an increased power consumption in digital domain. Therefore, there is an immaculate need to adapt new physical-layer algorithms and techniques for future 6G wireless systems, in order to tackle the challenges of system scalability with the adoption to more flexible architectures [8].

The major power hungry techniques in front-end digital signal processing involves nonlinearity mitigation and efficiency enhancement for nonlinear PAs [8]. In this regard, an important area of consideration is understanding the system tradeoffs of using these techniques, especially as the output power requirements on PAs are relaxed. In particular, the system level impact of deploying each method in terms of its complexity and system performance. This has been an active research area in the past few years [18, 23, 24] and this thesis is a continuous effort in this direction.

The research focuses on following topics regarding massive MIMO system behavior with nonlinear PAs,

1. The study of distortion behavior in massive MIMO with diverse channel scenarios, precodings and user distributions.

2. Analysis of PA efficiency and linearity trade-off in massive MIMO for different user configurations and precodings for relaxed PA output power requirements.

3. Study and analysis of PAPR reduction methods leveraging large antennas in presence of PA nonlinearity.

4. Development of improved low-complexity methods for PAPR reduction and efficiency enhancement in massive MIMO.

The following sections describe the key questions in each domain and the contribution of the work published in this regard.

3.1 Distortion Behavior from Large Antenna Systems

The behavior of distortion in single antenna systems is well understood [25]. When analyzing nonlinear distortions stemming from PAs in massive MIMO, many questions arise regarding the impact of nonlinear distortions depending upon number of antennas, user scenarios, channel environments and applied precoding schemes. The distortion behavior when number of antennas and served users increase in the system has been extensively analyzed [25, 26]. Our work presents the behavior in more complex channel environment considering a sophisticated geometrical channel model and applied precoding techniques. The key research questions tackled in this regard are,

- What is the distortion behavior in more complex channel environments compared to pure LoS?
- What is the distortion behavior depending upon applied precoding schemes?

The key conclusions from this work are,

- In LoS channel environments, the distortion behavior verifies the previous results [25]. In particular, it is directed towards the user in a single user scenario as high spatial channel correlation [12] can make distortion correlated with the main signal. For more users it spreads

Table 3.1: Received Power Difference between ZF and MRT at the User Position

No. of Users	**Distance between Users** (m)	**Inband Power Diff. in** dB **(ZF−MRT)**	**OOB Power Diff. in** dB **(ZF−MRT)**	**OOB Beam Appearance Behavior**
1	-	0	0	Single beam towards user.
2	0.5	8	16	Beamformed towards users.
2	1	2	6	Beamformed towards users.
2	2	1.5	≈ 6	OOB beams split in other directions.
2	10	≈ 0.1	1	OOB beams split in other directions.
2	30	no difference	no difference	OOB beams split in other directions.

in more directions given by the spatial inter-modulated product angles.

- In more complex channel scenarios such as COST2100, the signal behavior is very complex. Apparently the Out-Of-Band (OOB) emissions appears to impact the same clusters as inband signal. However, the inband signal is maximized at user position while for the OOB there are other points in space for which it appears maximized.

- The OOB in ZF appears to vary from MRT at random points in space. The OOB in ZF near users clearly show a same behavior as inter-user interference cancellation of inband signal. This also shows that the distortion in this case has some correlation to the main signal.

At the transmitter, the average power per antenna variation in a LoS channel for ZF and MRT are observed to be different depending upon the user positions. For MRT, the mean power is usually stable whereas, ZF has high variations if the distance between the users is less than approx. 2m since the matrix inversion becomes numerically unstable. This appears in OOB behavior as well. The received power difference between the inband and OOB at the user position for ZF and MRT is summarized in Table 3.1. It can be seen that as the distance between the users increase, the numerical stability of ZF improves and both precoders appears to have similar OOB dispersion behavior. A further extension of this work with more complex and diverse scenarios and even considering mobility can reveal more useful insights into the distortion behavior in more realistic environments.

3.2 PA Tradeoffs in Massive MIMO

In Paper I, we also study the linearity-efficiency tradeoffs for PA and PA power consumption model. This reveals PA input signal characteristics in massive MIMO such as PAPR and mean signal power per antenna. The key investigations in this area include,

1. The impact of using different precodings on PA efficiency, power consumption and performance.
2. How different number of users impact the PA behavior in massive MIMO?

The key conclusion is that the PA efficiency appears to improve when same system serves more users. The study of signal characteristics further reveals that the signal characteristics in massive MIMO are more PA friendly when serving more users. More precisely, the signal distribution appears more like Gaussian. Moreover, the use of different precoders has negligible impact on PA efficiency. However the impact on the performance is more visible, when a PA with higher 1dB compression point (backoff) is used in ZF than MRT, since IUI dominates in MRT.

3.3 Performance Tradeoffs of Antenna Reservation with Nonlinear PAs

Antenna reservation as discussed in section 2.7.3 provides a low complexity solution which can exploit the large number of antennas in the system to compensate for clipping error compensation. In Paper II, we study the system performance tradeoffs for adopting such a scheme in massive MIMO with PA nonlinearity and saturation. The main investigation in system tradeoffs include,

- How many antennas need to be reserved?
- What is the impact of clipping level on performance?
- How the relaxed PA requirements for using large number of antennas impact the OFDM-based massive MIMO system tradeoffs with different clipping levels and number of reserved antennas?

A major drawback of reserving antennas is the reduction in array gain. One possible way of dealing with such a situation is to increase power in main/primary antennas to meet the SNR requirements at the user. Assuming similar PAs are being used all over the array, the key findings therefore are,

- The interplay of signal saturation in main antennas and reserved antennas determines the performance that can be achieved when using antenna reservation.
- The clipping level also determines the overall performance improvement that can be achieved in the system and the required number of reserved antennas.

In summary a higher clipping level means more energy in the compensation signal. The compensation signal therefore puts high requirements on PAs of reserved antennas to deliver very high PAPR residual signal especially if a more rigorous iterative clipping and filtering (ICF) [27] is employed before compensation.

Nevertheless, antenna reservation with non-iterative hard clipping provides a low complexity solution to a problem of clipping error compensation in massive MIMO and it does not require transmission of any side information. However in a practical system, the requirement to transmit extra power to meet the SNR requirements at the receiver and signal saturation in primary and reserved antenna PAs are a major concern while utilizing such a technique.

3.4 Low Complexity PAPR Reduction for Massive MIMO

A major drawback of antenna reservation is the reduction in effective array gain. For a ULA, the array gain is given as $10\log_2 M$, where M is the number of active antennas. Reserving antennas for compensation reduces this gain, which reduces SNR and degrades system capacity. This limitation however can be alleviated by *sharing the same spatial subspace* between the information-bearing and compensation signals, thereby maintaining array gain while still achieving distortion cancellation.

The key questions in this regard are,

- Which antennas should be selected for compensation signal transmission?
- What are the system performance and power consumption tradeoffs for using partial or full set of antennas for compensation in a shared antenna scenario?

In antenna reservation, performance strongly depends on the clipping threshold and the number of reserved antennas. A higher clipping threshold or a smaller number of reserved antennas places stricter requirements on the linearity and maximum output power capability of the PAs being used. In practice, the limited power-handling capability of PAs in antennas, selected for compensation, can distort the compensation signal beams and reduce their ability to cancel clipping distortion effectively.

To resolve this, compensation signal transmission can also be shared among all the antennas, ensuring that cancellation occurs at the receiver without requiring dedicated antennas. Moreover, regularized precoding can further enhance the performance by balancing distortion suppression, interference mitigation, and array gain preservation.

As discussed before the clipping can offload high output power transmission requirements to limited set of reserved antennas, one can argue that having a part of array which can deal with high power requirements could be beneficial for such techniques. An assumption could be that a more efficient Doherty PA can be used while rest of the system can be built using cheaper class B PAs.

To analyze such a system, consider a massive MIMO system with 128 element linear array with $\lambda/2$ spacing between the elements, serving three users in a LoS. The PA model in each antenna is characterized by its 1dB compression point ($P1\text{dB}$). We define quality factor Q as a ratio of $P1\text{dB}$ in reserved antennas to $P1\text{dB}$ in main antennas. This means that $Q = 1$ implies that the same grade PA are deployed all over the array. The precoders are regularized by a factor of 0.01 and clipping ratio is set to 4dB. We evaluate the performance and power consumption tradeoff for antenna selection as compared to a baseline system with only PA (without any PAPR reduction).

We fix the quality of PAs in main antennas assuming PAs with output 1dB compression point $P1\text{dB}_{\texttt{out}} = 0.05\text{W}$ are used. The selected antennas are uniformly distributed over the linear array. The EVM performance for varying quality of PAs for different number of selected antennas is shown

in Fig. 3.1. It can be noticed that as the quality of PAs in compensating antennas improve, the EVM decreases despite using nonlinear PAs of lower quality i.e. $P1\text{dB}_{\texttt{out}} = 0.05\text{W}$ in main antennas.

We compare the above analysis with a scenario without any peak reduction (PA only) where quality of PA is varied to operate all PAs at a backoff. It can be noticed the EVM performance when using all antennas for compensation outperforms performance for all other scenario particularly for $Q = 1$. However, for the lower number of compensating antennas the performance improvement is also limited despite improvements in the PA quality. This is because the compensation signal energy is less distributed over spatial resources which leads to more signal saturation in PAs of those antennas.

The potential point of interest in Fig. 3.1 can be where EVM curves exceed or meet a certain performance criteria. For example, a power consumption analysis (assuming class B PAs) of EVM= 4% in Fig. 3.1 reveal same PA power consumption if only PA is used with $Q = 3$ or if 32 compensation antennas are available with $Q = 10$. In the former case, performance criteria has been met by deploying PAs at a backoff which reduces PA efficiency and thus increase its power consumption. In the later scenario, the main PAs are operated efficiently due to lower $P1\text{dB}$ whereas the compensation signal entails very high PA output requirements reducing PA efficiency and increasing power consumption. This can bring one to the point of diminishing returns as the power consumption is same for the achieved efficiency, unless otherwise a different PA such as Doherty with high backoff efficiency is used in compensation signal transmission. In summary, efficiency enhancement through PAPR reduction can create very high output power requirements on PAs of compensating antennas in order to meet performance requirements.

This leads us to the study and proposal presented in Paper III where all antennas are utilized for clipping distortion compensation. The analysis demonstrates that such an approach enables spatial filtering of distortion such that the performance at the user remains unaffected. The work provides a comprehensive performance analysis of the proposed approach under practical system constraints such as nonlinear PAs, CSI error and PA mismatch, thereby validating the robustness and applicability of the proposed scheme in realistic transmission scenarios.

Since the addition of a compensation signal can lead to peak regrowth, the proposed method is implemented in an iterative manner. Accordingly,

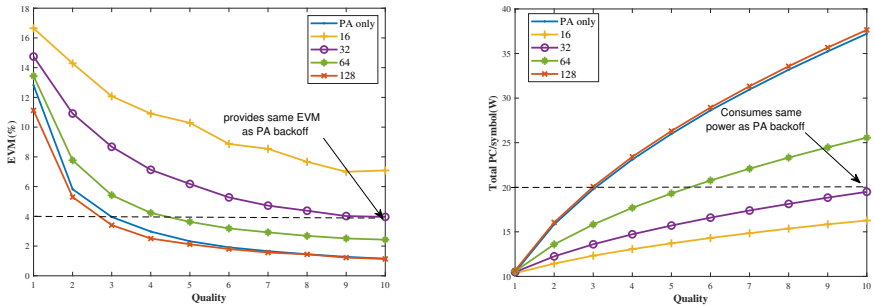


Figure 3.1: Performance and Power Consumption Analysis of a Massive system with Compensation Antenna Selection for Different Number and Quality Factor. The Main Ant. PA $P1\mathrm{dB}_{\mathrm{out}} = 0.05\mathrm{W}$.

its computational complexity has been compared with other iterative distortion compensation techniques available in the literature, while its performance has been extensively evaluated against distortion-based PAPR reduction techniques.

The key findings and conclusions of this work are summarized as follows:

- The proposed precoding-based approach offers a low-complexity and fast-converging solution for spatially filtering out distortion components directed towards the intended users.
- Performing user distortion mitigation prior to transmission provides a robust mechanism to minimize performance degradation at the receiver while simultaneously achieving PAPR reduction.
- The method demonstrates robustness against power amplifier mismatches and enhances performance under PA nonlinearity due to its improved PAPR characteristics. However, similar to the conventional zero-forcing precoding scheme, it remains sensitive to channel state information estimation errors.

The additional contributions from this section within the domain of antenna reservation/selection are concluded as,

- The array gain reduction in antenna reservation can be avoided by using antenna selection instead of reservation. This may require extra power delivered through those antennas.

- The above selection scheme can be best exploited if the compensating antennas are uniformly distributed over the linear array as compared to random or consecutive antenna selection. In a more practical channel conditions, a more optimized selection schemes may be deployed based on channel parameters and antenna shadowing.
- The method of Paper III provides an upper bound of performance using antenna selection where all the antennas are shared for both compensation and main signal transmission and provides best performance in the presence of nonlinear PA for the evaluated approaches.

3.5 Future Works

PA efficiency and linearity are two contradicting requirements. To meet these requirements in modern wireless communication systems, particularly those employing large-scale multiple-input multiple-output and orthogonal frequency-division multiplexing techniques, complex digital signal processing techniques pose significant scalability challenge. The low-complexity PAPR reduction methods in this regard, can help to improve PA efficiency, provided that the performance is not sacrificed. The precoding-based iterative clipping and distortion compensation has emerged as a promising approach in this regard, allowing the transmitter to counteract distortion before signal transmission, without sacrificing performance at user end.

With regard to this, a more optimum strategy can involve leveraging underutilized system resources, such as antennas that are not actively engaged in data transmission. Several criteria can guide the selection of such antennas for instance, the antennas that remain idle during a given coherence time interval, since the wireless channel is assumed to remain constant across multiple OFDM symbol durations within that interval. Moreover, channel parameters and antenna shadowing characteristics in realistic propagation environments can be further exploited to enhance the effectiveness of this approach, enabling adaptive and resource-efficient distortion mitigation.

In practical wireless communication systems, the transmit power exhibits significant variation due to the combined effects of modulation schemes, precoding strategies, user dynamics, and resource allocation policies. Consequently, a distortion-aware power control strategy becomes essential to exploit system resources efficiently while maintaining high linearity and energy efficiency. Such a strategy would ideally adapt transmission param-

eters based on instantaneous channel conditions and hardware limitations to optimize overall system performance [28]. In this thesis, a simplified power control approach has been adopted to maintain analytical and computational tractability. As a result, the intricate tradeoffs associated with more complex power adaptation, user scheduling, and resource allocation are considered beyond the scope of this work.

The proposed approach focuses on mitigating distortion through a linear projection technique, which effectively eliminates user-related distortion components from the transmit signal. As analyzed, this projection acts as a form of spatial filtering applied to the distortion across both time and frequency domains, analogous in spirit to the ICF method [27]. However, unlike conventional ICF techniques that primarily address OOB emissions, the proposed method provides in-band distortion mitigation, thereby improving signal fidelity at the receiver.

Implementing such spatial filtering directly in the time domain—without resorting to IFFT/Fast Fourier Transforms (FFT)-based transformations poses substantial computational challenge due to the coupled nature of spatial and temporal distortion components. To address this complexity, machine learning–based techniques may be investigated as a future direction. Data-driven models can potentially learn the nonlinear mapping between transmitted symbols, distortion characteristics, and optimal projection parameters, thereby enabling efficient and adaptive distortion compensation without explicit frequency-domain processing.

References

[1] Ericsson 5G Blog, 2025. https://www.ericsson.com/en/5g [Accessed: 2025-11-01].

[2] Sha Hu, Fredrik Rusek, and Ove Edfors. Beyond Massive MIMO: The Potential of Data Transmission With Large Intelligent Surfaces. *IEEE Transactions on Signal Processing*, 66(10):2746–2758, 2018. doi: 10.1109/TSP.2018.2816577.

[3] Jesús Rodríguez Sánchez. *Systems with Massive Number of Antennas: Distributed Approaches.* Doctoral thesis, 2022.

[4] Emil Björnson, Luca Sanguinetti, Henk Wymeersch, Jakob Hoydis, and Thomas L Marzetta. Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays. *Digital Signal Processing*, 94:3–20, 2019.

[5] Oliver Blume, Dietrich Zeller, and Ulrich Barth. Approaches to energy efficient wireless access networks. pages 1 – 5, 04 2010. doi: 10.1109/ISCCSP.2010.5463328.

[6] Improving energy performance in 5G networks and beyond, 2022. Available: https://www.ericsson.com/4a485f/assets/local/reports-papers/ericsson-technology-review/docs/2022/improving-energy-performance-in-5g-networks-and-beyond.pdf.

[7] Francois Horlin and Andre Bourdoux. *Digital Compensation for Analog Front-Ends: A New Approach to Wireless Transceiver Design.* 05 2008. ISBN 9780470517086. doi: 10.1002/9780470759028.

[8] Xin Liu, Guan-sheng Lv, De-han Wang, Wenhua Chen, and F.M. Ghannouchi. Energy-efficient power amplifiers and linearization tech-

niques for massive MIMO transmitters: a review. *Frontiers of Information Technology Electronic Engineering*, 21:72–96, 01 2020. doi: 10.1631/FITEE.1900467.

[9] Jeff Shepard. What is beamforming?, 2025. `https://www.5gtechnologyworld.com/what-is-beamforming/`.

[10] Mahmoud A. Albreem, Alaa H. Al Habbash, Ammar M. Abu-Hudrouss, and Salama S. Ikki. Overview of Precoding Techniques for Massive MIMO. *IEEE Access*, 9:60764–60801, 2021. doi: 10.1109/ACCESS.2021.3073325.

[11] Joao Vieira, Fredrik Rusek, and Fredrik Tufvesson. Reciprocity calibration methods for massive MIMO based on antenna coupling. In *2014 IEEE Global Communications Conference*, pages 3708–3712, 2014. doi: 10.1109/GLOCOM.2014.7037384.

[12] Emil Björnson, Jakob Hoydis, and Luca Sanguinetti. Massive MIMO networks: Spectral, energy, and hardware efficiency. *Foundations and Trends® in Signal Processing*, 11(3-4):154–655, 2017. ISSN 1932-8346. doi: 10.1561/2000000093.

[13] *Base Station (BS) radio transmission and reception*. 3GPP, 2021. Rel-16.

[14] Xiang Gao, Jose Flordelis, Ghassan Dahman, Fredrik Tufvesson, and Ove Edfors. Massive MIMO Channel Modeling - Extension of the COST 2100 Model. Joint NEWCOM/COST Workshop on Wireless Communications (JNCW) ; Conference, 2015.

[15] Jose Flordelis, Xuhong Li, Ove Edfors, and Fredrik Tufvesson. Massive MIMO Extensions to the COST 2100 Channel Model: Modeling and Validation. *IEEE Transactions on Wireless Communications*, 19(1): 380–394, 2020. doi: 10.1109/TWC.2019.2945531.

[16] Manlin Xiao and Wenyu Zhang. A Design and Implementation of High-Efficiency Asymmetric Doherty Radio Frequency Power Amplifier for 5G Base Station Applications. *Electronics*, 14:1586, 04 2025. doi: 10.3390/electronics14081586.

[17] Dennis R Morgan, Zhengxiang Ma, Jaehyeong Kim, Michael G Zierdt, and John Pastalan. A generalized memory polynomial model for digital predistortion of RF power amplifiers. *IEEE Transactions on signal processing*, 54(10):3852–3860, 2006.

[18] Muhammad Furqan Haider, Fei You, Songbai He, Timo Rahkonen, and Janne P. Aikio. Predistortion-Based Linearization for 5G and beyond Millimeter-Wave Transceiver Systems: A Comprehensive Survey. *IEEE Communications Surveys and Tutorials*, 24:2029–2072, 2022. ISSN 1553877X. doi: 10.1109/COMST.2022.3199884.

[19] Yasir Rahmatallah and Seshadri Mohan. Peak-to-average power ratio reduction in ofdm systems: A survey and taxonomy. *IEEE Communications Surveys and Tutorials*, 15:1567–1592, 2013. ISSN 1553877X. doi: 10.1109/SURV.2013.021313.00164.

[20] Francisco Sandoval, Gwenael Poitau, and Francois Gagnon. Hybrid Peak-to-Average Power Ratio Reduction Techniques: Review and Performance Comparison. *IEEE Access*, 5:27145–27161, 11 2017. ISSN 21693536. doi: 10.1109/ACCESS.2017.2775859.

[21] Shashi Kant, Mats Bengtsson, Gabor Fodor, Bo Goransson, and Carlo Fischione. EVM Mitigation With PAPR and ACLR Constraints in Large-Scale MIMO-OFDM Using TOP-ADMM. *IEEE Transactions on Wireless Communications*, 21:9460–9481, 11 2022. ISSN 1536-1276. doi: 10.1109/TWC.2022.3177136.

[22] Hemanth Prabhu, Ove Edfors, Joachim Rodrigues, Liang Liu, and Fredrik Rusek. A low-complex peak-to-average power reduction scheme for OFDM based massive MIMO systems. In *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 114–117. IEEE, 5 2014. ISBN 978-1-4799-2890-3. doi: 10.1109/ISCCSP.2014.6877829.

[23] Siqi Wang, Morgan Roger, and Caroline Lelandais-Perrault. Impacts of Crest Factor Reduction and Digital Predistortion on Linearity and Power Efficiency of Power Amplifiers. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 66:407, 2019. doi: 10.1109/TCSII.2018.2855084.

[24] Sandrine Boumard, Mika Lasanen, Olli Apilo, Atso Hekkala, Cédric Cassan, Jean-Philippe Verdeil, Jérome David, and Ludovic Pichon. Power Consumption Trade-Off between Power Amplifier OBO, DPD, and Clipping and Filtering. 2014.

[25] Christopher Mollen, Erik G. Larsson, Ulf Gustavsson, Thomas Eriksson, and Robert W. Heath. Out-of-Band Radiation from Large Antenna Arrays. *IEEE Communications Magazine*, 56(4):196–203, 2018. doi: 10.1109/MCOM.2018.1601063.

[26] Erik G. Larsson and Liesbet Van Der Perre. Out-of-Band Radiation From Antenna Arrays Clarified. *IEEE Wireless Communications Letters*, 7(4):610–613, 2018. doi: 10.1109/LWC.2018.2802519.

[27] J. Armstrong. Peak-to-average reduction for OFDM by repeated clipping and frequency domain filtering. *Electronics Letters*, 38:246 – 247, 03 2002. doi: 10.1049/el:20020175.

[28] Bin Liu and Sofie Pollin. Distortion-Aware Power Allocation for Multi-Stream Distributed Massive MIMO System With Nonlinear Power Amplifier. *IEEE Open Journal of the Communications Society*, 5: 1566–1578, 2024. doi: 10.1109/OJCOMS.2024.3373333.

Scientific publications

Author contributions

Paper I: Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption

I participated in developing the simulation framework for nonlinear modeling and performance analysis. I participated in writing the manuscript.

Paper II: System Design and Performance for Antenna Reservation in Massive MIMO

I participated in integrating the method in the simulation framework and performance analysis. I participated in writing the manuscript.

Paper III: Low Complexity Clipping Distortion Compensation for PAPR Reduction in Massive MIMO-OFDM for Frequency Selective Channels

I participated in developing the theory, simulation and analysis. I participated in writing the manuscript.

Paper I

S. Muneer, L. Liu, O. Edfors, H. Sjöland and L. V. der Petre
Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption
2020 54th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, 2020, pp. 974-978
doi:10.1109/IEEECONF51394.2020.9443341.

Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption

Sidra Muneer*, Liang Liu*, Ove Edfors*, Henrik Sjöland* and Liesbet Van der Perre†
*Department of Electrical and Information Technology, Lund University, Sweden
†Department of Electrical Engineering, KU Leuven, Belgium

***Abstract*—Massive MIMO enables a very high spectral efficiency by spatial multiplexing and opens opportunities to reduce transmit power per antenna. On the other side, it also introduces new challenges in tackling the power amplifier nonlinearity due to the increased number of antennas. The behavior of Out-of-Band radiation from PAs in Massive MIMO is non-trivial depending upon the applied precoding scheme, propagation environment, and how the users are distributed spatially. In this paper, we analyze the in-band and Out-of-Band power distribution in space for a Massive MIMO base station with nonlinear power amplifiers. We also study the tradeoff between amplifier power consumption and system performance (in terms of error vector magnitude at intended receivers) for PAs with different backoff operating levels. The spatial analysis of Out-of-Band under more realistic channel conditions is more reflective of the situation, to determine system emission requirements. The tradeoff between performance and power consumption will provide a basis for future investigations into the design of efficient Massive MIMO systems, taking nonlinearities into account.**

I. Introduction

In traditional base stations, the power amplifier (PA) in radio front ends is a major source of power consumption. The nonlinear behaviour of a PA causes Inband and Out-of-Band (OOB) distortion. To limit this, 3GPP defines error vector magnitude (EVM) and adjacent channel leakage ratio (ACLR) requirements [1]. Digital pre-distortion (DPD) is an efficient solution to reduce the effect of non-linearities and improve PA efficiency. But, with the advent of new communication technologies such as massive MIMO [2], there is a need to determine whether the same requirements and solutions are still appropriate. Recent research suggests that DPD deployment in a massive antenna regime imposes great challenges in terms of scalability, bandwidth and power consumption [3].

Massive MIMO (mMIMO) brings in spatial multiplexing using digital beamforming to achieve high data rates and spectral efficiency [4], [5]. The large array gain allows reduction of the total transmit power of the base station and the large number of antennas allow reduction of the power per each transmit chain. Thus, the power requirements per component are expected to be different in 5G systems deploying Massive MIMO. Research in [3], [4] suggest that this can allow the use of less precise components. However, to effectively use the benefits of technology, the design choices at individual transmit chains need to be analyzed and their impact needs to be understood. The research question in this regard is, what is the system level impact of using nonlinear components in mMIMO? Specifically, how severe is the spatial impact of OOB emissions from the use of non linear PAs in mMIMO? This impact has already been analyzed in [6]- [8] for line of sight and statistical channel models for different user positions. In this work, we analyze the impact of both in-band and OOB emissions using a more realistic channel model, COST2100 [12], considering different precoders and user distributions. This analysis is a helpful step in determining what is the most efficient way of coping with nonlinearity in mMIMO and whether we need different OOB requirements with respect to conventional systems. This work also analyzes the trade-offs between PA power consumption and system performance when using PAs offering varying nonlinearities and power requirements. According to authors knowledge, such type of analysis does not exist for PAs in mMIMO. We further study how these trade-offs change in realistic channels, using different precoders, and with different user distributions. This will help provide guidelines for optimal system design choices in massive MIMO.

A. Notation

All upper-case bold-faced letters represent matrices, e.g., $\boldsymbol{H}$, and lower-case bold-faced letters represent vectors, e.g., $\boldsymbol{y}$. The discrete Fourier transform is denoted $\mathcal{F}\{\cdot\}$ and its inverse $\mathcal{F}^{-1}\{\cdot\}$. Expected value is denoted $E\{\cdot\}$ and the Frobenius norm of a matrix $\|\cdot\|_F$.

II. System Model

In Fig. 1, we present the base station (BS) transmitter system of a mMIMO system. The base-band discrete-time (per-OFDM symbol), per-subcarrier, propagation model in which this transmitter operates can be described as,

$$\boldsymbol{r}_s = \boldsymbol{H}_s \boldsymbol{y}_s + \boldsymbol{n}_s, \tag{1}$$

where, $\boldsymbol{r}_s$ is a $K \times 1$ vector of received signals at the K users, $\boldsymbol{H}_s$ the $K \times M$ propagation channel matrix, $\boldsymbol{y}_s$ a $M \times 1$ vector of transmitted signals from the M BS antennas, and $\boldsymbol{n}_s$ a $K \times 1$ vector of independent and identically distributed (i.i.d) circularly-symmetric zero-mean complex Gaussian noise variables. Subscript s indicates the subcarrier index, belonging to the set $\mathcal{S}$ of used subcarriers. The per-OFDM symbol time index has been suppressed for reasons of brevity in the mathematical notation. Since we are addressing OOB signals caused by nonlinearlities, we will model and observe transmission on more subcarriers than used for data transmission.

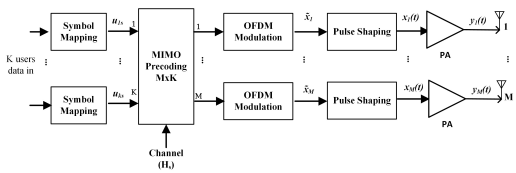


Fig. 1: Functional block diagram of Massive-MIMO downlink base station system model

With the propagation model described, let us return to the mMIMO transmission process at the BS. From the user-data input to the resampling and filtering, we again use a base-band discrete-time (per-OFDM symbol), per-subcarrier, model

$$\begin{bmatrix} \tilde{\boldsymbol{x}}_0 \\ \tilde{\boldsymbol{x}}_1 \\ \vdots \\ \tilde{\boldsymbol{x}}_M \end{bmatrix} = \mathcal{F}^{-1}\left\{\boldsymbol{W}_s \boldsymbol{u}_s\right\}_{s\in\mathcal{S}}, \tag{2}$$

where $\tilde{\boldsymbol{x}}_m$ is a $1 \times S$ row vector denoting the time-domain signal (before adding a cyclic prefix) intended for the m-th antenna, $\mathcal{F}^{-1}$ performs the OFDM-modulation row-wise, per antenna, $\boldsymbol{W}_s$ is the $M \times K$ mMIMO precoder matrix and $\boldsymbol{u}_s$ is the $K \times 1$ vector of user data-symbols. In this study, the precoder is either the maximal-ratio transmission (MRT) one,

$$\boldsymbol{W}_s = \alpha_{\text{MRT}} \boldsymbol{H}_s^H, \tag{3}$$

or the zero-forcing one,

$$\boldsymbol{W}_s = \alpha_{\text{ZF}} \boldsymbol{H}_s^H (\boldsymbol{H}_s \boldsymbol{H}_s^H)^{-1}, \tag{4}$$

where α_{MRT} and α_{ZF} are energy-normalizing constants.

Different power allocation schemes can be used at the base station depending upon different transmission policies. In general the deployed scheme should satisfy some system power constraint. In this case we assume that the total transmission energy of a BS for each OFDM symbol of length τ seconds in time, is limited to E. The total transmitted energy of the signal from all the antennas can be given as,

$$\sum_{m=1}^{M} \int_0^{\tau} |x_m(t)|^2 \, dt = E \tag{5}$$

provided that, a unitary basis is being used in symbol mapping and a unit gain amplifiers are assumed. $\boldsymbol{x}(t)$ is an up-sampled, filtered and pulse shaped version of $\tilde{\boldsymbol{x}}$. We consider a 4-fold up-sampling here and an ideal pulse shaping and filtering in order to avoid other sources of distortion than the ones being studied. If we deploy the above transmission policy to transmit equal power to all K users, each user gets an energy budget of E/K per OFDM symbol. The energy normalization constants for each user in (3) and (4) are defined as,

$$\alpha_k = \sqrt{\frac{E}{K}} \frac{1}{\|\boldsymbol{W_k}\|_F} \tag{6}$$

where, $\boldsymbol{W_k}$ is a precoder defined for each user over all subcarriers and antennas. The low pass complex equivalent response of each amplifier in a base station for a memory-less third-order nonlineairty [10], can be described as,

$$y_m(t) = x_m(t)\left(a_1 + \frac{3}{4} a_3 |x_m(t)|^2\right) \tag{7}$$

Although, more sophisticated amplifier models exists, this first order approximation is used here for a qualitative analysis of OOB. Thus the signal $\boldsymbol{y}_s$ in (1) represents a Fourier transform of each y_m i.e. $\mathcal{F}\{y_m\}$.

A. Simulation Setup

A Massive MIMO base station with 128 element linear array with $\lambda/2$ spacing and omni-directional antennas is considered. A QPSK symbol mapping have been considered and a precoder as defined in (3) and (4) have been considered at the transmitter to precode signal from multiple users to transmitting antennas. The framework is based on LTE based OFDM system with 20MHz channel bandwidth and a sub-carrier spacing of 15kHz, centered at a carrier frequency of 2.6GHz. It is assumed that a perfect channel knowledge can be acquired at base station.

B. Channel Models

We consider a static free space path loss based pure line of sight channel model as defined in [11] for each subcarrier s given by,

$$H_s = \frac{\lambda_s}{4\pi d} \exp(-\mathrm{j} 2\pi d/\lambda_s) \tag{8}$$

where d is the distance between any point in space and radiating element and λ_s is the wavelength for subacrrier s.

For analysis in more realistic channel environments, we consider COST2100 channel model, which is a stochastical-geometrical model with experimentally verified mMIMO extensions as presented in [12]. Table I provides a list of channel generation parameters for each subcarrier from COST2100.

Network	SemiUrban VLA 2.6GHz
Band	'Wide Band'
Antenna type	'MIMO VLA omni'
Scenario	'LOS'
Frequency band [GHz]	$2.57 - 2.63$
Center position of BS array[m]	$[0, -100, 2]$
Inter-element spacing[m]	$[0.0577, 0, 0]$
Number of positions at each BS	128

TABLE I: Summary of COST2100 parameters

C. PA Models and Evaluations

PA behaviour as defined in (7) is assumed to be used over all antennas chains in BS. For evaluations, three PA models with different 1dB compression points shown in Table II are considered. A same average transmit power of 50mW per antenna before PA has been chosen considering a link budget analysis to achieve an SNR of 10dB at the receiver for a single user case. To evaluate PA efficiency η and power consumption PC, we assume class B PA's with 60% achievable efficiency at saturation point, where efficiency η is directly proportional to

Amplifiers	Backoff operating level
Model 1	0dB
Model 2	5dB
Model 3	10dB

TABLE II: Summary of amplifier choices

the square root of input power as discussed in [13]. The drain power consumption P_{DC}, for such a PA can be determined using P_{OUT}/η. Hence total PA power consumption for a mMIMO BS with linear array can be estimated using,

$$PC = P_{in} + P_{DC} \tag{9}$$

Further, system performance is evaluated using mean error vector magnitude (EVM) for user at position (0,0).

III. Spatial Analysis of Inband and OOB

We have performed extensive simulations to observe the behaviour of transmitted signal both Inband and OOB, in pure line of sight (LoS) scenario as well as in a more realistic channel scenario with LoS component using COST2100 channel model. The analysis has been done for a planar area of 80m wide along x-axis and 160m along y-axis. Various user distributions with separation distance of 10m between the users have been considered. Here we only present the three user scenario as this can represent the general behaviour of other scenarios. Users lie in the same plane as BS, at a distance of 100m. We evaluate Inband and OOB power of received signal at each user position as well as the area around users and BS, by transmitting multiple OFDM symbols, assuming a same adjacent channel bandwidth of 20MHz for neighbouring channels.

A. Line of Sight channel

For a single user in LoS channel, Inband and OOB signal is always beamformed towards the user position (due to space constraints, we only show results with 3 users). When the number of users being served by base station are increased as shown in Fig. 2 Inband and OOB radiation appears to be beamformed towards user positions. However, when users are well separated ($d > 2$m in this case), OOB distortion appears to mainly radiate in the directions given by intermodulated spatial angles, $2\theta_m - \theta_n$ where, m and n are angles that determine users directions from BS. This results also corresponds to the result from [7]- [9].

B. COST2100 channel Model

OOB dissemination in COST2100 is different than LoS channel as shown in Fig. 2 for various precoders. A major difference comes from the geometry of the physical channel model, i.e. in COST2100 extensions to mMIMO [12], channel geometry is based on clusters and their associated circular visibility regions(VRs). The propagation of signal in such an environment mainly depends upon active clusters and corresponding VRs. These can be noticed as random circular regions in heat maps in Fig.2 (B). A slight LoS component is noticeable only for Inband. The power distribution largely appears to be similar for Inband and OOB propagation, with different overall power levels. This similarity indicates that both Inband and OOB signals propagate via the same clusters. However, the Inband signal power is focused at the individual user positions, both for LOS and COST2100, while the OOB signal is more geographically distributed and less focused in user positions for the COST2100 case than in LOS. If these observation holds in real environments, this may affect how adjacent channel interference regulations for mMIMO should be implemented.

C. Different Precoders

The signal propagation for Inband signal and OOB distortion in case of different precoders is not obvious from the individual maps. To highlight the difference, we show this difference in Fig.2 (C) for both channel models. All the differences have been set on same scale to make it comparable. We see that the difference in Inband and OOB signal dissemination in space for each precoder is different, provided that the same power is transmitted from the base station. The ZF appears to be 3dB higher or lower than MRT at random positions in space for COST2100. Moreover, near the user locations the OOB propagation in ZF appears to be better than MRT. Note that these results have been obtained only for an equal power transmission to all users.

IV. Power consumption and system performance in COST2100 using different PAs

Next, we present an analysis of system performance and power consumption when different users are being served by the BS. For a single user scenario the tradeoff for PA nonlinearity is similar to the 3-user case.

A. Three users in COST2100

In a realistic channel environment offered by COST2100, a massive MIMO BS with class B PAs operating according to Model 1 in Table II shows on average efficiency of less than 20% as shown in Fig.3 for three users. This is because power in most antennas is significantly low and peak to average ratio (PAPR) of OFDM is high. Using Model 2 instead, can give a significant advantage in terms of reducing OOB and improving EVM. This however results in poor efficiency deploying high power amplifiers. But if the power per antenna is extremely low, the cost of deploying such amplifiers should be compared to the cost of deploying complex nonlinear distortion cancellation techniques as the number of antennas grow very large. For more number of users, different precoders behave differently in terms of system performance. In MRT, received in-band distortion is dominated by user interference as indicated by EVM in Fig.3. Therefore, switching to PA Model 2 or Model 3 does not benefit much in terms of EVM as compared to ZF. In terms of power consumption and efficiency at transmitter end, though both precoders show similar performance.

B. Multiple user configurations

In general, if the total transmit power of base station is assumed to be fixed regardless of the number of users being served, a mMIMO system serving more number of users shows

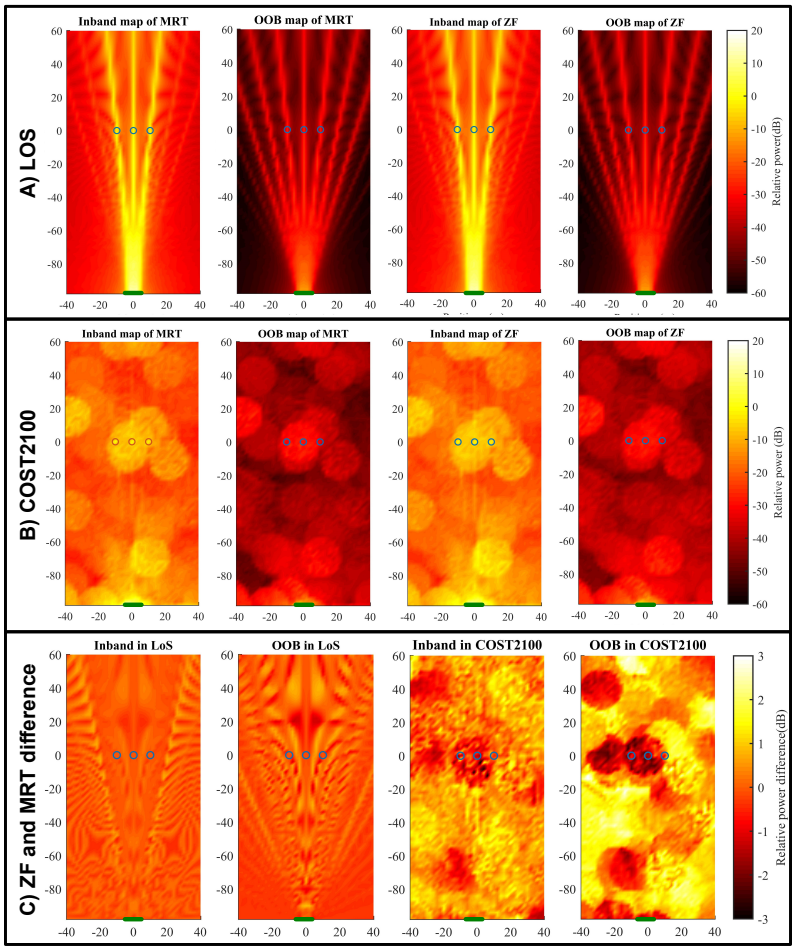


Fig. 2: Heat maps illustrating the available in-band and Out-of-Band power in a rectangular area, with x and y axis in meters, when using MRT and ZF precoders. In all plots, the BS is located at the green thick line and the three users at the three circles. LOS conditions are shown in A) and COST2100 in B) - both relative to the power received by user at (0,0). C) shows the power difference between the two pre-codings, for both channel conditions.

improved PA efficiency as shown in Fig. 4 for various user distributions with different spacing. This can be understood by analyzing the Pdf of power per antenna shown in Fig.5 for two scenarios with different number of users. Notice that the power distribution becomes more uniform resulting in an improved PA efficiency. Power consumption in this case increases instead of decreasing as input power to amplifiers increase.

V. Conclusion

In this work, an Inband and OOB spatial analysis in different channel environments for a mMIMO BS has been presented. We have presented results both for LOS and COST2100 channel conditions. LOS conditions allow us to compare our results to other studies, which typically use LOS channels, and the COST2100 allows us to find out what will happen in more realistic conditions. Our conclusion is that the distribution of power in LOS and COST2100 conditions are significantly different and, in particular, the distribution of OOB power is less focused and more widely distributed across geography in COST2100 channels. If this observation holds for real channels, it should influence adjacent-channel regulations for mMIMO systems. Further, the same methodology has been used to evaluate PAs with different power levels to measure the

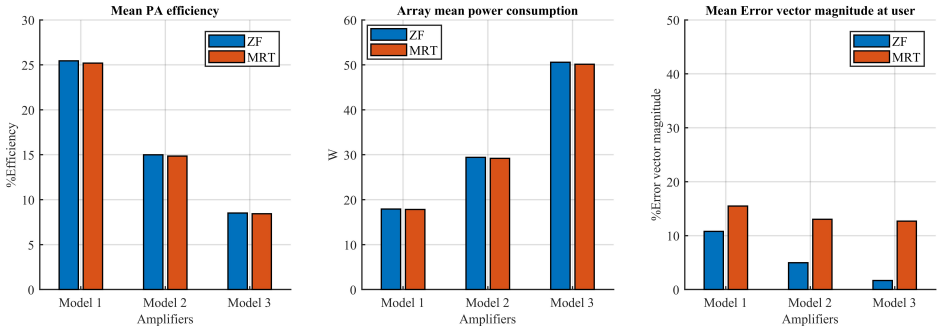


Fig. 3: PA performance for 3 users (20m apart) in COST2100.

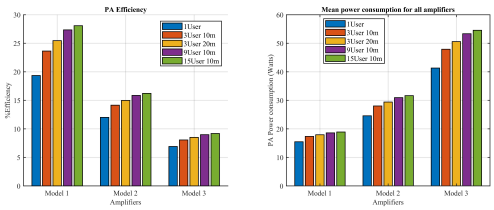


Fig. 4: PA efficiency and power consumption for various user distributions with different spacing using Zero-Forcing.

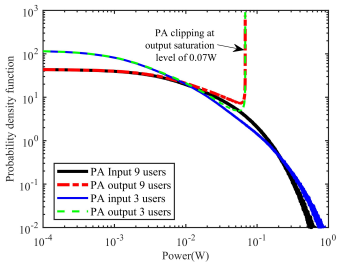


Fig. 5: Pdf of PA input and output for 3 and 9 user configurations

impact of nonlinearity on system performance along with their corresponding efficiency and PA power consumption. This analysis poses an important question of whether, for massive antenna systems, the best strategy for energy efficiency is to use a simple PA backoff or more complex compensation techniques? This question is, however, beyond the scope of this paper and left for future investigations.

ACKNOWLEDGMENT

The authors would like to thank Ericsson AB, Sweden for providing support under project 'Massive MIMO Technology and Applications' and other colleagues for sharing their work and knowledge.

REFERENCES

[1] 3rd Generation Partnership Project (3GPP), https://www.3gpp.org/release-16.

[2] 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, doi: 10.1109/TWC.2010.092810.091092.

[3] .C. Fager, T. Eriksson, F. Barradas, K. Hausmair, T. Cunha and J. C. Pedro, "Linearity and Efficiency in 5G Transmitters.." in IEEE Microwave Magazine, vol. 20, no. 5, pp. 35-49, May 2019, doi: 10.1109/MMM.2019.2898020.

[4] 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, doi: 10.1109/MCOM.2014.6736761.

[5] 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, doi: 10.1109/ACCESS.2017.2705561.

[6] E. G. Larsson and L. Van Der Perre, "Out-of-Band Radiation From Antenna Arrays Clarified," in IEEE Wireless Communications Letters, vol. 7, no. 4, pp. 610-613, Aug. 2018, doi: 10.1109/LWC.2018.2802519.

[7] C. Hemmi, "Pattern characteristics of harmonic and intermodulation products in broadband active transmit arrays," in IEEE Transactions on Antennas and Propagation, vol. 50, no. 6, pp. 858-865, June 2002, doi: 10.1109/TAP.2002.1017668.

[8] C. Mollen, E. G. Larsson, U. Gustavsson, T. Eriksson and R. W. Heath, "Out-of-Band Radiation from Large Antenna Arrays," in IEEE Communications Magazine, vol. 56, no. 4, pp. 196-203, April 2018, doi: 10.1109/MCOM.2018.1601063.

[9] L. Anttila, A. Brihuega and M. Valkama, "On Antenna Array Out-of-Band Emissions," in IEEE Wireless Communications Letters, vol. 8, no. 6, pp. 1653-1656, Dec. 2019, doi: 10.1109/LWC.2019.2934442.

[10] Horlin, Francois & Bourdoux, Andre. (2008). Digital Compensation for Analog Front-Ends: A New Approach to Wireless Transceiver Design.doi: 10.1002/9780470759028.

[11] Tse, D., & Viswanath, P. (2005). Fundamentals of Wireless Communication. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511807213

[12] J. Flordelis, X. Li, O. Edfors and F. Tufvesson, "Massive MIMO Extensions to the COST 2100 Channel Model: Modeling and Validation," in IEEE Transactions on Wireless Communications, vol. 19, no. 1, pp. 380-394, Jan. 2020, doi: 10.1109/TWC.2019.2945531.

[13] P. M. Lavrador, T. R. Cunha, P. M. Cabral and J. Pedro, "The Linearity-Efficiency Compromise," in IEEE Microwave Magazine, vol. 11, no. 5, pp. 44-58, Aug. 2010, doi: 10.1109/MMM.2010.937100.

Paper II

S. Muneer, J. R. Sanchez, L. Van der Perre, O. Edfors, H. Sjöland and L. Liu
System Design and Performance for Antenna Reservation in Massive MIMO
2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall), London, United Kingdom, 2022, pp. 1-5
doi: 10.1109/VTC2022-Fall57202.2022.10012708.

System Design and Performance for Antenna Reservation in Massive MIMO

Sidra Muneer*, Jesus Rodriguez Sanchez*, Liesbet Van der Perre†, Ove Edfors*, Henrik Sjöland* and Liang Liu*
*Department of Electrical and Information Technology, Lund University, Sweden
†Department of Electrical Engineering, KU Leuven, Belgium

***Abstract*—Peak to average power (PAPR) reduction of OFDM signals is critical in order to improve power amplifier (PA) efficiency in base stations. For massive MIMO, the complexity of these methods can become a real bottleneck in implementing low power digital signal processing chains. In this work, we consider an antenna reservation technique, which uses a low complexity clipping method to reduce signal peaks and leverages the benefit of massive antennas, by reserving a subset of antennas in order to compensate for the clipping distortion. Reserving antennas on the other hand reduces the potential array gain in the massive MIMO system, complicating the application of antenna reservation. This work explores various design space parameters in antenna reservation such as number of reserved antennas, amount of peak reduction and clipping methods. We investigate the impact of these parameters on the error vector magnitude at the user and on the adjacent channel power ratio at both transmitter and user positions. Our results enable a deeper understanding of antenna reservation as a low complexity PAPR reduction method in massive MIMO systems.**

I. Introduction

Massive MIMO (mMIMO) systems with hundreds of antennas at the base station require energy efficient solutions in order to incur low cost and sustainable infrastructures [1]. The Power amplifier (PA) is a major source of power consumption and thus its efficiency is a major concern due to the efficiency-linearity trade-off and high PAPR of OFDM signals. One way to improve PA efficiency is to operate it close to saturation, but this results in non-linear distortion emission from base stations transmitters, degrading the signal quality at the intended receivers as well as disturbing transmissions in the neighbouring frequency bands.

Traditionally, digital predistortion (DPD) has been used to handle PA nonlinearity, but the complexity and power consumption of DPD can be overwhelming if used for each antenna chain in mMIMO. To tackle this issue, PAPR reduction techniques leveraging large number of antennas in mMIMO have been proposed to relax DPD requirements. For instance in [2]–[5] low-PAPR precoding methods, including constant envelope precoding, have been proposed. These methods still suffer from high complexity because of iterative optimization solutions. [6] developed an attractive approach of antenna reservation, which reduces PAPR with much lower complexity.

In antenna reservation, transmitted signals are clipped using hard clipping (HC) before PA to reduce the PAPR of the signal. The inband distortion introduced due to clipping can be compensated by reserving set of antennas and transmitting compensation signal through those reserved antennas. Reserving antennas however comes with a drawback of reduced array gain in the system. Therefore in this paper we would like to find out how many antennas should be reserved and how much PAPR reduction is possible in order to balance the trade-off between distortion compensation capability of antenna reservation and array gain loss.

The OOB arising from clipping can be suppressed by low-pass filtering, but the process can regrow the peaks, diminishing the effectiveness of the antenna reservation method. We therefore also investigate different clipping methods such as iterative clipping and filtering (ICF) [7] to evaluate system performance with antenna reservation. In particular, we focus on down-link transmission and error vector magnitude (EVM) evaluation at user end for PAs with different 1dB compression points. Moreover, adjacent channel power ratio (ACPR) at both the base-station transmitter as well as user position is analyzed to evaluate the OOB distortion mitigation performance. Our results show that antenna reservation can be effective in improving PA efficiency by reducing PAPR. With carefully selected design parameters, antenna reservation is able to provide better performance, in terms of both EVM and ACPR, for mMIMO systems. In the following system model bold face upper case letters represent matrices and bold face lower case letters represent vectors.

II. System Model

A massive MIMO system with antenna reservation is shown in Fig.1. The transmitter consists of a set of main transmitting antennas M and a set of reserved antennas $\tilde{M}$, each capable of transmitting a set of data subcarriers S to K users using OFDM modulation. Subscripts m and $\tilde{m}$ represents signals at each main antenna and reserved antenna respectively, where $m = 1...M$ and $\tilde{m} = 1...\tilde{M}$. Consider $\boldsymbol{H}_s$, a $K \times M$ channel matrix for a particular subcarrier s for a set of M main antennas and $\tilde{\boldsymbol{H}}_{\boldsymbol{s}}$ a $K \times \tilde{M}$ channel matrix for a particular subcarrier s for $\tilde{M}$ antennas.

Data to be transmitted to each user, is precoded using a linear interference cancellation precoding scheme, only for M main antennas. A discrete-time domain signal $\boldsymbol{q_m}$ at each antenna is obtained after computing the Inverse Fast Fourier transform (IFFT) over all subcarriers for that antenna. The PAPR of the signal $\boldsymbol{q_m}$ is reduced by applying two different clipping and filtering methods. An error signal $\boldsymbol{e_m}$ at each main antenna is a difference between $\boldsymbol{q_m}$ and $\boldsymbol{v_m}$. This error is estimated for each user using frequency domain transformation and applying channel knowledge. This estimated user error is then precoded and modulated for $\tilde{M}$ antennas and transmitted

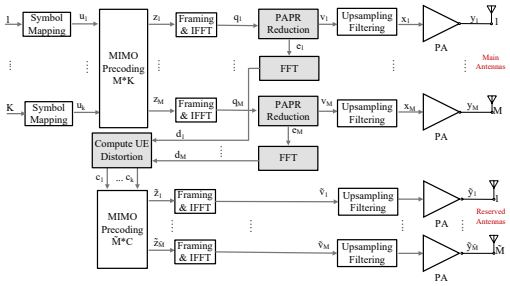


Fig. 1: Massive MIMO base station downlink system model with antenna reservation technique (gray shaded)

simultaneously with the corresponding peak reduced signal $\boldsymbol{v_m}$ in M main antennas, in order to reconstruct the signal at user position.

A. Precoding

Consider a precoding matrix $\boldsymbol{W_s}$ of size $M \times K$ which represents a linear regularized zero forcing precoder (RZF) for a subcarrier s in main antennas where,

$$\boldsymbol{W}_s = \boldsymbol{H}_s^H(\boldsymbol{H}_s\boldsymbol{H}_s^H + \delta \boldsymbol{I}_K)^{-1} \tag{1}$$

as defined in [8]. Here δ is a regularization factor chosen to maximize signal to interference and noise ratio (SINR) at the receiver and to avoid Gramian matrix ($\boldsymbol{H}_s\boldsymbol{H}_s^H$) inversion problems, and $\boldsymbol{I}_K$ represents identity matrix of size K.

If a vector $\boldsymbol{u}_s$ of length K represents users data symbols that must be allocated to a particular subcarrier s, consider $\boldsymbol{U} = [\boldsymbol{u}_1\boldsymbol{u}_2...\boldsymbol{u}_S]$ is a matrix containing such vectors with dimensions $K \times S$. Also consider $\boldsymbol{Z} = [\boldsymbol{z}_1\boldsymbol{z}_2...\boldsymbol{z}_S]$ of size $M \times S$ containing all subcarrier data for main antennas. Each $\boldsymbol{z}_s$ is $M \times 1$ vector containing all user information on subcarrier s defined as,

$$\boldsymbol{z}_s = \alpha \boldsymbol{W}_s\boldsymbol{u}_s \tag{2}$$

where α is a power scaling factor defined as,

$$\alpha = \frac{\sqrt{\beta P T_s/K}}{\|\boldsymbol{W}_k\|} \tag{3}$$

where, P is a total transmit power in Watts allocated fairly to all the users, T_s is the system sampling period in seconds and β is an array-gain loss adjustment factor due to antenna reservation. For reserving $\tilde{M}$ antenna at transmitter, we defined β as,

$$\beta = \left(1 + \frac{\tilde{M}}{M}\right)^2 \tag{4}$$

B. Framing and IFFT

Subcarrier data for each antenna from the precoder is further processed by an IFFT module to formulate together a composite time domain OFDM signal. If N represents the FFT size, We define,

$$\boldsymbol{Q} = \boldsymbol{F}^{-1}\boldsymbol{Z}^T \tag{5}$$

where $\boldsymbol{Q}$ is a $N \times M$ matrix and $\boldsymbol{F}^{-1}$ is a unitary Fourier inverse transform matrix of size $N \times N$. We assume that $\boldsymbol{Z}$ has been transformed to a size $M \times N$ matrix through insertion of guard band subcarriers. A high PAPR signal resulting from the IFFT is then processed by a PAPR reduction module to reduce the peaks in the signal.

C. PAPR Reduction

Error for all the antennas in matrix form is given as,

$$\boldsymbol{E} = \boldsymbol{Q} - \boldsymbol{V} \tag{6}$$

where $\boldsymbol{E}$ is $N \times M$. We have used a simple method of peak reduction as described in [6]. The Hard Clipping (HC) function is given as,

$$\boldsymbol{v} = \begin{cases} \boldsymbol{q}, \text{ if } |q| < T \\ Te^{j\angle \boldsymbol{q}}, \text{ otherwise} \end{cases} \tag{7}$$

where, T=$\sqrt{E[\|q\|^2] \cdot 10^{0.1\gamma} \cdot T_s)}$ is a clipping threshold determined based on the clipping level γ for the signal $\boldsymbol{v_m}$. Since filtering after clipping regrows the peaks to a certain extent, an iterative clip and filter (ICF) technique [7] has been considered for improved PAPR reduction. It uses a hard clipping and filtering approach in the frequency domain to eliminate OOB distortion and few iterations to achieve a PAPR closer to the clipping threshold (γ).

D. Distortion Prediction and Transmission

The distortion signal $\boldsymbol{c}$ for all the users in the subcarrier domain is expressed as,

$$\boldsymbol{c}_s = \boldsymbol{d}_s\boldsymbol{H}_s^T \tag{8}$$

where, $\boldsymbol{d}_s$ is the $1 \times M$ error vector in frequency domain and $\boldsymbol{c}$ has dimensions $1 \times K$. For simplicity we assume $\boldsymbol{H}_s = \boldsymbol{H}$, so we can drop index s such that:

$$\boldsymbol{C} = (\boldsymbol{F}\boldsymbol{E})\boldsymbol{H}^T \tag{9}$$

where, $\boldsymbol{C}$ is $S \times K$ matrix in frequency domain and $\boldsymbol{E}$ is $N \times M$ matrix in time domain. The precoded time domain signal at $\tilde{M}$ reserved antennas in matrix form is given as,

$$\tilde{\boldsymbol{V}} = \boldsymbol{F}^{-1}(\boldsymbol{C}\tilde{\boldsymbol{W}}^T), \tag{10}$$

E. Reception

Assuming an ideal amplifier, the final received signal as a result of transmission of precoded OFDM symbols, in frequency domain, after downsampling is given by,

$$\boldsymbol{R} = (\boldsymbol{F}\boldsymbol{V})\boldsymbol{H}^T + (\boldsymbol{F}\tilde{\boldsymbol{V}})\tilde{\boldsymbol{H}}^T \tag{11}$$

where $\boldsymbol{R}$ is an $S \times K$ matrix. Substituting the value of $\tilde{\boldsymbol{V}}$ and simplifying gives,

$$\boldsymbol{R} = (\boldsymbol{F}\boldsymbol{V})\boldsymbol{H}^T + \boldsymbol{C} = \boldsymbol{U} \tag{12}$$

Thus, signal information lost by clipping in order to reduce the PAPR can ideally be recovered completely through the use of antenna reservation. We next turn to system simulations in order to determine the impact on system performance when nonlinear PAs are used along with antenna reservation.

TABLE I: Simulation parameters

# antennas	128	Ant. spacing	$\lambda/2$
Bandwidth	20MHz	# subcarriers	1200
Tx power/ant.	0.05W	# users	3
Noise power	-174dBm/Hz	User sep.	20m
Oper. freq.	2.6GHz	Reg. factor, δ	$0.01\|(\boldsymbol{H}_s\boldsymbol{H}_s^H)\|$

TABLE II: Signal statistics at the transmitter

Clipping level (γ)	**2**dB	**4**dB
Original PAPR (dB) [$\boldsymbol{v_m}$]	9.6	9.6
PAPR before filtering (dB) [$\boldsymbol{v_m}$]	3	4.37
PAPR before PA (dB) [$\boldsymbol{x_m}$]	**5.27**	**6.23**
ACPR before PA (dB$_c$) [$\boldsymbol{x_m}$]	-44.65	-49.06

III. Performance evaluation and system design trade-offs

A. System setup and assumptions

An OFDM based mMIMO simulation framework has been developed considering the parameters given in Table I. We consider a free space line of sight channel (LoS) model with users situated approximately 100m away from the base station. In such a channel environment, the distortion radiating from nonlinear amplifiers in base stations is mainly beamformed in the direction of users [9], [10]. Thus, to measure communication performance EVM and ACPR have been evaluated at the user position using expressions,

$$\text{EVM}(\%) = E\left[\sqrt{\frac{(r-u)^2}{u^2}}\right] \times 100 \tag{13}$$

where r and u are received and reference constellation symbols respectively.

$$\text{ACPR}(\text{dB}) = 10\log_{10}\frac{\max(\sum \boldsymbol{P_l}, \sum \boldsymbol{P_r})}{\sum \boldsymbol{P_m}} \tag{14}$$

where $\boldsymbol{P_l}$ and $\boldsymbol{P_r}$ represent the power in left and right adjacent channel frequencies and $\boldsymbol{P_m}$ represents the power in main channel frequencies of the allocated spectral bandwidth.

The precoder regularization parameter δ has been chosen through simulation in order to maximize signal to interference and noise ratio at the user and to avoid a Gramian matrix instability problem. A 30 tap Chebychev filter with a 30 dB stop band attenuation is used to suppress OOB distortion due to clipping. The resulting ACPRs are shown in Table II, for clipping levels and a PA with 1dB compression point $P1\text{dB} = 0.05\text{W}$.

The PAPR at the PA input is given by that of the filter output $[\boldsymbol{x_m}]$. In Table II we also show the corresponding values of achievable PAPR at various signal points of Fig. 1.

The filtered signal is further fed to a third order memory-less nonlinear PA model.We consider two PAs based on different 1dB compression points (P1dB), which depicts two different scenarios of cost and energy consumption in mMIMO system. PA1 has a P1dB equal to the mean signal power, i.e. 0.05W, while PA2 has been chosen as a PA with higher saturation limit (higher cost) with a P1dB of 0.1W. For each evaluation, the same PA model is assumed across all transmitter chains.

B. Number of reserved antennas and clipping levels

In the antenna reservation technique, the number of reserved antennas and clipping level (γ) are two critical system design parameters. The reserved antennas $\tilde{M}$ contribute to transmit beams of a compensation signal $\tilde{\boldsymbol{y}}_{\boldsymbol{m}}$ towards the user. On the other hand, the clipping level determines the power of this compensation signal. We therefore analyze the trade-off between different clipping levels (γ), reserved antennas ($\tilde{M}$), and achievable performance for a three users scenario in a LoS channel. The results are shown in Fig. 2.

It can be noticed in Fig. 2 that for PA1 with clipping level $\gamma = 4\text{dB}$, the EVM initially improves as the number of reserved antennas grows. Further increasing the number of reserved antennas starts to degrade the EVM. This is because the EVM at the user is reflective of both PA distortion and clipping distortion. When the number of reserved antennas grows, the compensation capability of reserved antennas improves and reduces the impact of the error due to clipping. On the other hand, the main antenna mean power needs to be increased according to (4), which pushes the PA operation point closer to the saturation point. Thus, there is a trade-off between EVM performance and number of reserved antennas.

In terms of the ACPR at the user, the clipping process helps to reduce the PAPR of the transmit signal and thus reduces OOB distortion. However, increasing the number of reserved antenna does not help with ACPR reduction. This is because the antenna reservation method only compensates the in-band error due to clipping.

More aggressive clipping such as $\gamma = 2\text{dB}$, improves the PAPR afterwards leading to less distortion due to PA nonlinearity. On the other hand, it also increases the signal power in the error signal $\boldsymbol{e}_m$ which, in turn, calls for more reserved antennas than when $\gamma = 4\text{dB}$ is used, in order to avoid signal saturation in reserved antennas.

For PA2, with a higher saturation point, the number of reserved antennas required to reduce the EVM is higher than when using PA1. On the other hand, PA2 can deliver an approximately 1dB higher received power than PA1 at the same mean transmit power.

It is shown in the simulation results that, with a right balance of the number of reserved antennas and clipping level antenna reservation can provide better performance in terms of both EVM and ACPR. For instance, in the case of PA1 with 2dB clipping and 20 reserved antennas. At such a point, the distortion compensation capability of reserved antennas outperforms the distortion from amplifiers, providing improved performance at the user end as well as possibility of improving efficiency in main antennas.

More clipping results in higher signal power at reserved antennas, which can affect the standard compliance regarding the ACPR at the reserved antennas. This phenomenon can result in higher distortion received at the user position. To better illustrate this, we plot ACPR at the transmitter for both main and reserved antennas in Fig. 3 (PA1 is used in this analysis). The ACPR at main antennas increases with the

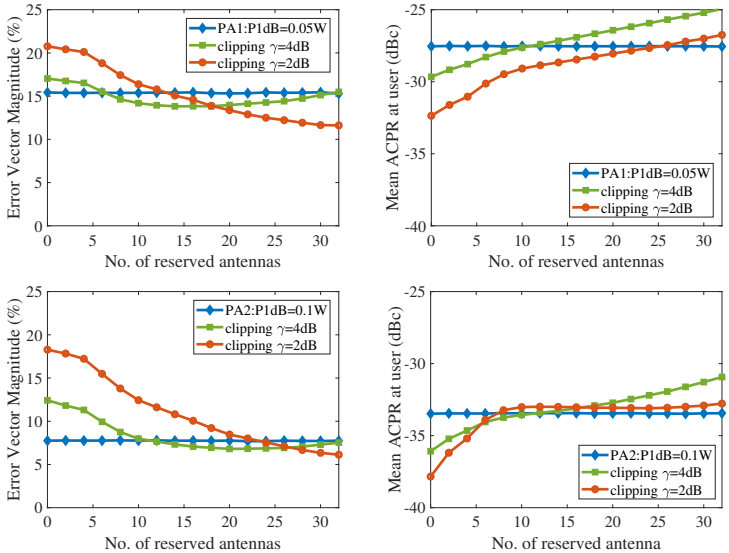


Fig. 2: Impact on received signal EVM and ACPR for reserving antennas in a mMIMO system with a 128 element linear antenna array serving 3 users. Each curve corresponds to a target clipping threshold of the signal in the main antennas, with a PA1 of $\mathrm{P1dB} = 0.05\mathrm{W}$ (top) and PA2 of $\mathrm{P1dB} = 0.1\mathrm{W}$ (bottom).

number of reserved antennas, while the ACPR at reserved antennas decreases. When the number of reserved antennas grows, the compensation signal power gets more distributed over a large number of reserved antennas, resulting in a lower power per antenna in the reserved part of the array. This reduces the OOB radiation from reserved antennas.

The problem of OOB from reserved antennas can be resolved by either using PAs with a higher saturation point in the reserved part of the array. This might result in lower efficiency in that part of the array, yet overall efficiency and power consumption when using antenna reservation must be considered. Here we conclude that the number of reserved antennas and PAPR clipping level along with the mean power operation point of the PA are important factors in determining the performance of the system if antenna reservation is deployed in mMIMO.

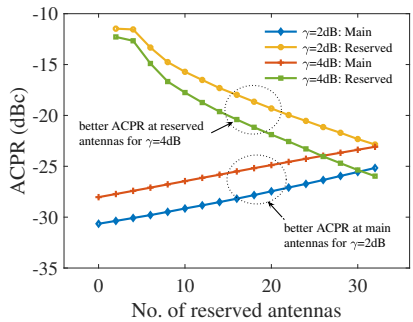


Fig. 3: A comparison of the ACPR at main and reserved antennas and for different clipping levels for a PA with $\mathrm{P1dB} = 0.05\mathrm{W}$.

C. Iterative clipping and filtering

Since different clipping methods result in different levels of peak regeneration after the upsampling-filtering process, we have evaluated the above mentioned system for ICF in order to improve PAPR reduction. For PA1, with $\mathrm{P1dB} = 0.05\mathrm{W}$, results are shown in Table III for 20 reserved antennas and $\gamma = 2\mathrm{dB}$. When a discrete time OFDM signal with a mean PAPR of $9.4\mathrm{dB}$ is passed through only the PA without reducing the PAPR, it introduces the EVM distortion of 15.36% while the mean ACPR at the main antennas is $-26.02\mathrm{dBc}$. Introducing only hard clipping before the PA to reduce peak of the signal results EVM of 20.92%, however it improves the ACPR at user by $4.83\mathrm{dB}$. With the antenna reservation and HC, the system can attain a mean EVM of 13.49% and can only maintain an ACPR improvement of $0.53\mathrm{dB}$ at the user position.

The PAPR can be further reduced by using ICF. In ICF, the clipping and filtering process is repeated until the PAPR becomes close to $2\mathrm{dB}$. The convergence process in ICF is logarithmic therefore we set the target PAPR level to 1.1γ which can be achieved in 12 iterations. This provides an effective PAPR reduction of $1.46\mathrm{dB}$ over HC, and in turn results in an ACPR improvement of $3.6\mathrm{dB}$ at the main antennas. ICF inherently results in loss in mean signal power

TABLE III: System performance of different configurations with and without antenna reservation

Setup	Clipping Method	PAPR before PA (dB)	ACPR main ant. (dB_c)	ACPR reserve ant. (dB_c)	ACPR at user (dB_c)	EVM at user(%)
Without	No Clipping	9.4	−26.02	-	−27.52	15.36
Antenna Reservation	Hard Clipping (HC)	5.31	−30.68	-	−32.35	20.92
With 20	Hard Clipping (HC)	5.31	−27.46	−19.38	−28.05	13.49
Reserved Antennas	Iterative Clip and Filter (ICF)	3.85	**-31.06**	−20.81	−31.13	14.08

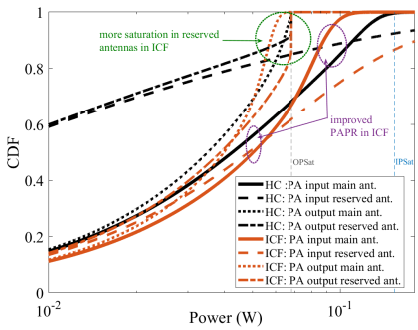


Fig. 4: A comparison of PA input and output power from main and reserved antennas when using iterative clip and filter (ICF) instead of hard clipping (HC) in antenna reservation. PA P1dB=0.05W and input and output saturation points are indicated as IPSat and OPSat

due to the iterative process which has been compensated by increasing the mean power per main antenna to that of the HC case, but the received signal at the user appears to be 1dB lower compared to HC. The exact reason for the received power loss in ICF needs further investigations, e.g., if ICF in any way compromises more the precoding.

In Fig. 4, the signal power distribution, when using ICF and HC, is shown. The power distribution in both main and reserved antennas at the PA input changes when ICF is used instead of HC. This is because ICF reduces signal PAPR at main antennas, while needs to transmit higher power at reserved antennas. This results in more PA saturation in reserved antennas as indicated in Fig. 4, which can degrade the compensation capability. Although the reserved antenna ACPR is better than that when using HC (−20.81dBc), as seen in Table III, the mean OOB power at reserved antennas is still higher. We conclude here that using ICF instead of HC to improve PAPR reduction can be beneficial in reducing ACPR but it does not guarantee an improvement of compensation performance over HC.

IV. Conclusion

In this paper we analyzed the performance of an antenna reservation technique in massive MIMO for various design parameters in a LoS channel. The main objective of any PAPR reduction technique is to improve PA efficiency, while improving, maintaining performance, or compromising it as little as possible. Our results show that antenna reservation can be a beneficial technique, especially in case of PAs operating closer to the saturation point. However, a vigilance is required with respect to the resulting reduction in array gain and corresponding increase in required mean output power. We used filtering in order to limit OOB due to clipping, which results in a certain regrowth of peaks. Therefore, an iterative clip and filter technique has been evaluated to investigate the possibility of performance improvement. The results show an improvement of 3.6dB in received ACPR from hard clipping, at the cost of off-loading more signal power to the reserved antennas. This also makes the number of reserved antennas impact overall system performance to a larger extent. This paper shows that antenna reservation has the potential to provide a low-complex means to control PAPR and enhance the total energy efficiency of massive MIMO systems. Further investigation is needed to mature and expand on these techniques.

Acknowledgment

The authors would like to thank Ericsson AB, Sweden for providing support under project 'Massive MIMO Technology and Applications'.

References

[1] 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.

[2] S. K. Mohammed and E. G. Larsson, "Per-Antenna Constant Envelope Precoding for Large Multi-User MIMO Systems," in IEEE Transactions on Communications, vol. 61, no. 3, pp. 1059-1071, March 2013.

[3] R. Zayani, H. Shaiek and D. Roviras, "PAPR-Aware Massive MIMO-OFDM Downlink," in IEEE Access, vol. 7, pp. 25474-25484, 2019.

[4] T. Kageyama, O. Muta and H. Gacanin, "Enhanced Peak Cancellation With Simplified In-Band Distortion Compensation for Massive MIMO-OFDM," in IEEE Access, vol. 8, pp. 73420-73431, 2020.

[5] S. Domouchtsidis, C. G. Tsinos, S. Chatzinotas and B. Ottersten, "Constant Envelope MIMO-OFDM Precoding for Low Complexity Large-Scale Antenna Array Systems," in IEEE Transactions on Wireless Communications, vol. 19, no. 12, pp. 7973-7985, Dec. 2020.

[6] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu and F. Rusek, "A low-complex peak-to-average power reduction scheme for OFDM based massive MIMO systems," 2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP), 2014, pp. 114-117.

[7] Armstrong, J.: 'Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering', Electronics Letters, 2002, 38, (5), p. 246-247.

[8] C. B. Peel, B. M. Hochwald and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," in IEEE Transactions on Communications, vol. 53, no. 1, pp. 195-202, Jan. 2005.

[9] S. Muneer, L. Liu, O. Edfors, H. Sjöland and L. V. der Petre, "Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption," 2020 54th Asilomar Conference on Signals, Systems, and Computers, 2020, pp. 974-978.

[10] C. Mollen, E. G. Larsson, U. Gustavsson, T. Eriksson and R. W. Heath, "Out-of-Band Radiation from Large Antenna Arrays," in IEEE Communications Magazine, vol. 56, no. 4, pp. 196-203, April 2018.

[11] 3GPP TS 36.104: "3rd Generation Partnership Project; Technical Specification Group Radio Access Network; NR; Base Station (BS) radio transmission and reception", Rel. 16, pp. 55-56, June 2020.

Paper III

S. Muneer, H. Sjöland, O. Edfors and L. Liu
Low Complexity Clipping Distortion Compensation for PAPR Reduction in Massive MIMO-OFDM for Frequency Selective Channels
IEEE Transactions on Green Communications and Networking
doi: 10.1109/TGCN.2025.3604969.

Low Complexity Clipping Distortion Compensation for PAPR Reduction in Massive MIMO-OFDM for Frequency Selective Channels

Sidra Muneer, Henrik Sjöland, Ove Edfors and Liang Liu
Department of Electrical and Information Technology, Lund University, Sweden

***Abstract*—Massive multiple input, multiple output (MIMO) systems are crucial for enhancing spectral efficiency and link reliability in next-generation wireless communications. However, high Peak-to-Average Power Ratio (PAPR) remains a significant challenge, leading to power inefficiencies and signal distortion. This paper presents a low complexity method for PAPR reduction using hard clipping tailored for frequency selective massive MIMO systems. The proposed technique involves distortion filtering in frequency and spatial domain, naturally restoring signal quality at the user end while reducing the peaks at the same time. The computational complexity of the proposed technique has been analyzed and compared with the state-of-the-art techniques. Performance analysis in the presence of a nonlinear PA shows an improvement over zero forcing in uncorrelated and correlated channels along with peak reduction, allowing efficient PA operation as well as reduced digital pre-distortion requirements for PAs in Massive MIMO systems.**

***Index Terms*—Massive MIMO, Peak to average power ratio, Power amplifier, Error compensation, Energy efficiency, Frequency selective channels.**

I. Introduction

Peak-to-average power ratio (PAPR) reduction techniques for Orthogonal Frequency Division Multiplexing (OFDM) signals are essential to avoid power amplifier (PA) backoff and improve PA efficiency in base station (BS) radios. In massive antenna systems, optimizing PA efficiency becomes more challenging as the number of PAs increases. Although reduced output power per antenna is beneficial, traditional PAPR reduction methods, along with per antenna digital pre-distortion (DPD) techniques, can significantly increase complexity and power consumption of digital signal processing [1]. Additionally, some practical challenges for PAPR reduction in 5G (fifth generation) and beyond networks also include the inability to use guard bands for nonessential subcarriers (e.g., tone reservation) and restrictions on sending non-standard side information due to stringent standards and regulatory requirements [2]. Therefore, a holistic system-level optimization approach is crucial in balancing PAPR reduction with system performance and complexity. This work thus explores a simplistic and practical method for improving the efficiency of the amplifier by reducing PAPR with minimum performance overhead, which is a critical factor in improving the energy efficiency of future wireless communication networks.

In massive MIMO systems, the availability of a large spatial domain enables techniques such as antenna reservation (AR) [3], which provides low-complexity peak reduction, such as hard clipping, while leveraging the large number of antennas to improve overall system performance. AR involves allocating a subset of antennas within a large antenna array to compensate for the distortion introduced by such techniques. However, reserving antennas reduces the number of antennas available for transmission of the primary signal, thereby diminishing the array gain in a massive antenna system. To meet the received signal to noise ratio (SNR) at the user end, the remaining active transmitting antennas must deliver a higher total power. This approach increases the signal saturation in primary PAs and can thus reduce the system performance, partially negating the benefits of AR. As shown in [4], carefully selecting the number of reserved antennas can improve performance in massive MIMO systems. Nonetheless, the compensation for clipping distortion is limited due to the saturation of PAs.

This paper advances the domain of clipping distortion compensation by proposing a novel approach in which all antennas contribute to error compensation. This eliminates the array-gain loss associated with AR and removes the dependency of compensation signal quality on the number of reserved antennas. Moreover, the method provides a robust low complexity PAPR reduction algorithm for PA efficiency enhancement in Massive MIMO.

This work provides a detailed analysis of complexity and convergence trade-offs of the proposed method compared to the latest optimization and compensation-based approaches in this domain. Our analysis demonstrates that the performance of the proposed method is outstanding in effectively eliminating in-band clipping distortion in the direction of the users. Additionally, simulation results with nonlinear PAs in different channel environments show improvements in system performance and total radiated power (TRP) due to enhanced quality of PA input signal. The performance is also evaluated in more realistic scenarios with channel state information (CSI) error and PA mismatches. A trade-off for convergence in spatially correlated scenario with precoder regularization is also highlighted.

II. Background and Related works

The PAPR problem in the context of OFDM is not new and more than a decade of research has been done on the subject. The major ground development in the domain of PAPR reduction for OFDM based systems has been highlighted in [5]. The least complex techniques in this domain comprise signal

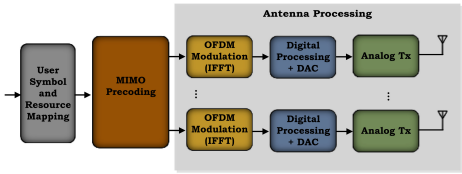


Fig. 1: Simplified block diagram of a downlink multi-user MIMO-OFDM transmitter.

distortion techniques, such as clipping, which provide a trade-off between PAPR reduction and system performance. Some additional work in this domain also involves hybrid techniques, which combine basic methods to achieve a reduction in PAPR as well as an improvement in performance but trade-offs complexity [6]. While higher order constellations combined with distortion-based PAPR reduction methods yield diminishing returns, distortion-less methods are to be avoided because of the high computational complexity.

A. Massive MIMO based PAPR reduction techniques

The primary challenge for PAPR problem in massive MIMO systems is that it arises from traditional precoding methods, regardless of whether OFDM is used [7]. To address this, extensive research has been conducted in recent years to develop PAPR-aware precoding techniques as summarized in [8]. Generally, these techniques tend to achieve a constant envelope per antenna and provide a balance between multi-user interference suppression and PAPR reduction [9]. However, these are predominantly time-domain precoding methods proposed for frequency-flat channels, raising concerns about their practicality for frequency-selective channels.

B. Peak Reduction methods for OFDM based Massive MIMO

In a large scale multi-user MIMO-OFDM systems with frequency domain precoding as shown in Fig.1, the PAPR issue arises after the Inverse Fast Fourier Transform (IFFT) in each antenna processing chain. Subsequent manipulation of the signal severely degrades system performance. Therefore, many optimization-based approaches have been proposed in the literature. In [10], a fast iterative truncation algorithm (FITRA) has been developed to solve a joint optimization problem of precoding, OFDM, and PAPR. Further, a perturbation-assisted optimization approach is proposed in [11] which adds null-space-constrained perturbations in the transmit signal to reduce PAPR while avoiding interference and unwanted radiation. The method offers faster convergence than [10] and can use any precoding technique. An accelerated proximal gradient method (APGM) developed in [12] also solves the joint optimization problem of [10] while offering lower complexity and improved performance and peak reduction. Later, a more improved approach linearized alternating direction method-of-multipliers (LADMM) is presented in [13], which achieves faster convergence and lower complexity over most of the existing optimization-based approaches in the literature. Although most of the above mentioned precoding techniques exploit null space properties, nevertheless inter-user interference performance is limited. Therefore, many practical approaches tend to use distortion-based methods combined with Zero-Forcing (ZF) or other precoders as discussed in [2].

C. Distortion error compensation based methods

In this domain, several works have attempted to reduce the performance degradation of distortion-based methods using extra degrees of freedom available in massive MIMO. A non-iterative AR technique, introduced in [3], mitigates distortion errors caused by clipping. However, as analyzed in [4], this approach involves trade-offs, particularly with nonlinear PAs. Similarly in [14], a peak cancelation (PC) compensation strategy has been deployed which results in lower complexity but requires a predetermined null space vector that has an extensive computation overhead for frequency selective channels. Therefore, the method is not scalable for massive MIMO. In [15], a block-based PC approach is proposed which uses low-dimensional null space in Massive MIMO-OFDM but time domain precoding is not scalable for frequency selective channels. In [16], a non-iterative peak cancellation technique with a similar approach for in-band distortion compensation as [3] is proposed. Additionally, a power controlled perturbation assisted approach is adopted for compensation assuming extra antennas at the transmitter. Similarly in [17], a nonlinear companding based peak reduction and an AR based compensation has been considered which achieves robust performance in the presence of a nonlinear PA. Recently, [18] also explores the optimum spatial arrangement for PC and AR based approach for a uniform planar array. It has been shown that an optimum placement strategy can outperform existing iterative approaches as well. Albeit all, a key drawback for approaches in [3], [16] - [18] is the reduction in array gain, which decreases received signal power at the user position unless compensated by increasing the power in primary antennas. While [4] addresses this by boosting power in the main antennas, it results in increased signal saturation and higher PA power requirements, diminishing the benefits of peak reduction. We therefore expand the horizon of the research in this domain, to include all antennas for the compensation. The proposed technique thus avoids the reduction of the gain of the array and improves the performance in the presence of nonlinear PA.

D. Research Gap

According to the best of authors knowledge, a complexity analysis among recent developments in optimization and distortion-based PAPR reduction methods in OFDM based Massive MIMO systems has not been weighed in the literature. Moreover, performance analysis for such approaches in correlated channel environments has been overlooked in relation to distortion behavior. The practical superiority of distortion based methods has been highlighted in [2], which discusses the implementation constraints for many of the above mentioned methods and motivates the distortion-based approaches.

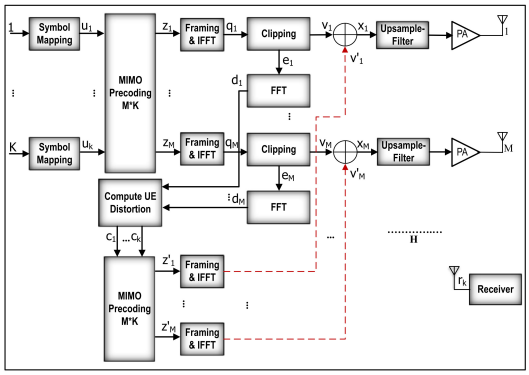


Fig. 2: Functional block diagram of Massive-MIMO downlink base station system model with distortion compensation using all antennas.

E. Contributions

1) We use a low complexity distortion based method to reduce PAPR and exploit the spatial domain of frequency selective channel in Massive MIMO to filter the distortions that exists in the direction of users, hence avoid degrading the performance at user end.
2) Our approach adds the compensation signal in the transmit signal prior to the transmission like [14] and [15]. It therefore avoids array gain loss like [3] and is applicable to frequency selective channels as well (unlike [14] and [15]).
3) As some peak regrowth is bound to occur after addition of compensation signal, the proposed method is hence iterative. We therefore compare it's complexity to the iterative frequency selective approaches of peak reduction for massive MIMO in the literature such as [2] and [13]. The convergence of the algorithm is analyzed for various clipping thresholds, and an optimum clipping strategy in this regard is highlighted.
4) We present the performance of the proposed approach in both uncorrelated and correlated channel environments in the presence of a nonlinear PA, channel state information (CSI) error and PA mismatches.
5) We emphasize the use of distortion based iterative scheme before up-sampling as it offers lower complexity and faster convergence. The signal PAPR statistics are presented for the proposed approach before and after the up-sample filter. The up-sample filter is considered, which provides a good trade-off between peak regrowth and Out-of-band (OOB) radiations before PA.

Section III provides an in-depth analysis of interference cancellation precoding along with clipping and compensation-based PAPR reduction for OFDM based Massive MIMO systems. The analysis shows how distortion due to clipping stemming towards users can be mitigated right at the transmitter using precoding. In Section IV, we compare the complexity of the proposed technique with some recently proposed approaches.

F. Notation

In the following system analysis $\circ$ represents element-wise multiplication of matrices, $\mathbb{F}$ represents discrete Fourier transform matrix, $E[.]$ is an expectation operator, $|.|$ represents absolute value , $\|.\|$ represents the Euclidean norm and $\angle$ represents the phase of the signal. The boldface upper case letters represent matrices and boldface lower case letters represent vectors. Symbol $\mathbb{C}$ denotes complex domain. $\exp(\cdot)$ represent exponential function. $(.)^T$ represents transpose.

III. System Model and Proposed Method

The approach used in this work in general is to reduce PAPR using clipping and then compensate for clipping-induced error/distortion using spatial domain. In this regard, the error compensation signal can be added or assigned to antennas in different ways such as,

1) Compensation precoding and transmission from reserved antennas.
2) Compensation precoding and transmission from the antennas that transmit the least power.
3) Compensation precoding and transmission from all antennas.

The first approach, as presented in [3] and evaluated with nonlinear PA in [4], involves antenna reservation, thereby causing array gain loss. The second approach compensates by utilizing unclipped antennas, adding the compensation signal to those with signal levels below a certain threshold. Thresholds can vary, e.g., antennas with the lowest power or all below mean power. In its simplest form, where all unclipped antennas are used, partial compensation occurs due to limited signal energy from antenna unavailability. The third approach avoids selection complexity by involving all the antennas in compensation signal transmission, therefore eliminating array gain loss due to antenna reservation. This work focuses on the third approach in particular to reduce the complexity of selection as well.

A detailed block diagram of a Massive MIMO system with proposed technique serving K users with M transmit antennas is shown in Fig.2. An IFFT block size of N is assumed, while N_d represents number of available data subcarriers at the transmitter. The components after upsample-filter such as digital to analog converter and frequency converters, are not shown. Unless otherwise stated, we assume a perfect channel state information (CSI) is available at the transmitter. Further for the simplicity of expressions, we assume $N_d = N$ and omit addition of cyclic prefix to combat inter-symbol-interference. The matrix notation used in the system model is defined in Table I. Each block of user data is assumed to be mapped on a constellation symbols, for instance using Quadrature phase shift keying (QPSK) or higher order constellation mapping schemes which constitutes a matrix $\boldsymbol{U}$.

A. Precoding

A precoding function in Fig.2 distribute signal over multitude of frequencies and spatial resources to further process it in individual processing chains to create, composite wideband time-domain OFDM waveforms for the transmission

TABLE I: Matrix Notation used in System Model

Notation	Description
$\boldsymbol{H} \triangleq [\boldsymbol{H}_1, \ldots, \boldsymbol{H}_N] \in (\mathbb{C}^{K \times M})^N$	A block of channel matrices for each subcarrier where each $\boldsymbol{H}_n \in \mathbb{C}^{K \times M}$
$\boldsymbol{W} \triangleq [\boldsymbol{W}_1, \ldots, \boldsymbol{W}_N] \in (\mathbb{C}^{M \times K})^N$	A block of precoding matrices for each subcarrier where each $\boldsymbol{W}_n \in \mathbb{C}^{M \times K}$
$\boldsymbol{P} \triangleq [\boldsymbol{P}_1, \ldots, \boldsymbol{P}_N] \in (\mathbb{C}^{M \times M})^N$	A block of projection matrices orthogonal onto the range of $\boldsymbol{H}$.
$\boldsymbol{U} \triangleq [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N] \in \mathbb{C}^{K \times N}$	A matrix of column vectors containing all user symbols where, $\boldsymbol{u}_n \in \mathbb{C}^{K \times 1}$
$\boldsymbol{Z} \triangleq [\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N] \in \mathbb{C}^{M \times N}$	A discrete domain signal for all subcarriers where each $\boldsymbol{z}_n \in \mathbb{C}^{M \times 1}$.
$\boldsymbol{Q} \triangleq [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_M] \in \mathbb{C}^{N \times M}$	A block of M discrete domain OFDM symbols of length N after IFFT.
$\boldsymbol{V} \triangleq [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_M] \in \mathbb{C}^{N \times M}$	A block of clipped OFDM symbols of all antennas each of length N.
$\boldsymbol{E} \triangleq [\boldsymbol{e}_1, \ldots, \boldsymbol{e}_M] \in \mathbb{C}^{N \times M}$	Difference between the signal before and after clipping.
$\boldsymbol{D} \triangleq [\boldsymbol{d}_1, \ldots, \boldsymbol{d}_N]^T \in \mathbb{C}^{N \times M}$	An error signal in frequency domain for all subcarriers and antennas.
$\boldsymbol{C} \triangleq [\boldsymbol{c}_1, \ldots, \boldsymbol{c}_N] \in \mathbb{C}^{K \times N}$	A compensation signal estimate for each user.
$\boldsymbol{Z'} \triangleq [\boldsymbol{z'}_1, \ldots, \boldsymbol{z'}_N] \in \mathbb{C}^{M \times N}$	A precoded compensation signal where each $\boldsymbol{z'}_n \in \mathbb{C}^{M \times 1}$.
$\boldsymbol{V'} \triangleq [\boldsymbol{v'}_1, \ldots, \boldsymbol{v'}_M] \in \mathbb{C}^{M \times N}$	A compensation signal after IFFT.
$\mathbb{F}$ & $\mathbb{F}^{-1} \in \mathbb{C}^{N \times N}$	Discrete domain Fast Fourier transform/Inverse Fast Fourier transform matrix.

through each antenna. If $\boldsymbol{H}_n \in \mathbb{C}^{K \times M}$ is a downlink channel estimation matrix for a particular subcarrier at the transmitter, a ZF precoder $\boldsymbol{W}_n$ as right Moore-Penrose inverse of a channel matrix $\boldsymbol{H}_n$ is given as,

$$\boldsymbol{W}_n = \boldsymbol{H}_n^H (\boldsymbol{H}_n \boldsymbol{H}_n^H)^{-1} \tag{1}$$

where $\boldsymbol{W}_n \in \mathbb{C}^{M \times K}$ matrix.

B. Orthogonal projections

An orthogonal projector $\boldsymbol{P}_n \in \mathbb{C}^{M \times M}$ can be defined in the range space of $\boldsymbol{H}_n^H$ such that $\boldsymbol{P}_n = \boldsymbol{W}_n \boldsymbol{H}_n$, where $\boldsymbol{P_n}$ is idempotent and Hermitian. This implies that $\boldsymbol{I}_M - \boldsymbol{P}_n$ is an orthogonal projector onto the kernel or null space of $\boldsymbol{H}_n$, where $\boldsymbol{I}_M$ is an identity matrix of dimension M [19]. A projection $\boldsymbol{P}_n$ can be used to extract components of vector $\boldsymbol{d}_n$ that lie onto signal space/range of $\boldsymbol{H}_n$ and projection $\boldsymbol{I}_M - \boldsymbol{P}_n$ can be used to separate components in its null space. This principal is thus used in the method being proposed. The vector $\boldsymbol{d_n}$ can be written as,

$$\boldsymbol{d_n} = \boldsymbol{P}_n \boldsymbol{d}_n + (\boldsymbol{I} - \boldsymbol{P}_n) \boldsymbol{d}_n \tag{2}$$

Note that any error in $\boldsymbol{H}_n$ will impact the projection matrix as it will not be a projection in a true sense, but the duality of the above expression still holds. The proposed technique is therefore robust against any CSI errors or regularization when other system nonlinearities are not present.

C. Frequency selective channels

In frequency selective channel environment, channel $\boldsymbol{H}$ is defined in Table I. A precoder matrix $\boldsymbol{W}$ is also defined where precoder for each subcarrier $\boldsymbol{W}_n$ is given by (1).

Let $\boldsymbol{U}$, $\boldsymbol{V}$ and $\boldsymbol{Q}$ denote matrices containing column vectors as listed in Table I. In Fig.2, a block of symbols $\boldsymbol{U}$ is precoded using a user interference cancellation precoding scheme such as (1). The precoded signal $\boldsymbol{Z}$ is given by,

$$\boldsymbol{Z} = \boldsymbol{W} \circ \boldsymbol{U} \tag{3}$$

The signal $\boldsymbol{z}_m$ at each antenna is then allocated to frames that include guard band intervals in order to combat inter symbol interference (ISI) and then converted to discrete sample domain using IFFT. For simplicity of exchanging dimensions we assumed $N_d = N$ here. The signal $\boldsymbol{Q}$ after IFFT is given as,

$$\boldsymbol{Q} = \mathbb{F}^{-1} \boldsymbol{Z}^T \tag{4}$$

The composite OFDM signal $\boldsymbol{q}_m$ in the discrete sample domain contains very high peaks, resulting in very high PAPR signal. Transmitting this signal directly would require either large PAs to be able to deliver the power of the peaks or a large PA backoff which eventually results in a poor PA efficiency and high system costs. The peaks in the signal must therefore be reduced.

D. Clipping

Clipping is a popular method that works by limiting the signal beyond a certain threshold value to achieve a lower PAPR. The Hard Clipping (HC) for an input signal vector $\boldsymbol{q}$ is given as,

$$\boldsymbol{v}_m = \begin{cases} \boldsymbol{q}_m, \text{ if } |\boldsymbol{q}_m| < T \\ T \exp(j \angle \boldsymbol{q}_m), \text{ otherwise} \end{cases} \tag{5}$$

where, $T = \sqrt{E[\|\boldsymbol{q}_m\|^2] \cdot 10^{0.1\alpha}}$ is a peak signal threshold determined based on the required PAPR (α) of the signal $\boldsymbol{v}$. We further refer to α as the clipping threshold and β as the final required PAPR in this work. Clipping introduces unwanted nonlinear distortion in the signal which degrades performance at the user end as well as disturbing transmission in adjacent frequency bands. Here we specifically highlight its impact on beamforming scenario in spatially correlated environments.

1) Impact of clipping in correlated channels: Recent studies in the behavior of nonlinear distortion radiating from large antenna systems suggest that distortion radiates strongest in the direction of users as number of users $K \to 1$ see [20], [21]. This is because the signal at the multiple antennas combine coherently in a beamforming systems. If transmit signal is spatially correlated any distortion also exhibits correlation. This causes constructive interference of distortion alongside the beam-formed directions. This phenomenon is clearly observable in more correlated channel environments see [22].

As an example, we can observe the OOB distortion stemming from a base station (BS) as a result of clipping. In Fig.3

one such scenario of OOB dispersion in free space line-of-sight (LoS) propagation channel for 3 uniformly spaced users placed at 100m distance from the BS is shown. On left, an OOB radiating from a clipping process in each antenna from a uniform linear array (ULA) situated at (x,y)=$(0, -100)$ is shown. Here, it can be noticed that the OOB distortion is strongest in the direction of users (situated on line $y = 0$). On right in Fig.3, it can be noticed that the received OOB power at user 1 experiences an array gain in a similar way as the main signal.

In general, distortion spreads in more directions than inband signal see [21] and thus experiences slightly lower gain as compared to inband signal. The scaling factor of distortion in such a scenario is highlighted in [22]. This is particularly problematic if a victim user operating in an adjacent frequency band is located in the same direction. Although the distortion (inband or OOB) behavior changes with more number of users and in uncorrelated channel environments, the system must be able to handle the worst case scenario in a best possible way.

The OOB radiation due to clipping can be suppressed via filtering (in exchange of some peak regrowth). We therefore consider post-filtering to control OOB from clipping while we further propose a solution to tackle inband distortion radiating towards users.

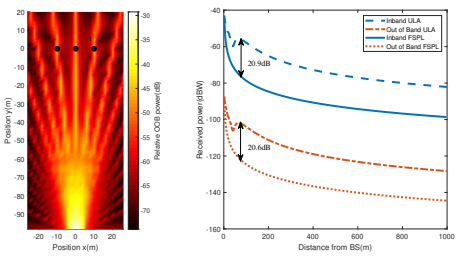


Fig. 3: Heat map of received signal power in adjacent channel/Out of band relative to inband received power at user 1, situated $100m$ away from base station$(-100, 0)$ (left). A 128 element uniform linear array (ULA) with $\lambda/2$ spacing between elements and a Line of Sight free space propagation is assumed. Receive signal power trace in the direction of user 1 (right). Notice that the distortion has a slightly lower gain than the inband signal because distortion signal energy spreads in more directions then the main signal/user directions see [21], [22].

E. Proposed method

In the proposed method, the inband distortion effect of clipping is reduced by error compensation. A distortion error $\boldsymbol{e}_m$is computed by considering the residual of the signal before and after clipping. The error matrix $\boldsymbol{E}$ for all M antennas is therefore,

$$\boldsymbol{E} = \boldsymbol{Q} - \boldsymbol{V}, \tag{6}$$

The error is further converted to frequency domain for compensation signal estimation. The distortion matrix $\boldsymbol{D}$ is then given by,

$$\boldsymbol{D} = \mathbb{F}\boldsymbol{E} \tag{7}$$

Further, user compensation estimate can be determined using,

$$\boldsymbol{C} = \boldsymbol{H} \circ \boldsymbol{D}^T \tag{8}$$

The compensation estimate for each user is further precoded using precoding matrix $\boldsymbol{W}$. Precoded user compensation signal $\boldsymbol{Z'}$ at each antenna is given as,

$$\boldsymbol{Z'} = \boldsymbol{W} \circ \boldsymbol{C} \tag{9}$$

$$\boldsymbol{Z'} = \boldsymbol{W} \circ \boldsymbol{H} \circ \boldsymbol{D}^T = \boldsymbol{P} \circ \boldsymbol{D}^T \tag{10}$$

The compensation signal for each antenna in sample domain after IFFT is given as,

$$\boldsymbol{V'} = \mathbb{F}^{-1}\boldsymbol{Z'}^T \tag{11}$$

The distortion estimate for each antenna is further superimposed on the clipped signal $\boldsymbol{v}_m$. The final transmit signal X is given as,

$$\boldsymbol{X} = \boldsymbol{V} + \boldsymbol{V'} \tag{12}$$

$$\boldsymbol{X} = \boldsymbol{V} + \mathbb{F}^{-1}(\boldsymbol{Z'})^T \tag{13}$$

Analyzing X in frequency domain to determine the transmit signal characteristics.

$$\boldsymbol{X}_f = \mathbb{F}(\boldsymbol{V}) + \boldsymbol{P} \circ \boldsymbol{D} \tag{14}$$

$$\boldsymbol{X}_f = \mathbb{F}(\boldsymbol{V}) + \boldsymbol{P} \circ (\mathbb{F}(\boldsymbol{Q} - \boldsymbol{V})) \tag{15}$$

Note that $\boldsymbol{P}^T = \boldsymbol{P}$. Also element wise vector multiplication is invariant of transpose. Rearranging,

$$\boldsymbol{X}_f = (\boldsymbol{I} - \boldsymbol{P}) \circ \mathbb{F}\boldsymbol{V} + \boldsymbol{P} \circ \mathbb{F}\boldsymbol{Q} \tag{16}$$

which implies that there are two distinct components in signal $\boldsymbol{X}$, a distortion-free signal $\boldsymbol{Q}$ which is directed towards users using orthogonal projection $\boldsymbol{P}$ onto the range of $\boldsymbol{H}$, while the other component is signal $\boldsymbol{V}$ which has direction $(\boldsymbol{I} - \boldsymbol{P})$ which is the null space of $\boldsymbol{H}$. The PAPR reduction comes from the fact that the energy of clipping distortion compensation signal is spatially filtered when mapping from antenna to user domain with the significant energy reaching points where there are no users. Then the estimated user distortion is converted back to antenna domain which results in a compensation signal with less energy and a different time domain waveform than the original clipping distortion. The compensation signal can then be added directly to the signal at each antenna for compensation of the clipping error before transmission. This results in a signal with lower PAPR than the original unclipped signal since the clipping noise is still preserved in other spatial directions. This is depicted in Fig. 4 in compensated signal power $\boldsymbol{x}$ after one iteration at antenna 1. There is some probability that the addition at some antennas and points in time causes the signal to exceed the clipping limit, the process can then be done iteratively (as shown in Fig.4), which can further reduce PAPR without introducing additional in-band distortion at the users. The proposed iterative algorithm is listed below.

Note that the algorithm 1 in contrast to peak cancellation based iterative methods presented in [14], [16] and [15] does not assume a pre-computed null space projection matrix and is applicable to frequency selective channels. In the proposed

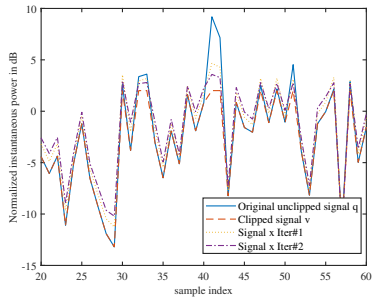


Fig. 4: Normalized instantaneous signal power of few samples around largest peak in an OFDM symbol at antenna 1, showing signal regrowth after each iteration of the proposed algorithm.

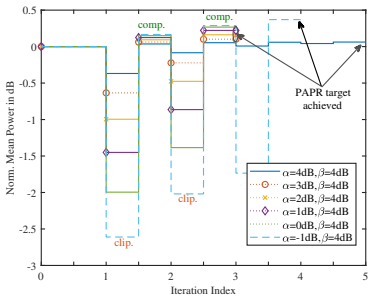


Fig. 5: Normalized mean signal power per antenna, showing power regrowth after each iteration of clipping and compensation for the proposed algorithm for various values of clipping threshold α.

Algorithm 1 Algorithm for PAPR Reduction

Input *Unclipped signal* $\boldsymbol{q}_m$: *Clipping threshold* α, *Target PAPR* β

Output: *Clipped signal* $\boldsymbol{x}_m$

Initialisation : $\boldsymbol{x}_m = \boldsymbol{q}_m$
while *PAPR*$(\boldsymbol{x}_m) > \beta$ **do**
  $\boldsymbol{v} \leftarrow clip(\boldsymbol{x}_m, \alpha)$;
  $\boldsymbol{e}_m \leftarrow \boldsymbol{q}_m - \boldsymbol{v}_m$;
  $\boldsymbol{d}_m \leftarrow fft(\boldsymbol{e}_m)$;
  $\boldsymbol{c}_n \leftarrow \boldsymbol{H}_n \cdot \boldsymbol{d}_m$;
  $\boldsymbol{z'}_m \leftarrow \boldsymbol{W}_n \cdot \boldsymbol{c}_n$;
  $\boldsymbol{v'}_m \leftarrow ifft(\boldsymbol{z'}_m)$;
  $\boldsymbol{x}_m \leftarrow \boldsymbol{v}_m + \boldsymbol{v'}_m$;
end while

technique, the effect of clipping is mitigated at the user positions, but distortion is still emitted to other locations, enabling the waveform of the PAs to be altered to reduce PAPR. The average power per antenna can slightly increase due to this depending upon the clipping threshold α. This is shown in Fig. 5 for various α when the target PAPR level is 4dB. All the processing can be done in digital baseband before upsample-filter, which can suppress the out of band distortion from clipping before the PA. A peak regrowth to some extent from such a process is expected and must be taken into account when evaluating PA efficiency gains.

IV. Complexity Analysis

The proposed method has two major complexity contributing parts. One consists of main signal precoding and IFFT which are non-iterative and the other is iterative as given by algorithm 1. For ZF precoding, the computation of precoder for each data subcarrier as given by eq. (1) contains product $(\boldsymbol{H}\boldsymbol{H}^H)$. This product requires MK^2 complex multiplications. Since the inverse of a square matrix requires K^3 multiplications, a more efficient Cholesky decomposition can be considered which requires only $K^3/2$ complex multiplications [23]. Further multiplication with $\boldsymbol{H}^H$ estimates total complexity of computing ZF precoder per data subcarrier as $2MK^2+K^3/2$. For precoding of user symbols MK, complex multiplications are also required. An IFFT is performed per antenna requiring M times computations in total. For a radix-2 decimation in Time (DIT) implementation of IFFT or FFT of block length N requires $\frac{N}{2}\log_2 N$ complex multiplications.

The total computational complexity of the baseline system without any PAPR reduction is $MN_DK(2K+1) + N_D(K^3/2) + (1/2)MN\log_2 N$. This is worth mentioning here many optimization-based approaches in the literature (e.g. FITRA [10], ADMM [11] etc.) avoid this baseline complexity by offering a standalone solution for mitigating multi-user interference and PAPR reduction. However, the distortion based approaches can achieve interference mitigation through traditional non-iterative techniques like ZF.

The proposed method uses clipping to reduce peaks in the signal. To implement clipping, the power of the time-domain OFDM signal at each antenna is compared to a threshold. This comparison requires N subtractions for each OFDM symbol. The power calculation of N complex valued samples requires $2N/3$ complex multiplications per antenna (considering 3 real multiplications correspond to one complex multiplication). The complexity with FFT and IFFT becomes same as ICF technique ([24]) which approximates to $\mathbb{O}(MNlog_2N)$ complex multiplications.

For the proposed method, in addition to clip and filter, few more steps are involved such as user distortion estimation and precoding. User distortion estimation in eq.(8) requires MN_DK complex multiplications.In addition, a precomputed precoding matrix (corresponding to the same coherence time interval) can be reused so that the precoding complexity is only MN_DK complex multiplications. The total complexity offered by proposed method is thus approximately $\mathbb{O}(2MN_DK + MN\log_2 N)$.

TABLE II: Computational Complexity of Proposed Method in comparison with optimization and distortion based methods ($M >> K$).

Method	Complexity (non-iterative baseline)	Complexity/Iteration (algorithm)[1]	Convergence[2]
LADMM [12]	–	$\mathbb{O}(MN\log_2 N + MN_DK)$	Slow
TOP-ADMM [2]	$\mathbb{O}(MN_DK^2 + N_DK^3 + MN\log_2 N)$	$\mathbb{O}(M^2N_D + MN\log_2 N)$	Medium
Proposed Method	$\mathbb{O}(MN_DK^2 + N_DK^3 + MN\log_2 N)$	$\mathbb{O}(MN_DK + MN(\log_2 N))$	Fast

[1]For a fair comparison up-sampling factor in complexity for the compared methods is not stated.
[2]Convergence is compared for the same system setups as considered in [12] and [2].

The complexity comparison of the proposed method with some recently proposed iterative optimization techniques as well as distortion-based methods is presented in Table II. The optimization based method LADMM [12] solves the joint problem of inter-user interference and PAPR reduction iteratively and outperforms the existing literature in this domain. On the other hand, there are distortion based iterative methods such as TOP-ADMM [2] and PCCNC [14], which only provide solution to PAPR problem and hence involves a non-iterative baseline complexity for precoding, IFFT etc. We do not consider PCCNC [14] here since it is only applicable to frequency flat channels. The total complexity is considered along with the convergence performance as well.

A. Upsampling and convergence

The convergence of the proposed algorithm is greatly impacted by use of up-sampling before or after the PAPR reduction. For a fair comparison, the up-sampling factor must therefore be considered in a similar way as, the method being compared in Table II. For example, for the LADMM setup in [12], an average PAPR of 4dB requires approximately 80 iterations, corresponding to approx. 1.58×10^7 operations. In contrast, the proposed method achieves the same PAPR in approx. 4 iterations, requiring approx. 2.06×10^6 operations.

In general, for the proposed method, up-sampling afterwards can reduce complexity and improve convergence. A more rigorous clipping can then be adopted in order to achieve the post sampling target PAPR. Figure 6 shows the number of iterations required for various clipping thresholds in a traditional base station setup with $M = 128$ antennas, $N = 2048$, $N_D = 1200$ and $K = 10$ users in an uncorrelated channel scenario. Here target PAPR β is kept the same as the clipping threshold α, with an initial average PAPR of approx. 9dB. It can be seen that if the target PAPR is 4dB a more aggressive clipping such as 2dB can achieve the target in lower number of iterations (i.e. 2). This is true in general for higher α values (greater than zero). This indicates that convergence is closely tied to the clipping threshold, suggesting that an optimal strategy is to use a threshold below the target/desired PAPR. Later in Section V, we adopt this strategy as well. We also show further how convergence can be improved in spatially correlated scenarios with fewer number of served users.

V. Simulation and Results

The proposed method in section III has a potential to completely eliminate the inband distortion due to clipping, propagating in the direction of users and can potentially improve power amplifier efficiency by reducing PAPR, while improving the total radiated power (TRP) at the same time.

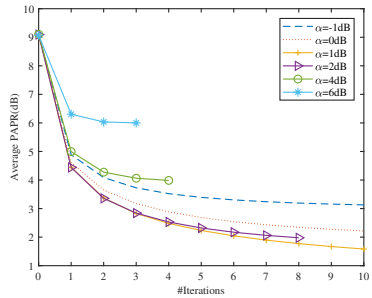


Fig. 6: Convergence of the proposed algorithm for various clipping thresholds in uncorrelated channel environment (K=10, M=128, 64-QAM).

In this section, we present the performance of the proposed method in both correlated and uncorrelated channels. For this matter, we adopt a practical approach to study the system trade-offs such as BER and EVM in the presence of a nonlinear PA.

A fair power allocation among K users and a strict average power constraint per antenna is assumed. For K users, each user gets a power budget of ς/K per OFDM symbol. The energy normalization constant for each user is defined as,

$$\gamma_k = \sqrt{\frac{\varsigma}{K}} \frac{1}{\|\boldsymbol{W}_k\|_F} \tag{17}$$

where $\boldsymbol{W}_k$ is a precoder matrix defined for each user over all subcarriers and antennas.

A. System setup

In the simulation setup, a uniform linear antenna array with 128 antenna elements and $\lambda/2$ spacing between the elements has been considered. An ideal digital to analog converter and frequency conversion of 2.6GHz is assumed to be present in each antenna processing chain along with an ideal or a nonlinear power amplifier. The simulation parameters are summarized in Table III. A unitary basis is being used in symbol mapping and a unit gain amplifiers are assumed. The signal before PA is an up-sampled pulse-shaped filtered version of $\boldsymbol{x}_m$. We consider a 4-fold up-sampling here and a OOB suppressing filter which provides a good trade-off between OOB and peak regrowth. For simulations, 48 OFDM symbols of length 2048 has been used.

PA nonlinear behavior is modeled using modified Rapp model since it approximates PA distortion more accurately closer to saturation region [25]. A low pass complex baseband

TABLE III: Simulation Parameters

# BS antennas (M)	128	# subcarriers	1200
Carrier frequency	2.6GHz	Bandwidth	20MHz
BS to user dist.	≈ 100m	Tx power/ant.	0.05W
Inter-use spacing	20m	Inter-element spacing	0.0577m
Noise power	-174dBm/Hz	# user separation	10m
FFT size (N)	2048	Upsampling factor	4

equivalent response for each amplifier in a base station can be described as,

$$\boldsymbol{y}_m = \boldsymbol{G}(|\boldsymbol{x}_m|) \cdot e^{j(\angle \boldsymbol{x}_m + \boldsymbol{\phi}(|\boldsymbol{x}_m|))} \tag{18}$$

where $\boldsymbol{x}_m$ is low-pass complex input and $\boldsymbol{y}_m$ is the corresponding output from the PA. The amplitude transformation (AM-AM), $\boldsymbol{G(.)}$ is given as,

$$G = \frac{A}{\left(1 + \left(\frac{A}{A_{\text{sat}}}\right)^{2p}\right)^{\frac{1}{2p}}} \tag{19}$$

where A is the input signal amplitude, A_{sat} is the amplifier output saturation point and p controls the smoothness of saturation. The amplitude to phase distortion transformation (AM-PM), $\phi(.)$ is given as,

$$\phi(A) = \phi_{\text{max}} \cdot \frac{\left(\frac{A}{A_{\text{sat}}}\right)^{q}}{1 + \left(\frac{A}{A_{\text{sat}}}\right)^{q}} \tag{20}$$

where, ϕ_{max} is the maximum deviation of the amplifier at saturation point and q is phase smoothness factor. For class B PAs we assume $\phi_{\text{max}} = 15^\circ$ and $q = 2$ [26]. The parameter p is estimated based on $1dBCP$ of the PA.

B. Performance in uncorrelated channel environment

The bit error rate (BER) performance for the proposed method is shown in Fig. 7 for a system with $K = 10$ users and $M = 128$ antennas, in an i.i.d. Rayleigh fading channel with perfect channel state information (CSI). The modulation scheme is 64-QAM (Quadrature amplitude modulation), with a target peak-to-average power ratio (PAPR) set to $4dB$. The achievable PAPR after upsample-filter for each scenario is shown in Fig. 8. The clipping threshold is chosen approx. 0.5dB lower than the target in order to minimize number of iterations. For comparison, the BER of ZF based system without PAPR reduction is also presented. Additionally, the widely used clip-and-filter technique from [24] and a simple non-iterative clipping for the same target PAPR are also evaluated.

Figure 7a demonstrates that the proposed method achieves a BER equivalent to that of the ZF system in an ideal scenario (without any PA nonlinearity or CSI error). The optimization approaches as discussed in the previous sections generally aim to approach the inter-user interference performance of ZF, but the proposed method accomplishes this with fewer iterations for any target PAPR. Furthermore, in Fig. 7b the performance of all the methods under nonlinear PA effects is shown. Notably, the proposed method with PA outperforms the other methods, even improving the BER performance over ZF under the conditions of PA nonlinearity and saturation. Such improvement has a potential to reduce the linearity requirements for the PA, thereby lowering the overall system cost. The enhanced performance can be attributed to the PAPR reduction, which improves the signal quality delivered to the PA, reducing nonlinear distortion emissions. Such techniques are especially desirable in massive antenna systems, where compensating for PA distortion through DPD poses significant challenges in terms of complexity.

Figure 7c shows the impact of CSI error on the performance of the proposed method. The evaluation has been done by introducing random normalized mean square error (NMSE) in channel matrix for the corresponding error variance at the transmitter. It can been seen that since ZF performance is limited in presence of CSI error, the proposed method is also impacted in the similar way as ZF.

Figure 7d shows the the performance of the proposed method with CSI error in the presence of PA nonlinearity. It can be seen that the proposed method still outperforms ZF in the presence of CSI error of for example, 1%. This can be regarded as a common scenario since many channel estimation algorithms can provide average NMSE lower than 0.01 at high SNR [27].

Figure 7e shows the performance for the proposed method in the presence of random mismatches between PA operating points, which is common in practical systems. It can be seen that the PA mismatch of few decibels has a very negligible impact on performance.

Figure 7f shows the the performance of the system in more realistic scenario when both erroneous channel state information as well as PA mismatches are present. It can be seen that the performance limitation of ZF in such scenarios equivalently impact the proposed method as well. One possible improvement can be to consider precoders which are robust against CSI error but this work is limited only to linear precoders.

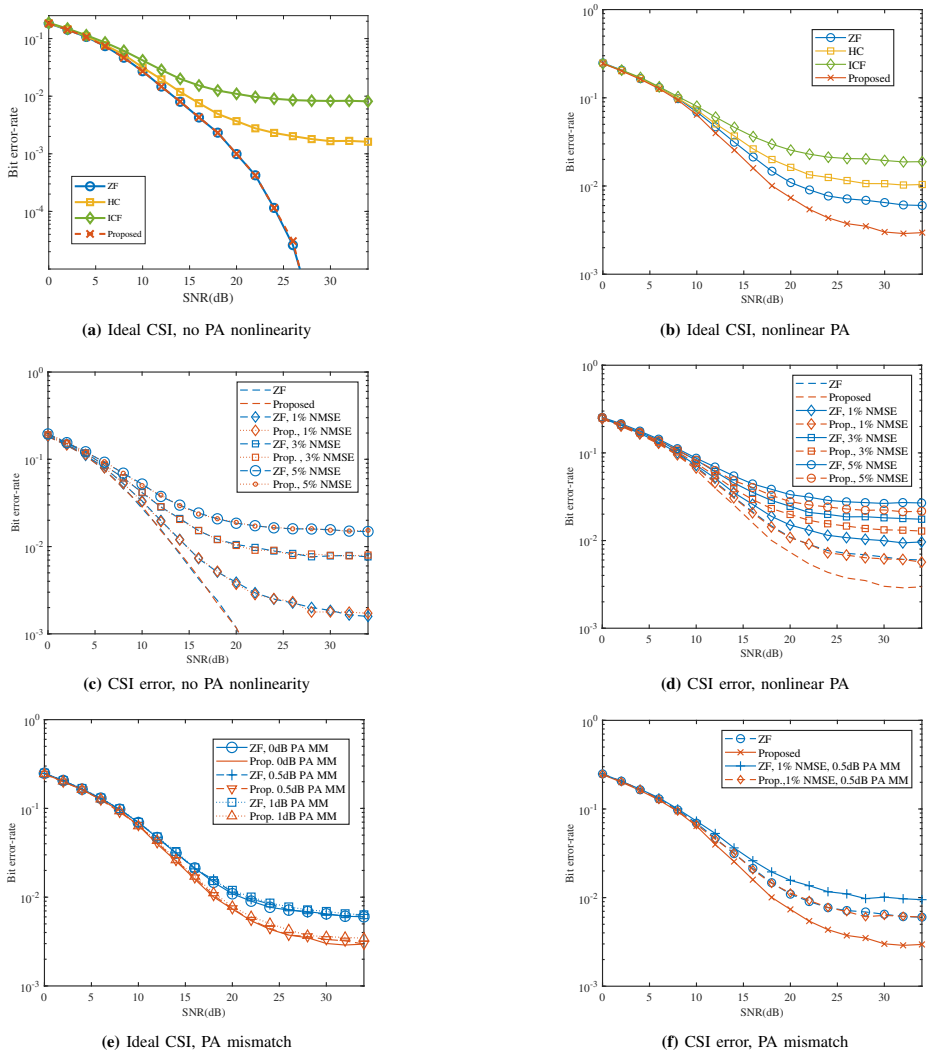
C. Peak reduction behavior

The PAPR reduction behavior can be explained by observing the complementary cumulative distribution function (CCDF) of PAPR before and after up-sampling for all scenarios. Mathematically, $\mathit{CCDF}(\boldsymbol{x}_m) = \mathit{Prob}(\mathit{PAPR} > \boldsymbol{x}_m)$. This is shown in Fig.8 for an average PAPR of 1600 OFDM symbols. The PAPR regrowth after upsample filter is also shown. Notice that the PAPR regrowth for ICF is the minimal, since it involves already involves iterative filtering of out-of-band distortions.

In summary, the PAPR reduction behavior is attributed to many factors. These include clipping threshold, upsample-filter that follows, precoder regularization, and the number of users being served. Many of these are considered further in the discussion of correlated channel environments.

D. Performance in correlated channel environment

In this section, we consider two different user configurations to analyze performance. Performance is evaluated in terms of



(a) Ideal CSI, no PA nonlinearity

(b) Ideal CSI, nonlinear PA

(c) CSI error, no PA nonlinearity

(d) CSI error, nonlinear PA

(e) Ideal CSI, PA mismatch

(f) CSI error, PA mismatch

Fig. 7: BER performance of proposed method in uncorrelated channel environment.

the average percentage of EVM degradation at the user. EVM is computed as,

$$\text{EVM}(\%) = E\left[\sqrt{\frac{(r-u)^2}{u^2}}\right] \times 100 \tag{21}$$

where r and u are received and reference constellation symbols respectively. For a target PAPR of 4dB, an improvement in EVM for a massive MIMO system in the LoS channel scenario with the proposed technique is shown in Table IV. For comparison, performance for ICF, AR and system without any PAPR reduction is also considered. The received power is kept same in all the scenarios with an SNR of 30dB per subcarrier at each receiver. This has been done by reducing the number of antennas from 128 to 58 in 3 user scenario. Reducing antennas is also more compatible with practical deployments where part of the transceivers are turned off to save power when there are very few served users.

In case of AR, performance always depends on the number of reserved antennas [4]. Here, no significant improvement can be achieved for the 10 or 20 reserved antennas, due to the reduction in array gain and an increase in PA saturation in order to maintain the same received power at the receiver. This also results in a degradation of the total radiated OOB

TABLE IV: Analysis of Different Methods in LoS channel.

	$K = 3(M = 58)$		$K = 13(M = 128)$	
Method	**PAPR before PA** (dB)	**EVM(%)**	**PAPR before PA** (dB)	**EVM (%)**
W/O peak reduction	9.4	25.8	9.4	14.19
ICF	5.2	37.5	5.2	23.47
AR with 10 reserved ant.	6.2	27.6	6.2	24.83
AR with 20 reserved ant.	6.2	24	6.2	14.68
Proposed ($\alpha = 2$dB, $\beta = 4$dB)	6.1	**20**	5.75	**10.02**

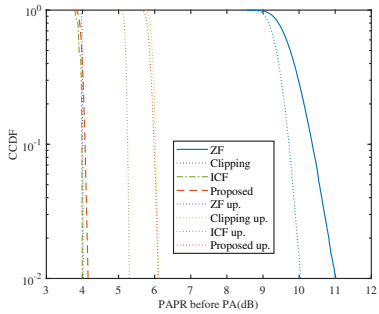


Fig. 8: Achievable average PAPR before and after up-sample filter.

power (TRP) by approximately 1dB-2dB, respectively. For ICF, an EVM can further degrade because of repeated clipping without compensation although OOB distortion is suppressed due to repeated filtering. Note that a different PAPR regrowth is expected for each method after upsample-filter. For the proposed method, a 4dB PAPR target requires approx. 500 and 5 iterations for 3 and 13 user scenario respectively. An improvement in EVM of approximately $4\% - 5\%$ over the baseline system can be observed for each case. It has also been observed that the proposed method improves TRP-OOB by approx. $1\text{dB} - 2\text{dB}$ by improving the input signal quality to the PA.

Note that for $K = 13$ even the baseline system without (w/o) PAPR reduction shows better EVM compared to the $K = 3$. This is because of the change in signal statistics for more number of served users in Massive MIMO as discussed in [21]. Practically in correlated channel scenarios with very fewer served users, besides distortion beam-forming, the convergence of algorithms is also problematic see [2]. Therefore, we discuss this scenario particularly in detail.

E. Convergence in correlated channels

In general for a faster convergence or lower iteration count, a final PAPR meeting criteria can be relaxed compared to the clipping threshold as discussed in Section IV. Convergence is generally challenging in correlated scenarios with fewer users see [2] but this poses a particular challenge in large antenna systems where distortion in case of fewer users is strongly beam-formed in the direction of the user. The proposed technique can significantly improve this by applying regularization in signal precoding as considered in [4]. Figure 9 illustrates this for different user configurations, showing that iterations for 3 users improve markedly with a regularized precoder, albeit with a performance trade-off. The corresponding EVM for each regularization factor is also highlighted. Note that the EVM cost can be further relaxed by using a regularized precoder particularly for the compensation signal precoding.

F. Discussions

The proposed method in general can be applied to wide variety of transmit beamforming architectures, user distributions and channel scenarios where PAPR improvement is crucial in order to improve system efficiency and performance. One such feasible architecture is directly-sampling transmitter presented in [28] and an example channel scenario is highlighted in [29]. Since the proposed method reduces PAPR in order to improve amplifier efficiency, it also improves the distortion stemming from front end components due to very high signal peaks. For example, in resolution constrained systems with limited DAC capabilities such as [30], a reduction in signal PAPR would guarantee a performance improvement as well. Since we also consider a LoS channel scenario where precoder regularization can improve convergence in worst case scenarios with few number of users, the work can benefit mmWave based massive MIMO [29] systems as well.

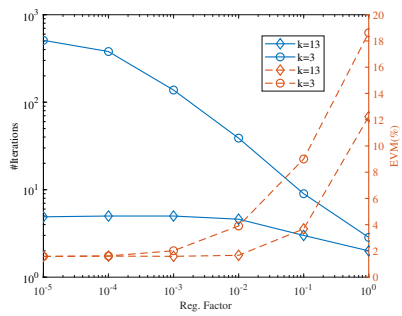


Fig. 9: Convergence of the proposed method using regularized precoder for compensation signal in spatially correlated channels and corresponding EVM trade-off.

VI. Conclusions

In this paper a low complexity clipping for PAPR reduction with a significant contribution towards clipping distortion compensation is proposed. A frequency selective, OFDM based Massive MIMO system has been analyzed. It has been shown through simulation and analysis that the distortion due to signal clipping can be completely mitigated in the direction of users. Although the algorithm is iterative, it

provides faster convergence than the recent optimization and distortion-based iterative approaches proposed in this domain. An optimum clipping strategy in this regard is also highlighted. The performance of the proposed methods has been compared in different channel environments with various state-of-the-art methods and practical system constraints such as CSI errors or PA mismatches. With a larger number of users served, the relative complexity of the proposed method is expected to decrease compared to the baseline system, and the convergence behavior improves. For fewer users, a regularization-based approach in precoding is beneficial in correlated channel environments.

VII. Future Works

Antenna selection for compensation transmission from the antennas that transmit the least power has been excluded for future investigations. For selection, different criteria can include fraction of time where amplitude in an antenna exceeds a certain threshold, i.e. its PA being in or near compression, and root mean square (RMS) amplitude. The selection process may involve choosing all antennas below a specified threshold, selecting the antennas with the lowest values based on the selection criteria, or using a combination approach. For instance, selecting all antennas below a threshold, but if insufficient, supplementing with a fixed number of antennas. However, the analysis in this work is restricted to transmission utilizing all antennas.

VIII. Acknowledgments

The authors thank Ericsson AB for the research support under Grant 'Massive MIMO Technology and applications' and ELLIIT project B2 "Baseband Processing for Beyond 5G Wireless".

References

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, pp. 186–195, 2 2014.

[2] S. Kant, M. Bengtsson, G. Fodor, B. Goransson, and C. Fischione, "EVM Mitigation With PAPR and ACLR Constraints in Large-Scale MIMO-OFDM Using TOP-ADMM," *IEEE Transactions on Wireless Communications*, vol. 21, pp. 9460–9481, 11 2022.

[3] H. Prabhu, O. Edfors, J. Rodrigues, L. Liu, and F. Rusek, "A low-complex peak-to-average power reduction scheme for OFDM based massive MIMO systems," in *2014 6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*. IEEE, 5 2014, pp. 114–117.

[4] S. Muneer, J. R. Sanchez, L. V. der Perre, O. Edfors, H. Sjoland, and L. Liu, "System Design and Performance for Antenna Reservation in Massive MIMO," in *2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. IEEE, 9 2022, pp. 1–5.

[5] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in ofdm systems: A survey and taxonomy," *IEEE Communications Surveys and Tutorials*, vol. 15, pp. 1567–1592, 2013.

[6] F. Sandoval, G. Poitau, and F. Gagnon, "Hybrid Peak-to-Average Power Ratio Reduction Techniques: Review and Performance Comparison," *IEEE Access*, vol. 5, pp. 27 145–27 161, 11 2017.

[7] C. Mollén, E. G. Larsson, and T. Eriksson, "Waveforms for the Massive MIMO Downlink: Amplifier Efficiency, Distortion, and Performance," *IEEE Transactions on Communications*, vol. 64, pp. 5050–5063, 12 2016.

[8] M. A. Albreem, A. H. Habbash, A. M. Abu-Hudrouss, and S. S. Ikki, "Overview of Precoding Techniques for Massive MIMO," *IEEE Access*, vol. 9, pp. 60 764–60 801, 2021.

[9] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Transactions on Communications*, vol. 61, pp. 1059–1071, 2013.

[10] C. Studer and E. G. Larsson, "PAR-aware large-scale multi-user MIMO-OFDM downlink," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 303–313, 2013.

[11] H. Bao, J. Fang, Q. Wan, Z. Chen, and T. Jiang, "An ADMM Approach for PAPR Reduction for Large-Scale MIMO-OFDM Systems," *IEEE Transactions on Vehicular Technology*, vol. 67, pp. 7407–7418, 8 2018.

[12] L. Hua, Y. Wang, Z. Lian, Y. Su, and Z. Xie, "Low-Complexity PAPR-Aware Precoding for Massive MIMO-OFDM Downlink Systems," *IEEE Wireless Communications Letters*, vol. 11, pp. 1339–1343, 7 2022.

[13] L. Hua, Y. Wang, Z. Lian, Y. Su, and Z. Xie, "LADMM-Based PAPR-Aware Precoding for Massive MIMO-OFDM Downlink Systems," *IEEE Transactions on Vehicular Technology*, vol. 72, pp. 735–746, 1 2023.

[14] T. Suzuki, M. Suzuki, Kishiyama, and K. Higuchi, "Complexity-reduced adaptive PAPR reduction method using null space in MIMO channel for MIMO-OFDM signals," in *IEICE Transactions on Communications*, vol. E103B. Institute of Electronics, Information and Communication, Engineers, IEICE, 9 2020, pp. 1019–1029.

[15] J. Saito, T. Hara, and K. Higuchi, "Peak Cancellation Signal-Based Parallel PAPR Reduction Method Using Low-Dimensional Null Space in Massive MIMO-OFDM," in *2024 International Conference on Computing, Networking and Communications, ICNC 2024*. Institute of Electrical and Electronics Engineers Inc., 2024, pp. 361–365.

[16] T. Kageyama, O. Muta, and H. Gacanin, "Enhanced Peak Cancellation with Simplified In-Band Distortion Compensation for Massive MIMO-OFDM," *IEEE Access*, vol. 8, pp. 73 420–73 431, 2020.

[17] M. Q. Khan, M. D. Nisar, and H. A. Khan, "On PAPR Reduction and Transmit Distortion Compensation in Massive MIMO OFDM Systems," in *IEEE Vehicular Technology Conference*, vol. 2021-April. Institute of Electrical and Electronics Engineers Inc., 4 2021.

[18] Z. Li, R. Nojima, O. Muta, M. Omoto, T. Makita, and K. Maruta, "In-band Distortion Compensation for Peak Cancellation Based PAPR Reduction in Massive MIMO-OFDM with Uniform Planar Array," *IEEE Access*, 2024.

[19] Sudipto and A. B. Roy, *Linear algebra and matrix analysis for statistics*. Crc Press Boca Raton, 2014.

[20] E. G. Larsson and L. V. D. Perre, "Out-of-Band Radiation From Antenna Arrays Clarified," *IEEE Wireless Communications Letters*, vol. 7, pp. 610–613, 8 2018.

[21] S. Muneer, L. Liu, O. Edfors, H. Sjoland, and L. V. der Petre, "Handling PA Nonlinearity in Massive MIMO: What are the Tradeoffs Between System Capacity and Power Consumption," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*. IEEE, 11 2020, pp. 974–978.

[22] C. Mollen, E. G. Larsson, U. Gustavsson, T. Eriksson, and R. W. Heath, "Out-of-Band Radiation from Large Antenna Arrays," *IEEE Communications Magazine*, vol. 56, pp. 196–203, 4 2018.

[23] A. Krishnamoorthy and M. Deepak, *Matrix inversion using Cholesky decomposition*. IEEE, 2013, pp. 70–72.

[24] J. Armstrong, "Peak-to-average power reduction for OFDM by repeated clipping and frequency domain filtering," *Electronics Letters*, vol. 38, pp. 246–247, 2 2002.

[25] 3rd Generation Partnership Project (3GPP), "On power amplifier model for the New Radio," 3GPP TSG-RAN WG4 Meeting #80, R4-165256, Aug. 2016, Gothenburg, Sweden, 22-26 Aug, 2016. [Online]. Available: https://www.3gpp.org/release-16.

[26] ZTE Corporation, "Discussion on PA Model," 3GPP TSG-RAN WG4 Meeting #80, R4-165569, Aug. 2016, Gothenburg, Sweden, 22-26 Aug, 2016.

[27] Y. Wu, Y. Gu, and Z. Wang, "Channel Estimation for mmWave MIMO With Transmitter Hardware Impairments," *IEEE Communications Letters*, vol. 22, no. 2, pp. 320–323, 2018.

[28] A. S. Bora, T. H. Singh, and P. T. Huang, "Digi-FH-OFDM: An all-digital wideband frequency-hopped OFDM system," *Physical Communication*, vol. 52, 6 2022.

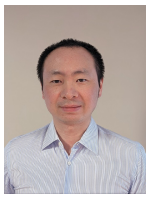
[29] P. Priya and D. Sen, "Particle Filter Based Nonlinear Data Detection in Presence of CFO for Frequency Selective mmWave MIMO-OFDM Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 5892–5907, 6 2021.

[30] S. Jacobsson, L. Aabel, M. Coldrey, I. C. Sezgin, C. Fager, G. Durisi, and C. Studer, "Massive MU-MIMO-OFDM Uplink with Direct RF-Sampling and 1-Bit ADCs," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–6.

Sidra Muneer (Graduate Student Member, IEEE) received her bachelor's degree in Telecommunication engineering (summa cum laude) from FAST University, Islamabad, Pakistan, in 2009, and her master's degree in Computer engineering from Mid Sweden University, Sundsvall, Sweden in 2013. She is currently part of department of Electrical and Information technology at Lund University, Lund, Sweden working towards her Ph.D. Her current research interests includes low complexity efficiency enhancement and linearization techniques for future wireless systems.

Henrik Sjöland (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree from Lund University, Lund, Sweden, in 1994 and 1997, respectively. In 1999, he was a Postdoctoral Researcher with the University of California at Los Angeles (UCLA), Los Angeles, CA, USA, on a Fulbright Scholarship. He has been an Associate Professor with Lund University since 2000 and a Full Professor since 2008. Since 2002, he has been part-time employed at Ericsson Research, Lund, where he is currently a Senior Specialist. He has authored or coauthored more than 180 international peer-reviewed journal articles and conference papers. He holds patents on more than 30 different inventions. He has successfully been the main supervisor of 14 Ph.D. students to receive their degrees. His research interests include the design of radio frequency, microwave, and mm-wave integrated circuits, primarily in CMOS technology. Dr. Sjöland has been a member of the Technical Program Committee of the European Solid-State Circuits Conference (ESSCIRC). He has previously been an Associate Editor of IEEE Transactions on Circuits and Systems II. He is also an Associate Editor of the IEEE Transactions on Circuits and Systems I.

Ove Edfors (Senior Member, IEEE) received M.Sc. degree in computer science and engineering and Ph.D. degree in signal processing from Luleå University of Technology, Sweden, in 1990 and 1996. Currently, he is a professor of radio systems with the Department of Electrical and Information Technology, Lund University, Lund, Sweden. His research interests include statistical signal processing and low-complexity algorithms with applications in wireless communications, and in the context of massive MIMO and large intelligent surfaces, his main research focuses on how realistic propagation characteristics influence system performance and baseband processing complexity.



Liang Liu (Member, IEEE) received his Ph.D. degree in microelectronics from Fudan University, Shanghai, China, in 2010. He joined Lund University as a Post-doc in 2010 and was Assistant and Associate Professor 2014-2024. In 2015, He received Docent. Since 2024, he is Professor at the Electrical and Information Technology (EIT) Department, Lund University and head of the Integrated Electronics Systems (IES) division. His research interest includes wireless communication system and digital integrated circuits design. Liang is active in several EU and Sweden projects is currently a Member of the Technical Committee of VLSI Systems and Applications and CAS for Communications of the IEEE Circuit and Systems Society.